Perception & Psychophysics
1989, 45 (1), 82-84

Notes and Comment

A reply to Miles, Madden, and Jones (1989): Mistakes and other flaws in the challenge to the cross-modal Stroop effect

NELSON COWAN
University of Missouri, Columbia, Missouri

Cowan and Barron (1987) demonstrated that spoken color words can interfere with performance in a visual color-naming task. This cross-modal Stroop effect has important implications for memory and attention, but Miles, Madden, and Jones (1989) questioned the reliability of the findings. After reading Miles et al.'s challenge, I urge readers to directly compare the two articles. There are clear, important differences in both method and interpretation. The basic points to be made about Miles et al.'s study are as follows: (1) Their Experiments 1a and 1b, which were the attempted replications, had serious, obvious methodological flaws that were not present in Cowan and Barron. (2) Their Experiment 2 did not require a spoken response, did not include spoken color-word interference, and is thus irrelevant to the main concerns of Cowan and Barron. Moreover, their discussion surrounding Experiment 2 was misleading on that point. (3) Cowan and Barron's method was criticized on grounds that are unwarranted. (4) Finally, although some of the theoretical points that Miles et al. brought up are interesting, their claim that Cowan and Barron's results are inconsistent with past research and current theory is incorrect. These points will be addressed in turn.

EXPERIMENTS 1a AND 1b OF MILES ET AL.

Experiment 1a of Miles et al. was an attempt to replicate the basic findings of Cowan and Barron. As in Cowan and Barron, a spoken response was required. Experiment 1b was similar, except that subjects were to whisper their responses. Unfortunately, both experiments contained serious departures from the method of Cowan and Barron.

Erroneously Determined Rate of Auditory Presentation

In their Method section, Cowan and Barron clearly stated that spoken color words were presented at the rapid rate of 1.66 words/sec (i.e., 0.6 sec/word). However, Miles et al. incorrectly took this figure to be 1.66 sec/word, a much slower rate. Then they decided that they would "increase" the rate of presentation in their own Experiments 1a and 1b to 1 word/sec, in order to present more phonetic material per unit of time and thereby increase the likelihood of spoken interference. However, they inadvertently *decreased* the rate of presentation substantially. Thus, according to their own logic, Miles et al. made their experiment less sensitive than that of Cowan and Barron.

It is not difficult to think of additional reasons why the rate of auditory presentation might be important. For example, the rate that Cowan and Barron used is roughly comparable to the subjects' mean response rate, which might promote confusions between the auditory input and a speech response buffer.

Insufficient Number of Subjects in Miles et al.

Cowan and Barron used 30 subjects in their experiment. In striking contrast, Miles et al. reported using 12 subjects in Experiment 1a and 8 subjects in Experiment 1b. They did not study each subject intensively; in fact, the number of trials per subject was roughly comparable to that of Cowan and Barron. If we had used so few subjects, it is doubtful whether we would have obtained a cross-modal Stroop effect either.[1]

The fact that Miles et al. obtained the conventional Stroop effect with so few subjects does not imply that their experiments had sufficient power to detect the cross-modal effect, because Cowan and Barron found that the cross-modal effect was much smaller in magnitude than the conventional Stroop effect.

It is also instructive to examine the mean reading times from the CW condition (color/word conflict) in Miles et al.'s Experiment 1a, in which simultaneous auditory and visual Stroop interference was presented. The direction and magnitude of the difference between means for spoken color words versus those for silence are comparable to what Cowan and Barron obtained. One might speculate that, with more subjects, a cross-modal effect might be obtained despite the slower auditory presentation rate, provided that visual Stroop stimuli are present to maximize the difficulty of the task. The means in Experiment 1b, which involved whispered responding, do not correspond to this speculation, but it appears that whispering slowed the subjects quite a bit (cf. Tables 1 and 2), and this could have further decreased the sensitivity of the task to cross-modal Stroop effects.

Other Discrepancies in Method

There are several other differences in method that might potentially be important. First, Cowan and Barron used

This research was supported by National Institutes of Health Grant R01-21338 awarded to the author. Comments in the present paper refer to the typewritten copy of Miles et al., inasmuch as there was no access to revisions that might be made within their galley proofs. I thank Jean M. Ispa for helpful comments. Address correspondence to the author at the Department of Psychology, 210 McAlester Hall, University of Missouri, Columbia, MO 65211.

a set of five colors in both the spoken and written materials, whereas Miles et al. used four colors. The task could be more difficult and susceptible to cross-modal effects when the possible response set is larger.

Despite a statement by Miles et al. to the contrary, Cowan and Barron used the same color names in both auditory and visual materials, as did Miles et al.

A final methodological point is that the trial composition of the experimental sessions differed in the two studies. In the Cowan and Barron study, each subject encountered spoken color-word interference on 2 out of 10 trials (i.e., 20%). In contrast, each subject in the Miles et al. study encountered spoken color-word interference on 3 out of 6 trials (i.e., 50%). Subjects might learn to overcome cross-modal Stroop interference with extended presentation of spoken color words, a possibility that I am currently investigating.

EXPERIMENT 2 OF MILES ET AL.

Irrelevance of this Experiment to the Main Concern of Cowan and Barron

Experiment 2 of Miles et al. involved a color-comparison task with a buttonpress response. No effects of irrelevant speech were found. The inclusion of this experiment in the article is puzzling, however, because the only auditory stimulus materials used were recorded lectures; spoken color-word presentations were *not* included. Because Cowan and Barron found that non-color-word speech input did not interfere with color naming, no effect of irrelevant speech would be expected either.

Misleading Allusions to Experiment 2

In their introduction, Miles et al. stated that "three experiments ... each [provide] evidence suggesting that irrelevant, spoken color items do not interfere with the visual Stroop task." This is misleading, given that Experiment 2 had no spoken color-word stimuli. Similar statements were made in several other places within the Miles et al. manuscript. (Note that this tangential experiment was the only one with an adequate number of subjects.)

UNWARRANTED CRITICISMS OF COWAN AND BARRON'S METHOD

Miles et al. made two criticisms of the method used by Cowan and Barron. First, they criticized the fact that the intensity levels of the spoken color-word condition and the three control conditions with sound (the alphabet, repetitions of the word "the," and music) were equated subjectively by the authors rather than being presented at equal decibel levels. In reply, it is not clear that an equal-intensity presentation is desirable, given that the stimuli for the different conditions could differ in perceived loudness if presented at equal decibel levels (because intensity and perceived loudness are affected differently by the unique spectral properties of each sound). Second, one can in no way attribute to intensity differences the basic effect under consideration: namely, the difference between the spoken color-word and silent conditions.[2]

Of the five auditory conditions of Cowan and Barron, only the spoken color-word condition produced different results from the silent control condition. One conceivably might be concerned that the specificity of the effect is an artifact of the presentation levels, and that there actually is an effect of spoken non-color words as well as of spoken color words. However, Miles et al. did not include non-color-word and color-word interference conditions in the same experiment, so they cannot speak to this possibility.

Miles et al. also criticized Cowan and Barron for presenting error rates even though these were very low. However, the main conclusions of the Cowan and Barron study would be the same if response times alone were reported. It should also be noted that Cowan and Barron carried out a logarithmic transformation on the error data in order to arrive at a measure more suitable for statistical analysis. Reporting error data was a strength, not a weakness, because it ruled out the possibility that effects in response time might reflect speed/accuracy trade-offs.

The cross-modal Stroop effect was a highly reliable one, even though it may have been attenuated to some extent by presentation order effects. The effects of auditory condition on response time was significant at the $p < .001$ level, as we reported. A reexamination of our data also indicates that the mean difference between the spoken color-word and silent conditions was in the correct direction in 24 subjects and the wrong direction in only 5 subjects, with 1 tie ($p < .01$, sign test).

DIFFERENCES IN THEORETICAL POSITION

Miles et al. asserted not only that cross-modal Stroop effects do not occur, but also that they should not occur according to past research and current theory. These assertions do not seem well-founded. Both the interpretation of cross-modal Stroop effects and its relation to other research must be reexamined.

Miles et al. discussed two different views of how irrelevant speech works. According to one view (Broadbent, 1983), irrelevant speech can affect the perception of visual stimuli. According to the other view (Baddeley & Salamé, 1986), which was supported by several empirical studies, irrelevant speech affects a speech memory buffer rather than affecting perception of the visual stimuli. Like Miles et al., I endorse the latter view, and the discussion provided by Cowan and Barron should make this clear. However, Miles et al. seemed to suggest that cross-modal Stroop effects would provide evidence that irrelevant speech affects perception. Perhaps they reached this conclusion because Broadbent (1983, p. 731, quoted in Baddeley & Salamé, 1986) suggested that irrelevant speech is "a Stroop effect in perception," and Baddeley and Salamé (1986, p. 526) questioned

whether unattended speech interferes with "the perception of ... visually presented letters in a Stroop-like manner."

In fact, however, there is sufficient evidence to suggest that Stroop effects occur closer toward the response end of the stimulus-response continuum and are greatly affected by stimulus-response compatibility (e.g., see Virzi & Egeth, 1985). Cowan and Barron's account of cross-modal Stroop effects is consistent with this. It is assumed that familiar, spoken color words are perceived automatically, even if little or no attention is directed toward them, and that these spoken color words automatically enter into a memory buffer that is used to make a spoken response. According to this account, irrelevant speech would not be assumed to alter the perception of the visual stimuli.

According to Miles et al., the cross-modal Stroop effect is not the kind of effect one would expect on the basis of past research on irrelevant speech. They claimed that interference from irrelevant speech should occur only when the subject must rehearse stimuli in order to recall them in their correct serial order. To support this view, they referred to Baddeley and Salamé's (1986) finding that subjects could judge whether pairs of letters rhymed without interference from irrelevant speech (Arabic prose). However, as Miles et al. noted, Baddeley and Salamé's experiments did not require a spoken response. In Baddely and Salamé's study, along with Cowan and Barron's study, two very different types of effect of irrelevant speech can be distinguished: (1) in Baddeley and Salamé, a very general effect that occurred when subjects rehearsed a list of items, and (2) in Cowan and Barron, an effect that occurred only when an auditorily presented item was a viable response candidate active in memory at the time the subject made a spoken response.

Although these different effects might depend upon different components within the processing system, one need not make that assumption in order for the cross-modal Stroop effect to be consistent with past research and current theory. Both effects could involve the "phonological store" and "articulatory control process" (ACP) discussed by Miles et al., if one simply assumes that the ACP is used both to enter items into the phonological store (as Baddeley and others have assumed) and also to select items from the store for speech output.

In the memory situation, subjects would use the ACP to refresh items in the phonological store. Irrelevant speech would have its effect by replacing items within that store when the store was already being used to its full capacity. In contrast, in the cross-modal Stroop situation, subjects would use the ACP to select items from the store for a spoken response. Spoken color words would enter the phonological store and make selection of the correct response more difficult, not because the store is used to capacity, but because the spoken color word competes with the correct candidate for selection and speech output.

In Baddeley's and Salamé's (1986) situation, in which subjects were to determine whether two letters rhymed, no interference from irrelevant speech would be expected according to this account (and none was obtained). The phonological store was not loaded to capacity, a spoken response was not required, and the irrelevant speech did not contain a string of viable response candidates. This also was the case in Experiment 2 of Miles et al.

The above account may be only one of several reasonable scenarios showing that there is no conflict between Cowan and Barron's data and past data or current theory. Indeed, the above considerations suggest that Cowan and Barron's discussion helps to clarify and further develop current theory.

I do not wish to imply that replication and extension of the cross-modal Stroop effect is unnecessary. Cowan and Barron presented only a single experiment, and the boundary conditions of the cross-modal Stroop effect have not been established. However, the experiments of Miles et al. do little to satisfy the need for additional research.

REFERENCES

BADDELEY, A., & SALAMÉ, P. (1986). The unattended speech effect: Perception or memory? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 525-529.

BROADBENT, D. E. (1983). Recent advances in understanding performance in noise. In *Noise as a public health problem: Proceedings of the Fourth International Congress* (pp. 719-738). Milan: Edizioni Tecniche a cura del Centro Ricerche e Studi Amplifon.

COWAN, N., & BARRON, A. (1987). Cross-modal, auditory-visual Stroop interference and possible implications for speech memory. *Perception & Psychophysics*, **41**, 393-401.

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.

VIRZI, R. A., & EGETH, H. E. (1985). Toward a translational model of Stroop interference. *Memory & Cognition*, **13**, 304-319.

NOTES

1. It is important to point out a simple mistake in Cowan and Barron that is inconsequential when corrected. We stated in our Method section that there were 32 subjects, but we inadvertently counted the subject who was excluded because she could not identify the colors and another who failed to complete the experiment. As we stated in our Results section, all of the analyses were carried out with 30 subjects, not 32.

2. Moreover, all speech sounds were presented at about 85 ± 5 dB(A) against a background of less than 40 dB(A) measured with a GenRad 1565-B sound level meter and 9-A type artificial ear, although the earphones were nonstandard for sound measurement. The music tape was more variable (80 ± 10 dB).

(Manuscript received March 24, 1988;
revision accepted for publication June 10, 1988.)