

Joint Influence of Visual and Auditory Words in the Stroop Task

Wendy S. Francis¹ · Colin M. MacLeod² · Randolph S. Taylor¹

Published online: 14 October 2016
© The Psychonomic Society, Inc. 2016

Abstract We conducted four Stroop color–word experiments to examine how multiple stimuli influence interference. Experiments 1a and 1b showed that interference was strong when the word and color were integrated, and that visual and auditory words made independent contributions to interference when these words had different meanings. Experiments 2 and 3 confirmed this pattern when the word information and color information were not integrated, and hence when overall interference was substantially less. Auditory and visual interference effects are comparable except when the visual distracter is integrated with the color, in which case interference is substantially enhanced. Overall, these results are interpreted as being most consistent with a joint influence account of interference as opposed to a capture account.

Keywords Selective attention · Interference · Stroop effect · Modality

There are a great many situations in our experience where irrelevant information disrupts our processing of relevant information—the phenomenon known as cognitive interference. None, however, is more familiar than the classic Stroop color–word interference effect (Stroop, 1935). Incongruent color words cause robust interference in naming print colors when those words are presented visually. In fact, incongruent color words also interfere with color naming

when the words are presented auditorily (e.g., Cowan, 1989; Elliott, Cowan, & Valle-Inclan, 1998; Elliott et al., 2014; Roelofs, 2005). To say that the visual Stroop color–word task has been extensively studied would be a major understatement (see MacLeod, 1991, for an extensive review), but auditory versions have been much less thoroughly explored (see, e.g., Green & Barber, 1981; McClain, 1983), and there has been little work on cross-modal interference, the subject of this article.

Thackray and Jones (1971) initially reported that cross-modal interference did not occur between auditory and visual stimuli in the Stroop task. However, their use of a manual response, where participants press buttons for colors, likely limited interference overall (see MacLeod, 1991). Cowan and Barron (1987) instead used the traditional oral response and the classic multiple color–word stimuli per card version of the Stroop task. They showed that, relative to auditory controls (hearing letters of the alphabet or the word *the* presented repeatedly), hearing an auditory recording of random color words interfered with naming print colors. They also suggested that the interfering effects of auditory and visual distracter words were additive. This cross-modal auditory interference effect has since been replicated using trial-by-trial manipulations of word–color congruency (e.g., Elliot et al., 1998; Roelofs, 2005), more analogous to the standard visual procedure, but auditory interference was not directly compared to visual interference.

Most recently, Elliott et al. (2014) used the trial-by-trial version of the Stroop task with vocal responding. Their controls were noncolor words and a silent condition. Contrary to Cowan and Barron (1987), they observed that combining the two modalities did not increment interference (see also Miles, Madden, & Jones, 1989). Given this empirical conflict, the relative influences of auditory and visual distracters remain unclear. Our first empirical goal, therefore, was to determine how these two modality-specific effects combine.

✉ Wendy S. Francis
wfrancis@utep.edu

¹ Department of Psychology, University of Texas at El Paso, 500 W. University Ave, El Paso, TX 79968, USA

² University of Waterloo, Waterloo, ON N2L 3G1, Canada

Combined Effects of Multiple Sources of Interference

Only a few studies have investigated whether presenting two incongruent visual distracters increases interference relative to presenting only one incongruent visual distracter, or have compared the effects of pairs of visual distracters that were consistent or inconsistent with each other. Moreover, these studies have differed in the paradigms used for presenting the second stimulus and whether the distracters were integrated with or separate from the color stimulus. Generally, no differences have been observed (Kahneman & Chajczyk, 1983; MacLeod & Bors, 2002; MacLeod & Hodder, 1998; Yee & Hunt, 1991); that is, interference has not been greater for multiple interfering visual words. Understanding the degree to which multiple incongruent stimuli—within modality or across modalities—influence interference is our second empirical goal of this research, growing out of our first goal.

The apparent lack of increased interference from a second distracter, or from having two inconsistent rather than two consistent distracters, was first explained using a *capture* account. The idea, developed in the context of visual attention, was that whichever word was attended first captured attention and the other was not processed sufficiently to produce additional interference (Kahneman & Chajczyk, 1983). By mixing incongruent and congruent word stimuli, however, MacLeod and Bors (2002) showed that the second word did have an influence. Specifically, if one word was congruent and the other was incongruent, the interfering effect of the incongruent word was reduced, thus calling the capture account into question. However, there still was no difference in the interference observed for two inconsistent versus two consistent incongruent visual distracters. The alternative explanation offered by MacLeod and Bors was that there is *joint influence* of the two distracters, that the “word dimension continues to be monitored during the attempt to identify and produce the name of the color” (p. 789). More recent work also leads to the conclusion that attention shifting can occur within Stroop stimuli (Cho, Choi, & Proctor, 2012; Cho, Lien, & Proctor, 2006).

We set out in this study to determine whether having incongruent color word inputs from both the auditory and visual modalities would lead to greater interference than having incongruent input from only one modality, or whether there would be a pattern of joint influence. One possible reason to expect a pattern of joint influence would be that visual and auditory distracters might influence color naming at different stages. For example, unlike visual distracters, auditory distracters might have their primary effect at the level of retrieving the phonology of the response word. Another possible reason to expect joint influence would be that auditory and visual distracters may tap different pools of modality-specific

attentional resources (Wickens, 1984), allowing both modalities to influence responding.

Only a couple of previous studies have combined auditory and visual distracters. Cowan and Barron (1987) showed that playing a recording of random color words interfered with naming print colors of a mixture of congruent and incongruent color words. The magnitude of the visual effect was larger than that of the auditory effect. However, because incongruent and congruent auditory color words were mixed, the interfering effect of the auditory distracters was likely underestimated. Also, because the distracting words were presented in a continuing stream, there could be no separation of trial types—particular combinations of auditory and visual color words could not be examined. Nevertheless, these results, along with the lack of interaction between the presence of the auditory recording and the congruency of the printed color words, suggest that auditory and visual inputs together cause more interference than either alone. Elliot et al. (1998) extended these results, reaching a similar conclusion. In contrast, as mentioned, Elliott et al. (2014) recently presented findings using the single trial Stroop method that suggested no enhanced interference when distracter words were presented in both modalities compared to only one modality. Our goals are to resolve this empirical discrepancy at the same time as we tackle the capture vs joint influence theoretical distinction.

The Present Study

The experiments that we report all used large sample sizes and quite large numbers of trials per condition to optimize power to address the critical questions. Experiments 1a and 1b compared interference effects for auditory distracters, visual distracters, and combinations of auditory and visual distracters that were either congruent or incongruent with the target response and with each other. In both of these experiments, visual distracters were integrated with the target color information, as in the classic Stroop task, and Experiment 1b was a direct replication of Experiment 1a. Experiment 2 was a replication of Experiment 1, but with visual distracters that were spatially separated from the target color stimulus, a change well established to sharply reduce overall interference (see, e.g., Kahneman & Chajczyk, 1983; MacLeod, 1998). Experiment 3 involved only visual distracters and examined whether spatially integrated and spatially separated distracters would produce nonredundant interference. In all cases, the capture account predicts that combined conditions should produce interference comparable to that of one of the distracters (the one that captures): There should be no increment from one to two distracters. In contrast, the joint influence account predicts that two incongruent distracters should produce more interference than one incongruent distracter. Similarly, the

capture account predicts that two inconsistent incongruent distracters should produce equivalent interference to two consistent incongruent distracters, whereas the joint influence account predicts stronger interference for two inconsistent incongruent distracters, each of which would make a separate contribution to overall interference.

Experiment 1

The purpose of Experiment 1 was to compare interference effects for auditory and visual distracters and to determine whether their effects were or were not redundant. Participants performed color naming with incongruent auditory, visual, or both auditory and visual stimuli that were the same or different from each other in meaning. They also named colors with congruent auditory, visual, or both auditory and visual stimuli. In all cases, the visual words were integrated with the to-be-named color information and responding was vocal: Both of these features follow the standard Stroop procedure (see MacLeod, 1991, for a review).

Researchers have often debated the best control condition in a Stroop experiment (see MacLeod, 1991; Melara & Mounts, 1993). Should one use meaningless symbols, as Stroop (1935) himself did, or is it more appropriate to use noncolor words, which require reading but without specific competition (see also Klein, 1964, for consideration of multiple possible comparison conditions)? The answer in experimental psychology is always that there is no control condition that is perfect regardless of the question being asked. Here, we have chosen to use an auditory tone, the best analogy we could make in the auditory modality to a meaningless string of symbols in the visual modality. While this control condition might be seen as overestimating interference, it certainly can be argued that it does not itself have the potential to interfere with the color naming response, making it a baseline if not a perfect control. We chose this simple, “no information” control to align with the bulk of the Stroop literature.

Experiment 1a

Method

Participants Participants were 50 students (10 men, 40 women) from the University of Texas at El Paso whose primary language was English (i.e., they reported speaking no other languages fluently). Their median age was 19 years. Most were recruited from introductory psychology classes and received credit toward a class research requirement; the rest were unpaid volunteers.

Apparatus Visual stimuli were presented on the monitor of a Macintosh computer. Sequencing of stimuli and timing of responses were controlled using PsyScope software (Cohen, MacWhinney, Flatt, & Provost, 1993). Auditory stimuli were presented via headphones. Vocal response times were measured using a high-impedance microphone attached to a PsyScope button box (New Micros, Dallas).

Stimuli Visual word stimuli were the color words red, blue, green, and purple; the control visual stimulus was xxxxx. The visual stimuli served as the color carriers (i.e., the color was integrated with the visual stimulus). Auditory stimuli were the same color words recorded by a native English speaker. The control sound was an unramped tone (middle C, 262 Hz) generated by the computer at approximately the same volume as the auditory word stimuli and lasting for 500 ms.

Design The eight trial types are illustrated in Table 1. For each trial type, there were 48 trials, with every possible color combination used equally often, resulting in 384 total trials. It is important to note that there were two types of “incongruent both”: The auditory and visual word could be the same or they could be different.

Procedure Participants were tested individually in sessions lasting approximately 30 minutes. They were instructed to name aloud into the microphone the print color of each visual stimulus as quickly and accurately as possible. The visual stimulus on each trial appeared in lowercase letters at the center of the screen in Arial 36-point font. On each trial, the auditory and visual stimuli were presented simultaneously, and the visual stimulus remained on the screen until a vocal response was registered. The intertrial interval was 1,250 ms. After 10 practice trials, the 384 experimental trials were presented in a random sequence, with a rest after every 96 trials. The experimenter recorded unexpected responses and voice relay misfires on a worksheet preprinted with the expected responses so that these trials could be removed prior to analysis.

Results

Only correct and valid responses were included in the RT analysis. Trials with response errors or voice relay errors were excluded, as were trials with outlier RTs (over 3,000 ms or under 300 ms); 97.5 % of trials were retained for analysis. For each participant, a mean RT was calculated for each condition, and these were the scores that were submitted to analysis. Interference (facilitation) scores were obtained for each participant by subtracting the RT of the control condition from the RT of each of the other conditions. Thus, a positive score indicates interference whereas a negative score indicates

Table 1 Experiments 1a and 1b: Response Times and Interference Scores as a Function of Condition

Condition	Stimulus Examples			Experiment 1a				Experiment 1b			
	Auditory Stimulus	Visual Stimulus	Print Color	RT (ms)		Interference (ms)		RT (ms)		Interference (ms)	
				Mean	(SE)	Mean	(SE)	Mean	(SE)	Mean	(SE)
<i>Control</i>	<tone>	xxxxx	Blue	680	(13)	–	–	758	(18)	–	–
<i>Congruent</i>											
Visual	<tone>	blue	Blue	692	(16)	12	(6)	781	(21)	22	(6)
Auditory	blue	xxxxx	Blue	715	(13)	35	(5)	779	(17)	21	(5)
Visual + Auditory	blue	blue	Blue	750	(16)	69	(7)	838	(20)	80	(7)
<i>Incongruent</i>											
Visual	<tone>	red	Blue	802	(16)	122	(7)	872	(20)	114	(8)
Auditory	red	xxxxx	Blue	739	(16)	58	(7)	804	(17)	46	(5)
Vis + Aud Same	red	red	Blue	834	(16)	153	(8)	878	(17)	120	(8)
Vis + Aud Diff	red	green	Blue	855	(19)	174	(10)	901	(19)	143	(10)

Each cell is based on a maximum of 48 trials (before elimination of error and equipment-failure trials).

facilitation. The principal data, the response times and the interference/facilitation scores derived from those response times, are displayed in Table 1. Table 4 displays the error rates and interference/facilitation scores derived from those error rates; their corresponding analyses are presented in the Appendix. Patterns evident in the response times typically were also apparent in the error rates; there was no evidence of speed–accuracy tradeoff.

Congruent conditions None of the congruent conditions showed facilitation; rather, all actually showed interference despite the word–color congruence ($ps \leq .05$). A 2 (auditory word present or absent) \times 2 (visual word present or absent) repeated-measures ANOVA showed that, relative to the control condition, interference increased due to the presence of a word whether that word was visual, $F(1, 49) = 18.67$, $MSE = 1,435$, $p < .001$, partial $\eta^2 = .28$, or auditory, $F(1, 49) = 118.75$, $MSE = 894$, $p < .001$, partial $\eta^2 = .71$. The interfering effects of congruent visual and auditory words interacted, $F(1, 49) = 14.86$, $MSE = 420$, $p < .001$, partial $\eta^2 = .23$, such that their combined effect was greater than the sum of their individual effects. The unanticipated interference on the congruent trials was part of the motivation for carrying out a faithful replication of Experiment 1a—Experiment 1b—reported below.

Incongruent conditions Interference was significant in each of the four incongruent conditions relative to the control condition ($ps < .001$). To test whether the effects of interference from the two modalities interacted or were independent, two 2 (visual word present or absent) \times 2 (auditory word present or absent) repeated-measures ANOVAs were conducted. These analyses differed only in which visual + auditory condition (same or different) was included. For the analysis including

the visual + auditory same condition, relative to the control condition, the presence of a word increased interference both when the word was visual, $F(1, 49) = 388.81$, $MSE = 1,512$, $p < .001$, partial $\eta^2 = .89$, and when the word was auditory, $F(1, 49) = 84.52$, $MSE = 1,188$, $p < .001$, partial $\eta^2 = .63$. In this case, the interaction of the auditory and visual words, $F(1, 49) = 11.28$, $MSE = 786$, $p = .002$, partial $\eta^2 = .19$, indicated that the combined effect was less than the sum of the individual effects. For the analysis including the visual + auditory different condition, again the presence of a word increased interference whether that word was visual, $F(1, 49) = 388.02$, $MSE = 1,821$, $p < .001$, partial $\eta^2 = .89$, or auditory, $F(1, 49) = 92.51$, $MSE = 1,650$, $p < .001$, partial $\eta^2 = .65$. In this case, auditory and visual words did not interact, $F < 1$, partial $\eta^2 = .02$, indicating that, when visual and auditory words have different meanings, their effects are independent.

Experiment 1b

Experiment 1b was a direct replication of Experiment 1a conducted with a new participant sample and using new equipment. At issue was the unexpected appearance of interference in the congruent conditions, but we also sought to confirm the entire pattern observed in Experiment 1a.

Method

Participants Participants were 50 students (14 men, 36 women) from the University of Texas at El Paso who were fluent speakers of English. These individuals were from the same source as Experiment 1a, but in a subsequent academic year. Their median age was 19 years. Most were recruited from

introductory psychology classes and received credit toward a class research requirement; the rest were unpaid volunteers.

Apparatus Visual stimuli were presented on the monitor of an iMac desktop computer. Sequencing of stimuli and timing of responses were controlled using PsyScope X software (Cohen et al., 1993). Auditory stimuli were presented on headphones. Vocal response times were measured using a high-impedance microphone attached to an ioLabs button box.

Design, materials, and procedure These all were identical to those of Experiment 1a.

Results

Only correct and valid responses were included in the RT analysis. Trials with response errors or voice relay errors were excluded, as were trials with outlier RTs (over 3,000 ms or under 300 ms); 98.5 % of trials were retained for analysis. The principal data, the response times and the interference/facilitation scores derived from those response times, are displayed in Table 1. Error rates and interference/facilitation scores derived from those error rates are displayed in Table 4, with the corresponding analyses presented in the Appendix. Patterns evident in the response times were typically also apparent in the error rates; there was no evidence of speed–accuracy tradeoff.

Response times were about 70 ms slower overall in Experiment 1b than in Experiment 1a, a change that we attribute to two factors: (1) there was no language restriction in Experiment 1b, so participants who were bilingual were included, and (2) participants seemed to place a somewhat higher value on accuracy in Experiment 1b than in Experiment 1a. This overall slower responding did not, however, result in any substantial change in the pattern of response times across the eight conditions with respect to Experiment 1a, as Table 1 shows.

Congruent conditions None of the congruent conditions showed facilitation; rather, all actually showed interference ($p < .001$). A 2 (auditory word present or absent) \times 2 (visual word present or absent) repeated-measures ANOVA showed that, relative to the control condition, the presence of a visual word increased interference, $F(1, 49) = 59.94$, $MSE = 1,366$, $p < .001$, partial $\eta^2 = .55$, as did the presence of an auditory word, $F(1, 49) = 56.04$, $MSE = 1,377$, $p < .001$, partial $\eta^2 = .53$. The interfering effects of congruent visual and auditory words interacted, $F(1, 49) = 29.84$, $MSE = 544$, $p < .001$, partial $\eta^2 = .38$, such that the combined effect was greater than the sum of the individual effects. Essentially, the pattern that was unexpected in Experiment 1a was replicated in Experiment 1b.

Incongruent conditions Interference was significant in each of the four incongruent conditions ($p < .001$). To test whether

the effects of interference from the two modalities interacted or were independent, the same 2 (visual word present or absent) \times 2 (auditory word present or absent) repeated-measures ANOVAs were conducted as in Experiment 1a. For the analysis including the visual + auditory same condition, the presence of a word increased interference relative to the control condition both when the word was visual, $F(1, 49) = 176.30$, $MSE = 2,484$, $p < .001$, partial $\eta^2 = .78$, and when it was auditory, $F(1, 49) = 36.54$, $MSE = 949$, $p < .001$, partial $\eta^2 = .43$. In this case, the interaction of the auditory and visual words, $F(1, 49) = 38.91$, $MSE = 511$, $p < .001$, partial $\eta^2 = .44$, indicated that the combined effect was less than the sum of the individual effects. For the analysis including the visual + auditory different condition, the presence of a word increased interference both when the word was visual, $F(1, 49) = 185.42$, $MSE = 2,988$, $p < .001$, partial $\eta^2 = .79$, and when it was auditory, $F(1, 49) = 62.67$, $MSE = 1,153$, $p < .001$, partial $\eta^2 = .56$. The interaction of the auditory and visual words, $F(1, 49) = 8.27$, $MSE = 413$, $p = .006$, partial $\eta^2 = .14$, indicated that the combined effect was less than the sum of the individual effects. Again, the critical patterns observed in Experiment 1a replicated here.

Discussion

The results of Experiment 1a and Experiment 1b correspond very well. We will shortly consider the principal results—those from the incongruent condition. But the results from the congruent condition, although not the focus of this study, seem puzzling: Rather than facilitation, there was consistent interference resulting from words congruent with the display color both across conditions within each experiment and across the two replications. Certainly, the average interference in the congruent conditions (overall mean of 40 ms) is substantially lower than the average interference in the incongruent conditions (overall mean of 116 ms), but this unanticipated interference from congruent words still appears to be reliable and thus warrants an explanation, although any such explanation is necessarily post hoc.

Our suspicion is that the apparent interference in the congruent condition has to do not with the congruent condition being slow but with the control condition being fast. Notice that for the incongruent conditions, the lowest interference occurs in the auditory-only condition where the to-be-ignored visual stimulus is xxxxx (mean interference across the replications = 54 ms), and the next lowest interference occurs in the visual-only condition where the to-be-ignored auditory stimulus is the tone (mean interference across the replications = 118 ms), contrasting with the cases where there is a word in both modalities: visual + auditory same (mean interference across the replications = 137 ms) and visual + auditory different (mean interference across the replications = 158 ms).

We note that this same pattern is evident for the congruent conditions: Interference is lowest in the visual-only and auditory-only congruent conditions. The “placeholders” in each modality—tone and xxxxx—are quite distinctive both in physical form and in frequency of occurrence relative to the words appearing on all of the other trials. In particular, their combination in the control condition may have signaled to participants that the color was all that mattered on such trials, alerting them and making it easier for them to focus on the color, thereby speeding responding. This would, of course, lead to apparent interference in the congruent condition. We note that this apparent interference in the congruent condition disappears in Experiments 2 and 3 where the distinctiveness of xxxxx in particular, but also of the tone, is reduced.

Our principal results of interest, though, are those from the incongruent conditions. Here, auditory words produced less interference than visual words, as suggested by previous research but not previously tested within a single experiment nor using the individual item method to permit examination of specific relations. Because we used the SOA previously shown to maximize auditory interference (Roelofs, 2005), we are confident that this result is not merely an artifact of the serial nature of auditory stimulus presentation. It is likely the case, however, that the separation of the auditory word and the visual color—as opposed to their integration when both are visual—contributes to this difference.

Two incongruent words led to greater interference than did a single incongruent word, as previously reported in a visual-only study by MacLeod and Bors (2002). Even two distracters with the same meaning (e.g., seeing and hearing the word *green* when the response should be “red” to the print color) produced stronger interference than did a visual or auditory word alone. When the two distracters had the same meaning, their effects were underadditive, indicating that they were not independent. When the two distracters differed in meaning, interference was greater: In fact, interference in the combined condition was approximated by the sum of the interference from the visual and auditory words alone, suggesting that the effects of the two different words were independent. These results are inconsistent with the capture account because only one word should “capture” attention, so that a second word should not influence responding. These results are, however, consistent with the joint influence account, where multiple distracters can each influence responding. Indeed, these findings even demonstrate that independent contributions can be made by words in different modalities, where one might expect the visual word—the color carrier—to uniquely capture distracter processing. Our results also demonstrate that, in accord with intuition, experiencing the same distracter twice has less impact than does experiencing two different distracters.

Experiment 2

Although somewhat counterintuitive, in that integrated stimuli would seem to encourage capture more than separated stimuli, it is possible that results consistent with the joint influence account like those found in Experiments 1a and 1b could be limited to the situation where one distracter is perceptually integrated with the target stimulus and the other is not. For this reason, we carried out a conceptual replication of Experiments 1a and 1b in which the visual distracter stimuli were spatially separated from the color stimulus, a method that has been used frequently in the Stroop literature (see MacLeod, 1998, for a brief review). Interference was expected to be smaller, as is typically the case with separated as opposed to integrated Stroop stimuli (cf. MacLeod, 1998; Roberts & Besner, 2005), but the question was whether the interference pattern would still be the same.

Method

Participants Participants were 48 students (19 men, 29 women) from the University of Texas at El Paso whose primary or only language was English. Their median age was 20 years. All were undergraduate students, most of whom were recruited from introductory psychology classes and received credit toward a class research requirement; the rest were unpaid volunteers.

Materials, design, and procedure The design and procedure were identical to those of Experiment 1a with one exception—rather than being integrated with the color information, the visual words were presented in black as flankers to the colored target stimulus, which was always xxxxx. The colored target stimulus appeared in the center of the screen, with the flanker words presented immediately above and immediately below the target. The experimental conditions are listed in Table 2.

Results

RT data were processed in the same manner as in Experiments 1a and 1b. RTs and their derived interference scores are given in Table 2. As in Experiments 1a and 1b, the error analyses appear in the Appendix, with Table 5 displaying the error data for Experiment 2.

Congruent conditions No reliable facilitation or interference effects were observed in the congruent condition ($ps > .10$).

Incongruent conditions Interference was reliable for all incongruent conditions ($ps < .001$). As in Experiments 1a and 1b, we carried out two 2×2 ANOVAs. In the analysis

Table 2 Experiment 2: Response Times and Interference Scores as a Function of Condition

Condition	Stimulus Examples				RT (ms)		Interference (ms)	
	Auditory	Flanker	Target	Color	Mean	(SE)	Mean	(SE)
<i>Control</i>	<tone>	xxxxx	xxxxx	Blue	722	(22)	–	–
<i>Congruent</i>								
Visual	<tone>	blue	xxxxx	Blue	722	(22)	0	(4)
Auditory	blue	xxxxx	xxxxx	Blue	732	(23)	10	(6)
Visual + Auditory	blue	blue	xxxxx	Blue	725	(23)	2	(6)
<i>Incongruent</i>								
Visual	<tone>	red	xxxxx	Blue	758	(21)	36	(5)
Auditory	red	xxxxx	xxxxx	Blue	768	(22)	46	(5)
Vis + Aud Same	red	red	xxxxx	Blue	811	(22)	88	(7)
Vis + Aud Diff	red	green	xxxxx	Blue	808	(22)	86	(6)

involving the visual + auditory same condition, the presence of a word increased interference, relative to the control condition, both when that word was visual, $F(1, 47) = 96.92$, $MSE = 756$, $p < .001$, partial $\eta^2 = .67$, and when it was auditory, $F(1, 47) = 110.81$, $MSE = 1,047$, $p < .001$, partial $\eta^2 = .70$. The auditory and visual words did not interact, $F(1, 47) = 2.01$, $MSE = 291$, $p = .163$, partial $\eta^2 = .04$. In the analysis including the visual + auditory different condition, again the presence of a word increased interference both for visual words, $F(1, 47) = 121.21$, $MSE = 569$, $p < .001$, partial $\eta^2 = .72$, and for auditory words, $F(1, 47) = 105.49$, $MSE = 1,049$, $p < .001$, partial $\eta^2 = .69$. And again, the auditory and visual words did not interact, $F < 1$, partial $\eta^2 = .02$.

Discussion

Experiment 2 showed that the interference magnitudes of auditory words and of spatially separated visual words were comparable; if anything, the spatially separated visual words appeared to have a smaller basic interference effect. Auditory color–word distracters and spatially separated visual color–word distracters produced independent interference effects in color naming, whether they were the same word or two different words. Thus, although effect sizes were smaller in Experiment 2 because of the word and color being separated in the visual display, the critical interference effects seen in Experiments 1a and 1b were replicated, again more consistent with the joint influence account than with the capture account. Note that the use of xxxxx as the color carrier on all trials made that visual stimulus more frequent and less distinctive than it was in Experiments 1a and 1b, helping to explain why the apparent interference observed for congruent trials in Experiments 1a and 1b was no longer evident in Experiment 2.

Experiment 3

Finally, we considered the possibility that using the auditory modality was somehow special in producing nonredundant effects. Auditory and visual words could have independent effects because, as separate modalities, they might tap different pools of attentional resources (e.g., Wickens, 1984). If so, then auditory and visual words might exert joint influence, but multiple visual words might not. To address this issue, in Experiment 3, only visual distracters were included. A visual distracter word was presented either integrated with the target color, separated from the target color (the word above or below the colored stimulus), or both. In the case of both, paralleling the preceding experiments, the two words could be the same or different. The critical change was that there were no auditory stimuli in Experiment 3. Would the pattern for two visual stimuli—one integrated and one separated—be analogous to that for the case of an auditory stimulus and a visual stimulus? How large a role has using two modalities played in the results observed thus far?

Method

Participants Participants were 52 students (19 men, 33 women) from the University of Texas at El Paso whose primary or only language was English. Their median age was 19 years. All were undergraduate students, most of whom were recruited from introductory psychology classes and received credit toward a class research requirement; the remainder were unpaid volunteers.

Stimuli As previously, visual word stimuli were the color words *red*, *blue*, *green*, and *purple*, and the control stimulus was xxxxx. On each trial, three visual stimuli were presented. The target stimulus was presented in color and two distracter

stimuli were presented in black, one immediately above and one immediately below the target, thus flanking the target. The flanker stimuli were always identical to each other. The center target stimulus was used for presenting integrated distracters, and the black flanking stimuli were used to present the spatially separated distracters.

Design The eight included trial types are illustrated in Table 3. In the integrated-only condition, the target stimulus was the name of a color word, and the flankers were both xxxxx. In the separated-only condition, the target stimulus was xxxxx, and the flanker stimuli were both the same color word. In the integrated + separated same condition, the target and flanker stimuli were the same color word. In the integrated + separated different condition, the target word and flanker words were different color words. For each trial type, every possible color combination was used equally often, and there were 48 trials of each type.

Procedure Participants were given instructions and were tested as in Experiment 1. On each trial, the target and flanker stimuli appeared simultaneously in the center of the computer monitor and remained until a vocal response was registered.

Results

RT data were processed in the same manner as in Experiments 1a, 1b, and 2. RTs and interference scores derived from the RTs are given in Table 3. As in the prior experiments, the error analyses appear in the Appendix, with Table 6 displaying the error data for Experiment 3.

Congruent conditions None of the congruent condition RTs differed from that of the control condition, $ps > .20$.

Incongruent conditions Interference was reliable in all incongruent conditions, $ps < .001$. Two 2 (integrated word present or absent) \times 2 (separated word present or absent) repeated-measures ANOVAs were conducted. The first analysis was conducted using the integrated + separated same condition. The presence of an integrated word increased interference, $F(1, 51) = 192.47, MSE = 5,188, p < .001, \text{partial } \eta^2 = .79$; the presence of a separated word marginally increased interference, $F(1, 51) = 4.05, MSE = 1,032, p = .050, \text{partial } \eta^2 = .07$. The effects of the integrated and separated words interacted, $F(1, 51) = 22.65, MSE = 1,508, p < .001, \text{partial } \eta^2 = .31$, such that their combined effect was less than the sum of their individual effects. The second analysis was conducted using the integrated + separated different condition. The presence of a word increased interference both for integrated stimuli, $F(1, 51) = 225.17, MSE = 5,808, p < .001, \text{partial } \eta^2 = .82$, and for separated stimuli, $F(1, 51) = 32.43, MSE = 1,345, p < .001, \text{partial } \eta^2 = .39$. In this analysis, the effects of integrated and separated words did not interact, $F(1, 51) = 1.28, MSE = 1,282, p = .262, \text{partial } \eta^2 = .03$.

Discussion

As in previous research (e.g., MacLeod, 1998; Roberts & Besner, 2005), integrated color–word distracters produced greater interference than did spatially separated color–word distracters, $t(51) = 11.13, p < .001, \text{partial } \eta^2 = .71$. The effects of these two distracter types were nonredundant when the distracter words were different from each other: Both contributed to total interference. The observed pattern of interference shows that it is not necessary to have different modalities of distracter presentation to produce non-redundant interference effects in color naming. The data pattern in Experiment 3 is consistent with those of Experiments 1a, 1b, and 2, again

Table 3 Experiment 3: Response Times and Interference Scores as a Function of Condition

Condition	Stimulus Examples			RT (ms)		Interference (ms)	
	Flanker Stimulus	Target Stimulus	Target Color	Mean	(SE)	Mean	(SE)
<i>Control</i>	xxxxx	xxxxx	Blue	813	(13)	–	–
<i>Congruent</i>							
Integrated	xxxxx	blue	Blue	815	(14)	2	(7)
Separated	blue	xxxxx	Blue	816	(13)	3	(6)
Integ + Sep	blue	blue	Blue	807	(14)	-6	(7)
<i>Incongruent</i>							
Integrated	xxxxx	red	Blue	977	(21)	164	(13)
Separated	red	xxxxx	Blue	848	(14)	35	(6)
Integ + Sep Same	red	red	Blue	961	(17)	148	(10)
Integ + Sep Diff	green	red	Blue	1001	(18)	188	(12)

fitting the joint influence account better than the capture account.

General Discussion

Consistent with previous comparisons done across different studies, Experiment 1 showed that the interference caused by an auditory color–word was weaker than the standard visual effect when presentation of the visual word is integrated with the color. In fact, though, Experiment 2 demonstrated that the interference from auditory distracters was comparable to the interference observed for visual distracters when the visual word is presented in a spatially separate location. Taken together, these findings imply that it is not the visual modality itself but the fact that the visual stimuli are ordinarily integrated that accounts for the differential modality effect.

Across the present series of experiments, two distracters elicited larger interference effects than one distracter, whether the two distracters were the same word or two different words. This was true for an auditory distracter and an integrated visual distracter (Experiment 1a) and for an auditory distracter and a spatially separated visual distracter (Experiment 2). When auditory and visual distracters were different words, they made independent contributions to color naming interference, whether the visual distracter was spatially integrated with (Experiment 1a) or spatially separated from (Experiment 2) the target color stimulus. Similarly, when both distracters were visual, spatially integrated and spatially separated distracters had independent effects on color naming (Experiment 3).

Our argument has been that these results conflict with predictions of a capture account but are consistent with a joint influence account. Under a capture account (see, e.g., Kahneman & Chajczyk, 1983; MacLeod & Hodder, 1998), while trying to attend to the color, a simultaneously available nominally irrelevant word can intrude on attention, causing interference. But ordinarily only one word would do this on a given trial, the basis for using the word *capture*: The first word that is attended gates out other words that are also present. In contrast, under a joint influence account (see, e.g., MacLeod & Bors, 2002), multiple nominally irrelevant words can attract and share attention, with each word potentially influencing color naming and contributing to overall total interference.

How do our results align with the predictions of these two accounts? First, we have created three situations where multiple nominally irrelevant words are present: one visual and integrated plus one auditory (Experiment 1), one visual and separated plus one auditory (Experiment 2), and one visual and integrated plus one visual and separated (Experiment 3). The patterns of results across these three situations correspond well, as we have just summarized. Interference from an

auditory word (necessarily separated) or from a separated visual word is relatively small and consistent, as seen in Experiments 1 and 2. Interference grows substantially when the irrelevant word is integrated with the color information—the classic Stroop procedure. Having the same irrelevant word appear in two places (either auditory + visual or visual integrated + visual separated) does little to alter the interference. But having two different irrelevant words occur on a single trial does further increment interference. This last result implies that all words that are present on a trial are processed and that, when these include more than one incongruent word, interference is heightened. Only the joint influence account correctly anticipates this outcome.

Is it nevertheless possible for a version of the capture account to handle this pattern? Certainly, attention can be engaged and then disengaged (the “catch” and “release” idea), as the work of Posner and his colleagues (e.g., Posner, Walker, Friedrich, & Rafal, 1984) and others has clearly demonstrated. So following an initial capture by one dimension of a stimulus event, attention could disengage and then subsequently be captured by another dimension of that same stimulus event. In this way, each dimension of a stimulus could influence performance, thereby mimicking joint influence. It may really be a question, then, of whether the two elements are exerting their influence in parallel or sequentially. Although we see our results as more in keeping with a joint influence explanation, particularly given other results in the literature (e.g., La Heij, van der Heijden, & Plooi, 2001; MacLeod & Bors, 2002), we acknowledge that the successive capture account is difficult to discriminate from the joint influence account, and that no single study will be able to do so.

This study has provided an approach to understanding interference that converges on the one used by MacLeod and Bors (2002), who reached the same conclusion favoring joint influence. We suspect that prior studies that failed to observe more interference for two distracters than for one, or more interference for two different distracters than for the same distracter presented twice (e.g., Kahneman & Chajczyk, 1983; MacLeod & Hodder, 1998; Yee & Hunt, 1991), did not have sufficient power to observe these effects, due to substantially smaller sample sizes (and perhaps also number of trials per condition) than we used in this series of experiments: Our relatively small standard errors are testament to this. We first established the joint influence using auditory and visual distracters, but it is also clear that joint influence occurs even within a single modality, given the pattern for visual-only stimuli in Experiment 3.

Taking this analysis a step farther, the data from Experiments 1 to 3 display an interesting regularity. When there is no integration of color and word, as in Experiment 2, interference is relatively small—about 40 ms as seen in the incongruent visual (separated) condition. But whenever a Stroop stimulus is integrated—the

color is carried by the words itself (e.g., RED in green)—this results in a considerably larger amount of interference—in our experiments, at least twice as large. The incongruent visual condition of Experiment 1 is the best illustration, with interference incremented by 80 ms from the 40 ms in Experiment 2 to the 120 ms in Experiment 1 as a consequence of integration. Adding a second, necessarily nonintegrated distracter, raises interference by about 40 ms. This can best be seen in Experiment 1 in moving from the single (integrated) visual distracter to the auditory + visual dual distracters, and in Experiment 2 in moving from the single (separated) visual or auditory distracter to one of each.

The Stroop task continues to serve as a fertile testing ground for developing our understanding of interference. As Melara and Algom (2003) pointed out, Stroop chose just the right combination of stimulus and response conditions to maximize interference. Indeed, our experiments reconfirm that his choice of integrating color and word is crucial to observing large interference. It is the case, however, that nonintegrated stimuli can exert interference, and that they can do so even in the presence of integrated stimuli by augmenting the amount of interference. Moreover, this is true even for stimuli in a different modality—the focus of the present study. We appear routinely to process multiple stimuli in Stroop settings, as indicated by their apparent joint influence on interference.

Author Note We gratefully acknowledge the assistance of Jacqueline Carbajal, Ana Flores-Chan, and J'me Jennifer Garcia Phelps in data collection and Arlene Luévano in data processing. Experiment 2 was conducted as a senior honors thesis by the third author. The findings of Experiment 1a were presented at the 50th Annual Meeting of the Psychonomic Society.

Address correspondence to Wendy S. Francis, Department of Psychology, University of Texas at El Paso, El Paso, TX 79968, wfrancis@utep.edu

Appendix: Error Rate Analyses

Experiment 1a

Error rates and error interference scores for Experiment 1a are given in Table 4.

Congruent conditions None of the congruent conditions showed significant facilitation or interference, $ps > .10$.

Incongruent conditions Interference was significant in the visual word condition and in both of the visual + auditory conditions, $ps < .001$. However, interference was not significant in the auditory word condition, $t < 1$. To test whether the effects of interference from the two modalities interacted or were independent, two 2 (visual) × 2 (auditory) repeated-measures ANOVAs were conducted on the control, visual-only, auditory-only, and visual + auditory condition error rates. The first analysis used the visual + auditory same condition. Presence of the visual word increased interference, $F(1, 49) = 28.01$, $MSE = .00206$, $p < .001$, partial $\eta^2 = .36$, whereas presence of the auditory word did not, $F < 1$, partial $\eta^2 = .02$. Effects of the auditory and visual words did not interact, $F < 1$, partial $\eta^2 = .01$. The second analysis used the visual + auditory different condition. Presence of the visual word increased interference reliably, $F(1, 49) = 39.43$, $MSE = .00231$, $p < .001$, partial $\eta^2 = .45$, whereas presence of the auditory word did not, $F(1, 49) = 3.22$, $MSE = .00057$, $p = .079$, partial $\eta^2 = .06$. Effects of the auditory and visual words did not interact, $F(1, 49) = 2.09$, $MSE = .00076$, $p = .155$, partial $\eta^2 = .04$.

Table 4 Experiments 1a and 1b: Error Rates and Interference Scores

Condition	Experiment 1a				Experiment 1b			
	Error Rate (%)		Interference (%)		Error Rate (%)		Interference (%)	
	Mean	(SE)	Mean	(SE)	Mean	(SE)	Mean	(SE)
Control	0.8	(.23)	–	–	0.5	(.18)	–	–
<i>Congruent</i>								
Visual	0.5	(.18)	–0.3	(.26)	0.0	(.04)	–0.5	(.19)
Auditory	0.8	(.19)	0.0	(.21)	0.4	(.17)	–0.1	(.26)
Visual + Auditory <i>Incongruent</i>	0.4	(.12)	–0.4	(.24)	0.2	(.10)	–0.3	(.19)
Visual	4.5	(.74)	3.7	(.68)	2.0	(.37)	1.5	(.36)
Auditory	0.8	(.22)	0.0	(.28)	0.3	(.16)	–0.2	(.26)
Vis + Aud Same	3.9	(.82)	3.1	(.72)	1.2	(.26)	0.7	(.31)
Vis + Aud Diff	5.7	(.90)	4.9	(.81)	2.2	(.47)	1.7	(.52)

Experiment 1b

Error rates and error interference scores for Experiment 1b are given in Table 4.

Congruent conditions None of the congruent conditions showed significant facilitation or interference.

Incongruent conditions Interference was significant in the visual word condition, $t(49) = 4.21$, $p < .001$ and in the visual + auditory different condition, $t(49) = 3.30$, $p = .002$. However, interference was not significant in the auditory word condition, $t < 1$, or the visual + auditory same condition, $t(49) = 2.27$, $p = .028$. To test whether the effects of interference from the two modalities interacted or were independent, two 2 (visual) \times 2 (auditory) repeated-measures ANOVAs were conducted on the control, visual-only, auditory-only, and visual + auditory condition error rates. The first analysis used the visual + auditory same condition. Presence of the visual word increased interference, $F(1, 49) = 22.76$, $MSE = .00031$, $p < .001$, partial $\eta^2 = .32$, as did presence of the auditory word, $F(1, 49) = 4.70$, $MSE = .00025$, $p = .035$, partial $\eta^2 = .09$. Effects of the auditory and visual words did not interact, $F(1, 49) = 1.83$, $MSE = .00027$, $p = .183$, partial $\eta^2 = .04$. The second analysis used the visual + auditory different condition. Presence of the visual word increased interference reliably, $F(1, 49) = 22.17$, $MSE = .00065$, $p < .001$, partial $\eta^2 = .31$, whereas presence of the auditory word did not, $F < 1$, partial $\eta^2 = .00$. Effects of the auditory and visual words did not interact, $F(1, 49) = 1.08$, $MSE = .00016$, $p = .304$, partial $\eta^2 = .02$.

Experiment 2

Error rates and error interference scores for Experiment 2 are given in Table 5. Facilitation did not reach significance in any of the congruent conditions, $ps > .04$, nor did interference reach significance in any of the incongruent conditions, $ps > .025$.

Experiment 3

Error rates and error interference scores for Experiment 3 are given in Table 6.

Congruent conditions There was no significant facilitation in any of the congruent conditions, $ps > .20$.

Incongruent conditions Interference was significant in the integrated word condition, the integrated + separated same condition, and the integrated + separated different condition,

Table 5 Experiment 2: Error Rates and Interference Scores

Condition	Error Rate (%)		Interference (%)	
	Mean	(SE)	Mean	(SE)
<i>Control</i>	0.9	(.21)	–	–
<i>Congruent</i>				
Visual	0.8	(.21)	-0.1	(.28)
Auditory	0.5	(.17)	-0.4	(.19)
Visual + Auditory	0.4	(.18)	-0.5	(.23)
<i>Incongruent</i>				
Visual	1.1	(.22)	0.2	(.26)
Auditory	1.2	(.27)	0.3	(.29)
Vis + Aud Same	1.6	(.28)	0.7	(.32)
Vis + Aud Diff	1.6	(.31)	0.7	(.37)

$ps < .001$. However interference was not significant in the separated word condition, $t < 1$. To test whether the effects of interference from the integrated and separated words interacted or were independent, two 2 (integrated) \times 2 (separated) repeated-measures ANOVAs were conducted on the control, integrated-only, separated-only, and integrated + separated condition error rates. The first analysis used the integrated + separated same condition. The presence of the integrated word increased interference, $F(1, 51) = 42.49$, $MSE = .00238$, $p < .001$, partial $\eta^2 = .45$. However, the presence of the separated word did not, $F < 1$, partial $\eta^2 = .01$, nor did the effects of the integrated and separated words interact, $F < 1$, partial $\eta^2 < .01$. The second analysis used the integrated + separated different condition. The presence of the integrated word increased interference, $F(1, 51) = 42.69$, $MSE = .00241$, $p < .001$, partial $\eta^2 = .46$, but the presence of the separated word did not, $F < 1$, partial $\eta^2 = .01$. The effects of the integrated and separated words did not interact, $F < 1$, partial $\eta^2 < .01$.

Table 6 Experiment 3: Error Rates and Interference Scores

Condition	Error Rate (%)		Interference (%)	
	Mean	(SE)	Mean	(SE)
<i>Control</i>	1.1	(.26)	–	–
<i>Congruent</i>				
Integrated	1.0	(.27)	-0.1	(.25)
Separated	0.9	(.24)	-0.2	(.21)
Integ + Sep	1.2	(.38)	0.0	(.29)
<i>Incongruent</i>				
Integrated	5.7	(.84)	4.6	(.78)
Separated	1.0	(.28)	-0.1	(.29)
Integ + Sep Same	5.2	(.83)	4.1	(.76)
Integ + Sep Diff	5.3	(.78)	4.2	(.76)

References

- Cho, Y. S., Choi, J. M., & Proctor, R. W. (2012). Likelihood of attending to the color word modulates Stroop interference. *Attention, Perception, & Psychophysics*, *74*, 416–429.
- Cho, Y. S., Lien, M.-C., & Proctor, R. W. (2006). Stroop dilution depends on the nature of the color carrier but not on its location. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 826–839.
- Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavior Research Methods, Instruments, and Computers*, *25*, 257–271.
- Cowan, N. (1989). The reality of cross-modal Stroop effects. *Perception & Psychophysics*, *45*, 87–88.
- Cowan, N., & Barron, A. (1987). Cross-modal, auditory-visual Stroop interference and possible implications for speech memory. *Perception & Psychophysics*, *41*, 393–401.
- Elliot, E. M., Cowan, N., & Valle-Inclan, F. (1998). The nature of cross-modal color–word interference effects. *Perception & Psychophysics*, *60*, 761–767.
- Elliott, E. M., Morey, C. C., Morey, R. D., Eaves, S. D., Shelton, J. T., & Lutfi-Proctor, D. A. (2014). The role of modality: Auditory and visual distractors in Stroop interference. *Journal of Cognitive Psychology*, *26*, 15–26.
- Green, E. J., & Barber, P. J. (1981). An auditory Stroop effect with judgments of speaker gender. *Perception & Psychophysics*, *30*, 459–466.
- Kahneman, D., & Chajczyk, D. (1983). Tests of the automaticity of reading: Dilution of Stroop effects by color-irrelevant stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, *9*, 497–509.
- Klein, G. S. (1964). Semantic power measured through the interference of words with color-naming. *American Journal of Psychology*, *77*, 576–588.
- La Heij, W., van der Heijden, A. H. C., & Plooi, P. (2001). A paradoxical exposure-duration effect in the Stroop task: Temporal segregation between stimulus attributes facilitates selection. *Journal of Experimental Psychology: Human Perception & Performance*, *27*, 622–632.
- MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*, 163–203.
- MacLeod, C. M. (1998). Training on integrated versus separated Stroop tasks: The progression of interference and facilitation. *Memory & Cognition*, *26*, 201–211.
- MacLeod, C. M., & Bors, D. A. (2002). Presenting two color words on a single Stroop trial: Evidence for joint influence, not capture. *Memory & Cognition*, *30*, 789–797.
- MacLeod, C. M., & Hodder, S. L. (1998). Presenting two incongruent color words on a single trial does not alter Stroop interference. *Memory & Cognition*, *26*, 212–219.
- McClain, L. (1983). Stimulus-response compatibility affects auditory Stroop interference. *Perception & Psychophysics*, *33*, 266–270.
- Melara, R. D., & Algom, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review*, *110*, 422–471.
- Melara, R. D., & Mounts, J. R. W. (1993). Selective attention to Stroop dimensions: Effects of baseline discriminability, response mode, and practice. *Memory & Cognition*, *21*, 627–645.
- Miles, C., Madden, C., & Jones, D. M. (1989). Cross-modal, auditory-visual Stroop interference: A reply to Cowan and Barron (1987). *Perception & Psychophysics*, *45*, 77–81.
- Posner, M. I., Walker, J. A., Friedrich, F. J., & Rafal, R. D. (1984). Effects of parietal injury on covert orienting of attention. *The Journal of Neuroscience*, *4*, 1863–1874.
- Roberts, M. A., & Besner, D. (2005). Stroop dilution revisited: Evidence for domain-specific, limited-capacity processing. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 3–13.
- Roelofs, A. (2005). The visual-auditory color–word Stroop asymmetry and its time course. *Memory & Cognition*, *33*, 1325–1336.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662.
- Wickens, C. D. (1984). Processing resources and attention. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 63–102). New York: Academic Press.
- Yee, P. L., & Hunt, E. (1991). Individual differences in Stroop dilution: Tests of the attention-capture hypothesis. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 715–725.