The role of visual-auditory "compellingness" in the ventriloquism effect: Implications for transitivity among the spatial senses

DAVID H. WARREN
University of California, Riverside, California 92521

ROBERT B. WELCH
University of Kansas, Lawrence, Kansas 66044

and

TIMOTHY J. McCARTHY
University of California, Riverside, California 92521

A magnitude estimation response procedure was used to evaluate the strength of visual-auditory intersensory bias effects under conditions of spatial discrepancy. Major variables were the cognitive compellingness of the stimulus situation and instructions as to the unity or duality of the perceptual event. With a highly compelling stimulus situation and single-event instructions, subjects showed a very high visual bias of audition, a significant auditory bias of vision, and a sum of bias effects that indicated that their perception was fully consonant with the assumption of a single perceptual event. This finding reopens the possibility that the spatial modalities function as a transitive system, an outcome that Pick, Warren, and Hay (1969) had expected but did not obtain. Furthermore, the results support the model for intersensory interaction proposed by Welch and Warren (1980) with respect to the susceptibility of intersensory bias effects to several independent variables. Finally, a new means of assessing intersensory bias effects by the use of spatial separation threshold was demonstrated.

Occasionally, a spatial discrepancy occurs between the visual and auditory sources of an event as, for example, when the visual portion of a movie is projected on the screen while the sound track emanates from a speaker off to one side. In situations such as this, one usually hears the sound to be located in the direction of the visual source, and one perceives no intersensory discrepancy. This is known as the "ventriloquism effect" (Howard & Templeton, 1965, p. 306), one of a number of perceptual phenomena involving various combinations of sensory modalities and referred to generically as "intersensory bias." The most common laboratory example of this class of perceptual events is "visual capture," in which the felt position of one's finger or hand is strongly biased in the direction of its prismatically displaced visual image (e.g., Hay, Pick, & Ikeda, 1965).

The ventriloquism effect has proved to be susceptible to a wide range of factors, including the cog-

nitive compellingness of the stimulus situation, the amount of discrepancy between visual and auditory sources, the relative cadence of the visual and auditory signals, and the reporting requirements (Welch & Warren, 1980). Unfortunately, the variety of reported methods for evaluating the strength of the effect precludes a valid comparison of much of this literature. Methods range from an all-or-none technique in which the subject is simply asked to judge whether the sources are in the same place (e.g., Jack & Thurlow, 1973) to one in which the subject indicates the precise position at which the object appears to be located (e.g., Pick, Warren, & Hay, 1969). Furthermore, in some studies a verbal report is used (e.g., Jackson, 1953), whereas in others a pointing response occurs, either with the subject's forefinger (e.g., Pick et al., 1969) or with a pistol-pointing device (e.g., Canon, 1970).

Pick et al. (1969) studied intersensory bias using all possible pairs of the three localization modalities: vision, proprioception, and audition. A major purpose of their experiment was to determine the extent to which the three modalities act as an integrated system with respect to the localization of dis-

crepant targets. It was hypothesized, for example, that if vision were found to bias proprioception [V(P)] strongly, and proprioception to bias audition [P(A)] strongly, then vision should be found to bias audition [V(A)] even more strongly. This pattern was not obtained: V(A) was found to be far weaker (47%) than was expected, given the strength of V(P), 64%, and of P(A), 60%. Pick et al. concluded from these results that the predicted transitivity had failed to occur and that the spatial modalities do not appear to act as a system in this respect.

Warren (1979) argued that the conclusion of nontransitivity may be premature, however, noting that the degree of compellingness of the various stimulus situations used by Pick et al. (1969) was probably not equivalent. Prior research has demonstrated clearly that compellingness significantly influences the ventriloquism effect. Jack and Thurlow (1973), for example, showed that the degree of synchrony between a speaking puppet's mouth and the speech ostensibly emitted by the puppet affected the strength of V(A), and Jackson (1953) found that V(A) is enhanced when the nature of the auditory signal (e.g., a whistle sound) is made to correspond to the apparent emitter of the signal (e.g., a visible whistle with a puff of steam emerging from it). On the other hand, less research has been done on the susceptibility of the V(P) effect to such factors, although recent work by Welch and Warren (Note 1) and Welch, Warren, With, and Wait (Note 2) suggests that these factors may be similarly involved. In any event, it is clear that before it may validly be concluded that the three spatial modalities do not operate as an integrated system, it will be necessary to use stimulus situations that are equated, as nearly as possible, with respect to their compellingness.

Welch and Warren (1980), in their review of the intersensory bias literature, drew a similar conclusion. In particular, they noted that the visual-proprioceptive situation used by Pick et al. (1969) provided the subject with a direct view of his own forefinger (the localization target). Since the subjects had, of course, had extensive experience with the identity of the feel and sight of their own forefingers, it is reasonable to assume that they were quite convinced that the two modalities were presenting information about a single event. On the other hand, the visual-auditory situation used by Pick et al. involved the view of a small earplug speaker with a plastic cone mounted on top and a series of repetitive clicks. The subjects had had no particular experience with this speaker and may have been less inclined to assume naturally that the visual and auditory information represented a single event.

In developing their model of perception under intersensory discrepancy conditions, Welch and Warren (1980) noted the need to take account of the cognitive compellingness of the stimulus situation.

They suggested that when presented with information from two modalities, the subject makes an assumption (usually unconscious) about whether the two information sources represent a single event or two separate events. This "unity assumption" (UA) is thus a cognitive variable, whose graded strength depends on both experience and various factors in the stimulus setting, including: (1) the number and importance of cues, such as spatial and temporal characteristics, that present redundant as opposed to discrepant information; (2) the observer's general and specific history with cue redundancies and/or discrepancies in similar situations; and (3) specific instructions or other situational factors that may strengthen or weaken the UA. If these factors produce a strong UA, then the subject's perceptual experience tends to occur in a way that is consonant with the presence of a single event. If the UA is weak, perception tends to be consonant with the presence of separate events. Specifically, if the UA is strong, then substantial intersensory bias effects should occur, and the subject should tend to report experiencing little or no discrepancy between the two modalities. In fact, if the UA is sufficiently strong, the subject should judge the event to be in the same location regardless of the modality for which he is asked to make his judgment. This would lead to the outcome in which the sum of the bias effects that the two modalities exert on one another should approximate 100%.

Applying this analysis to the Pick et al. (1969) results, the suggestion by Welch and Warren (1980) that the visual-auditory and visual-proprioceptive situations involved differing degrees of compellingness is supported. The visual-proprioceptive results suggested that a single event was perceived: V(P) ranged from 59% to 72% in several replications, and proprioceptive bias of vision [P(V)] was 16% to 40%. The sum of V(P) and P(V) ranged from 88% to 101%, or very close to 100%. In contrast, the V(A) values were 35% to 59%, the auditory bias of vision [A(V)] values were 1% to 15%, and the sum of V(A) and A(V) ranged from 49% to 62%, showing that the subjects were responding as though there were separate visual and auditory target locations. (There *was*, though, visual bias of audition: The visual event *did* affect the localization of the auditory event.) Informal reports presented by Pick et al. corroborate this picture: In a postsession interview, it was rare in the visual-proprioceptive situation for a subject to report experiencing a discrepancy between the seen and felt locations of his forefinger, whereas it was common in the visual-auditory situation for a subject to report experiencing a discrepancy between the seen and heard locations of the speaker.

Experiment 1 used the visual-auditory location discrepancy paradigm to study two factors hypothesized

by Welch and Warren (1980) to affect the UA and the perceptual outcome. One factor was compellingness. In the high-compellingness condition, the subject saw a person's face on a TV monitor and heard the person's voice via a separate (and invisible) speaker, a procedure used by Radeau and Bertelson (1977). In the medium-compellingness condition, the video and audio signals were the same, but a 150-msec lag was introduced between them, creating a temporal asynchrony. According to the formulation, this change should produce a situation that is less compelling for the subject. That is, he should be less compelled to the assumption that he is in the presence of a single perceptual event. In the low-compellingness condition, the same audio signal was used, but a small tape mark replaced the person's face on the video monitor. In this situation, the subject should be rather unconvinced that there is any necessary identity between the visual and the auditory information and thus should form only a weak UA.

The second factor, referred to as "unity instructions," had two levels and was completely crossed with the first. In the single-event condition, the subject was instructed that there was a single event and that an optical device was being used that could account for any possible perceived discrepancy between the visual and the auditory signals. Thus, the attempt was to produce a strong UA. In the dual-event condition, the subject was told that the audio and video signals could be positioned independently of one another at different locations. The goal was to cause the subject to assume that the events were not necessarily one and the same, thus making a weak UA.

The primary prediction was that the combination of a strongly compelling stimulus situation (the speaker's face and voice in synchrony) and the single-event instructions would produce V(A) and A(V) effects that would sum to 100%, as has been reported in previous research with highly compelling visual-proprioceptive stimulus situations, and that as a consequence there should be relatively little perceived spatial discrepancy between the auditory and visual stimuli. The UA should be weakened, however, by dual-event instructions and/or by a reduction in the compellingness of the stimulus situation, with the result that weaker bias effects should be found and, specifically, that the sum of V(A) and A(V) should not reach 100%.

In contrast with much of the previous localization discrepancy research that has used a pointing response, the present study made use of a magnitude estimation procedure. Such a procedure was used in the context of intersensory bias by Lederman (1979). Magnitude estimation has the advantage that it is modality free and thus avoids any possible contamination caused by interaction between the modalities involved in the discrepancy situation and the modality used for response.

EXPERIMENT 1

Method

Subjects. A total of 72 college student subjects were assigned to six groups of 12. Each group received a separate experimental situation, the six being constituted by crossing the two independent variables of compellingness (low, medium, high) and unity instructions (single event, dual event). An experimental session lasted about 30 min, and the subjects were either paid or given course-related credit for their participation.

Apparatus. The subject sat at a semicircular shelf with his head fixed into a chinrest and forehead restraint combination. The subject viewed the visible targets monocularly through a clear Plexiglas window that allowed a 60-deg field of view. A shutter prevented vision except when it was directly involved in the task. Beyond the perimeter of the shelf was another shelf on which a 9-in. TV monitor was placed. The monitor could be moved along the perimeter to a variety of target locations. The speaker in the monitor was deactivated, and the audio signal was fed ¨ through a separate speaker, mounted on a curved track at the perimeter of the shelf. Thus, the video and the audio signals , could be placed at independent locations in order to produce a spatial discrepancy. A discrepancy of 10 deg was used. The track-mounted speaker was hidden behind a cloth fence so that it was not visible to the subject. Thus, it was possible to give the subject an auditory-only target (for which the eyepiece shutter remained closed), a visual-only target (speaker not activated), or a visual + auditory target with a 10-deg discrepancy introduced. In this experiment, the displacement always occurred with the speaker placed 10 deg to the subject's left of the video location (thus approximating the optical effect of a 20-D base-left wedge prism). The target locations were at 0, ±5, ±10, ±15, and ±20 deg, where 0 represents straight ahead of the subject.

The session was conducted in the dark, so that the video screen provided the only source of illumination.

Procedure. A magnitude estimation procedure was used for all localization responses. The subject was instructed to use zero to denote the location straight ahead of his nose and to use positive numbers for locations to his right and negative numbers for locations to his left. No further guidance was given about the magnitudes of numbers to be used, and, in fact, subjects varied considerably in their spontaneous use of numbers. Before data collection was begun, a familiarization procedure was used in which the subject was presented with visual targets alone or auditory targets alone across the entire 40-deg range and was asked to make magnitude estimation judgments of location. Ordinarily, 15 practice trials were judged sufficient to ensure that the subject understood the procedure and was using his scale with good consistency.

Each subject participated in four localization tasks. The first two were control tasks, one with visual targets alone and the other with auditory targets alone. For each of these tasks, two targets were presented at each of the nine target locations, so that the subject made a total of 18 judgments. Half of the subjects in each group performed the visual control task first. The last two tasks both occurred in the presence of visual-auditory discrepancy. The target presentation was identical for the two tasks: The video and audio locations were set into place, displaced by 10 deg, and the shutter was lifted at the same time that the speaker was activated, so that the subject received the two discrepant sources simultaneously. After 3 sec, the shutter was lowered and the speaker simultaneously deactivated. Seven target location combinations were used, ranging in 5-deg steps from visual −10 deg/auditory −20 deg at one end to visual 20 deg/auditory 10 deg at the other end. Each target combination was used twice, for a

total of 14 trials in each task. The target order was randomized. In the visual conflict task, the subject was asked to give a number to indicate where he saw the speaker's mouth (or the spot, in the low-compellingness condition) and, in the auditory conflict task, to give a number to indicate where he heard the speaker's voice. Half of the subjects in each group performed the visual conflict task before the auditory conflict. Thus, there were, within each condition, four task order subgroups, each with three subjects. Since a preliminary analysis showed that there was no effect of or interaction with task order, this factor was not considered in subsequent analyses.

After the four localization tasks, two supplementary tasks were conducted to assess the magnitude of the subject's perceived visual-auditory discrepancy and unity assumption. In order to measure perceived discrepancy, the subject was asked to use his number scale to indicate how much discrepancy, if any, he experienced between the seen and heard target locations. A single response was recorded and later converted to degrees. In order to assess the strength of the unity assumption, the subject was asked: "When you were seeing the face (spot) and hearing the voice, how confident were you that you were hearing and seeing were actually the same thing? Use a scale from 1 to 7, where 7 indicates that you were very certain that what you heard and saw were the same, 4 indicates that you were completely uncertain, and 1 indicates that you were certain that what you heard and saw were different."

For the three compellingness conditions, the groups were treated identically except for the following differences in stimulus situation: In the high- and medium-compellingness conditions the videotape showed a person's face as he read the material, but in the high-compellingness condition the auditory stimulus was synchronous with the visual, whereas in the medium-compellingness condition the auditory stimulus lagged behind the visual by 150 msec, thus creating the situation that we may encounter in low-budget movies. In the low-compellingness condition, the TV screen remained illuminated but without picture, and a 1 × 2 cm piece of tape was placed on the screen at the location of the speaker's mouth in the other conditions.

The unity instruction conditions were differentiated as follows: In the single-event condition, after the subject performed the warm-up trials, the experimenter had the subject lean back while he inserted a clear Plexiglas window in the eyepiece, saying, "Now before we start let me put this into the goggles, like this." The control tasks were then conducted. In introducing the discrepancy tasks, the experimenter said: "Now I'll let you see the face (spot) and hear the voice at the same time. Using the same number scale, I'll ask you to indicate either where you hear it or where you see it. These goggles may displace your vision to one side or the other—so when you see the face (spot) and hear the voice, it may appear to be located other than where you hear it. It's really the same, though—any difference that you sense is because of the goggles." Then, before each of the specific discrepancy tasks, the experimenter said: "Remember that because of the goggles, where you see it and where you hear it may not seem the same. They really *are* the same, though—you are to indicate where you *see* (*hear*) it." Thus, the attempt in these instructions was to influence the subject to believe that a single event was involved and that any discrepancy should be attributable to the goggles (actually, of course, there was no optical device in the goggles, the discrepancy being created by the difference in locations between the monitor and the audio speaker). In contrast, in the dual-event condition, the attempt was made to convince the subject that there were separate events. This was done by showing the subject, in the introduction to the experimental session, the track on which the audio speaker traveled and showing him that the sound could be placed independently of the location of the video monitor. Then, in the instructions to the discrepancy tasks, the experimenter said: " . . . when you hear the voice and see the mouth (spot), it

may appear to be located other than where you hear it—that's because I can move the speaker separately from the location of the TV screen."

Results

Calculation of percent bias effects. Percent bias effects were calculated for individual subjects as in previous research (e.g., Pick et al., 1969). In essence, the no-conflict control tasks are used as points of reference against which responses in the conflict tasks are assessed. For the present paradigm, the mean visual control value (V_c) was the mean of the magnitude estimates for the visual targets occurring from -10 to $+20$ deg target locations. The mean auditory control value (A_c) was the mean of the magnitude estimates for the auditory targets occurring from -20 to $+10$ deg target locations. (In each case, this was the range in which that type of target occurred during the conflict tasks.) The mean magnitude estimate for the visual conflict task (V_e) represented the subject's judgment about visual location when discrepant auditory information was also available. The influence, or percent bias, of audition on visual localization [A(V)] was calculated for each subject by the formula:

$$A(V) = \frac{V_c - V_e}{V_c - A_c} \times 100.$$

Similarly, the mean magnitude estimate for the auditory conflict task (A_e) represented the subject's judgment about auditory location when discrepant visual information was also available. The percent bias of vision on auditory localization [V(A)] was calculated for each subject by the formula:

$$V(A) = \frac{A_e - A_c}{V_c - A_c} \times 100.$$

The mean V(A) and A(V) values for each of the six conditions appear in Table 1, with 95% confidence intervals noted for each mean.

Sum of V(A) and A(V). As noted earlier, the Welch and Warren (1980) model makes certain predictions about the sum of two complementary bias effects. These sums appear in Table 1. The sum of V(A) and A(V) for each subject was entered into a two-way ANOVA with instructions and compellingness as between-groups factors. The effect of instructions was significant [$F(1,66) = 10.17$, $p < .01$], with the greater sum occurring in the single-event condition. Compellingness also had a significant effect [$F(2,66) = 9.40$, $p < .01$]. The high-compellingness condition showed a greater sum of bias effects than did the medium and low conditions, which were not apparently different from one another. However, these main ef-

Table 1
Experiment 1: Mean Bias Effects, V(A) and A(V), and Mean
Sum of Bias Effects, All in Percent, With 95%
Confidence Intervals (CI)

	Compellingness					
	High		Medium		Low	
	Mean	CI	Mean	CI	Mean	CI
	Single-Event Instructions					
V(A) + A(V)	96.8	14.2	40.2	11.5	31.9	21.4
V(A)	78.9	12.5	36.8	14.9	35.1	23.1
A(V)	17.1	10.5	3.3	7.5	-4.0	9.3
UA	4.6		3.8		2.9	
PD	.1		1.6		4.9	
	Dual-Event Instructions					
V(A) + A(V)	35.5	18.8	31.0	18.3	35.4	13.1
V(A)	22.2	17.3	19.7	16.5	18.0	13.2
A(V)	13.3	13.0	11.3	4.7	17.4	9.6
UA	4.2		3.8		3.1	
PD	2.8		2.9		3.3	

Note—Mean unity assumption (UA) and perceived discrepancy (PD) judgments also appear. UA is measured by a 1-7 scale, and PD is represented in degrees.

fects must be viewed in the context of a significant interaction of the instructions and compellingness factors [F(2,66) = 8.38, p < .01]. It is evident from Table 1 that the interaction was produced largely by the difference between the high-compellingness/single-event group (mean = 96.8%) and the remaining five groups, which averaged about 35%. Among the latter five groups, a comparison of the highest with the lowest showed no difference (F = 1.17).

It is clear that, as hypothesized, the combination of both experimental factors produced a large sum of bias effects. In fact, the mean of 96.8% for the high-compellingness/single-event group was not significantly different from 100%. This means that these subjects indicated the same location whether they were asked to localize using the auditory information or the visual information. Mean V(A) was greater (79.8%) than mean A(V) (17.1%), showing that the visual information was given substantially more weight than the auditory information in arriving at the perceptual resolution of the conflicting visual and auditory location information. The important point, however, is that the sum of bias results was such as to indicate that a successful perceptual resolution had occurred, that is, that the perceptual outcome was consonant with the perception of a single event rather than of two events.

A further prediction of the model is that the single-event perceptual outcome should be accompanied by a strong UA and by low perceived discrepancy. Table 1 contains the mean scores for these measures: the single-event/high-compellingness condition had the highest mean UA and the lowest perceived discrepancy of all conditions.

Thus, the situation for subjects in this group may be described as follows. Because of the single-event instructions and the high degree of compellingness, the subjects made the assumption that they were perceiving but a single event (despite the 10-deg imposed discrepancy), and they therefore responded in accordance with this assumption, both in reporting a single location regardless of modality and in perceiving no discrepancy between the visual and auditory sources of information.

The difference in sum of bias scores between this group and the remaining five is striking, as is the similarity among the remaining five. In brief, reducing the compellingness of the stimulus situation *or* giving instructions about the possibility of dual events produced a dramatically decreased sum of bias effects. That is, these subjects responded to different locations when asked for visual or for auditory judgments. At the same time, though, there was no group in which the sum of bias effects was not greater than zero, indicating that even the strongest attempt to destroy the subject's unity assumption (the low-compellingness/dual-event condition) did not eliminate intermodality bias effects entirely. Whether the effects could be abolished by other operations, such as placing more stimulus qualities into discrepancy, remains unanswered by the present data.

The very high sum of bias effects in the high-compellingness/single-event condition is noteworthy in its similarity to the results obtained previously (e.g., Pick et al., 1969) from the visual-proprioceptive situation in which the subject views his forefinger as the target, where the sum of V(P) and P(V) was not significantly different from 100%. As Welch and Warren (1980) pointed out, the visual-proprioceptive situation in which the finger is viewed is an extremely compelling one, and it is not surprising that the subject's perception is consonant with a single-event situation. The current visual-auditory results from the highly compelling condition are clearly more like the Pick et al. (1969) visual-proprioceptive than the visual auditory results, where the sum of V(A) and A(V) was far less than 100%. Thus, it may be concluded that, given a high degree of visual-auditory compellingness, together with instructions concerning the unity of the event, the sum of V(A) and A(V) can reach its theoretical maximum, as suggested by Warren (1979) and Welch and Warren (1980).

V(A). Mean V(A) values appear in Table 1. The scores were submitted to a two-way ANOVA identical to that used for the bias sum. The effect of instructions was significant [F(1,66) = 19.57, p < .01], with the single-event instructions producing greater bias. The effect of compellingness was also significant [F(2,66) = 5.17, p < .01], with the entire effect apparently due to higher bias in the high-compellingness condition than in the medium and low conditions.

However, the interaction of instructions and compellingness was also significant [F(2,66) = 3.80, p < .01]. The interaction was due primarily to the difference between the high-compellingness/single-event group (78.9%) and the others (18% to 37%). Although single-event instructions appear to produce higher bias for all three compellingness conditions, tests of simple main effects of instructions showed no significant differences at the low- and medium-compellingness levels. Thus, V(A) was very high, although significantly less than 100%, in the high-compellingness/single-event condition and uniformly lower in the remaining conditions, although always significantly greater than 0. The value of 79% in the high-compellingness/single-event condition is clearly beyond the range of 35% to 59% found by Pick et al. (1969) for V(A); in fact, it is higher than the V(P) values in the Pick et al. study, which ranged from 59% to 72%. It should also be noted that the mean V(A) in this condition was far higher than the mean A(V) in the same condition, indicating that the visual information is relied upon far more heavily than is the auditory information in this situation.

A(V). Mean A(V) values also appear in Table 1. Data were analyzed by means of two-way ANOVA. The effect of instructions was significant [F(1,66) = 4.71, p < .05], with the results showing less bias in the single-event conditions. The effect of compellingness was not significant [F(2,66) = 1.93], but the interaction of instructions and compellingness was significant [F(2,66) = 3.40, p < .05]. The interaction results from the fact that A(V) was relatively stable and significant in the 11%-to-17% range with dual-event instructions, while A(V) was not significant in the low- and medium-compellingness conditions but was significant (17%) in the high-compellingness condition.

The A(V) findings are provocative. We note, first, that Pick et al. (1969) reported A(V) means ranging from 1% to 15%, with no mean value significantly different from zero. The present A(V) mean values showed a similar range, from −4% to 17%, but four of the six group means were significantly greater than 0. Several questions occur about these small but nonetheless statistically significant effects. First, why was there a significant A(V) effect under dual-event instructions in all conditions of compellingness, while A(V) tended not to be significant in the single-event conditions? It seems likely that the significant A(V) effects in the dual-event groups occurred because the subjects in these groups, assuming that there were separate events, tended to divide their attention between the two modalities that were receiving information. This divided attention creates the situation that leads to intersensory bias effects, according to the Welch and Warren (1980) model.

This suggestion, though, does not account for the significant A(V) effect (17%) in the single-event/high-compellingness condition. We have no ready expla-

nation for this result. We note, though, that this group differed from all the others in showing a sum of bias effects that was not significantly less than 100%. In this respect, this group showed results akin to those from previous research (e.g., Pick et al., 1969), in which a 100% sum of bias effects was found for the highly compelling stimulus situation in which the subject views his finger through a wedge prism. In that situation, a significant bias of vision by proprioception has been routinely found.

EXPERIMENT 2

The most important results in Experiment 1 are the demonstration that V(A) and A(V) summed to 100% in the single-event/high-compellingness condition and that V(A) was dramatically higher in that condition than in the work reported by Pick et al. (1969). However, besides differing in compellingness and instructions from previous work, Experiment 1 differed in that it made use of a magnitude estimation procedure rather than manual target pointing. It is thus necessary to examine the possibility of a response procedure artifact. Consequently, in Experiment 2, the single-event/high-compellingness condition was repeated exactly, with the exception that a pointing response was used rather than magnitude estimation. Twelve subjects were tested.

The results of Experiment 2 proved to be more similar to those of the corresponding condition from Experiment 1 than to previous results that used pointing responses. Specifically, the mean V(A) with pointing was 79.5% ± 11.5%, virtually identical to the 78.9% from Experiment 1. The mean A(V) was 7.3% ± 4.8%, significantly greater than 0 and not significantly different from the mean of 17.7% found with the magnitude estimation procedure. Finally, the mean sum of V(A) and A(V), 86.8% ± 9.7%, was not significantly different from the corresponding sum in Experiment 1. It should be noted, however, that the 95% confidence interval around the mean of 86.8% does not extend to 100%, suggesting that in this condition the subjects' responses indicated slightly different locations when visual and auditory response requests were made. The mean UA judgment was also somewhat lower, 3.8 as compared with 4.6 in Experiment 1.

In brief, the results of this experiment show considerable similarity to the magnitude estimation group with high-compellingness and single-event instructions, and thus it is evident that the striking results found in that group cannot be attributed to the response procedures.

EXPERIMENT 3

Stimulus redundancy across participating modalities is hypothesized to affect the subject's assumption of unity and therefore the nature of intermodal-

ity bias effects. The high-compellingness condition used in Experiments 1 and 2 was designed to include a very important redundancy, temporal correspondence, and, in fact, high UA and the related strong bias were found, as predicted. It should be noted that the subjects responded as if in the presence of but a single event despite the occurrence of a significant nonredundancy, the spatial discrepancy itself. The 10-deg spatial discrepancy apparently did not represent a nonredundancy of sufficient consequence to overcome the UA-producing influence of temporal redundancy and the cognitive identity of the speaker's voice and face. Presumably, if the spatial discrepancy were made larger, a point would be reached beyond which the redundancies present in the situation are no longer strong enough to counteract the discrepancy. The UA should be found to weaken and the sum of bias effects should be significantly less than 100%.

In Experiment 3, the same high-compellingness/single-event condition was used as in Experiment 1, but the discrepancy was increased to 20 deg. Twelve subjects were tested.

The results were as expected. The mean sum of V(A) and A(V) was 48.8% ± 17.7%, or significantly less than 100%. V(A) averaged 50.8% ± 18.2%, while A(V) averaged −2.1% ± 3.3%. It is notable that A(V) was not significantly greater than 0, in contrast with the 17% finding in Experiment 1. Clearly, with the larger spatial discrepancy, UA strength was lessened (to 4.2), as was, correspondingly, the subject's ability to perceive as if in a single-event situation. Furthermore, perceived discrepancy was 8.6 deg, as compared with .1 deg in the corresponding 10-deg discrepancy condition of Experiment 1.

EXPERIMENT 4

In Experiment 4, a new approach was taken to the exploration of visual-auditory interaction, one involving the determination of spatial separation thresholds.

Method
Six subjects participated in each of five conditions of visual-auditory discrepancy. For each of the conditions, the order of which was nonsystematic, the visual and auditory signals were presented simultaneously, with the spatial separation of the signals varied from trial to trial. A forced-choice, paired comparisons approach was used, in which the subject was required to report whether the auditory signal was to the left or to the right of the visual signal. The spatial separations were varied in 2-deg steps from 0- to 8-deg separation in each direction. Each subject received 20 trials at each separation in each condition. Thus a subject's threshold for a condition was based on 180 trials, administered in unsystematic order within a single session. No more than one session was administered to a subject on a single day. The five conditions varied in terms of their presumed compellingness. They were as follows:

Voice and mouth. The auditory signal was a recorded passage read from a perception textbook by a male voice. The visual signal was the videotaped view of the speaker's face, whose mouth could be seen to move synchronously with the auditory signal. The speaker's face occupied virtually all of the 9-in. monitor screen, with the mouth in the center of the screen. As in the preceding experiments, this was presumed to be a highly compelling situation.

Voice and mouth, asynchronous. This was like Condition 1, except that the auditory signal was delayed by 150 msec, so that a distinct asynchrony between the visual and auditory signals was produced.

Voice and spot. The auditory signal was the same male voice, but the visual signal was a vertically aligned 1 × 2 cm tape mark on the face of the video screen. The speaker's face did not appear on the screen.

Click and spot. The auditory signal was a repetitive clicking sound, while the spot was as described in Condition 3.

Click and mouth. The auditory signal was the repetitive click (as in Condition 4), and the visual signal was the speaker's face and mouth (as in Conditions 1 and 2).

Results
A threshold was calculated for each subject in each of the five conditions by plotting the percentage of "auditory to the right of visual" responses against the spatial separation, and determining the spatial separation of the 25% and 75% intersections. It was predicted that the more compelling the visual-auditory condition, the farther apart the two stimuli would need to be placed before the observer could reliably perceive the direction of their separation.

The mean thresholds, in degrees, for the five conditions, with standard deviations in parentheses, were, respectively, 4.6 (.74), 3.9 (1.30), 3.2 (1.27), 3.1 (1.12), and 3.2 (1.44). These results were just as expected: Condition 1 was designed to be highly compelling, and the spatial separation required for the subject to be able to determine reliably the relative locations of the visual and auditory signals was the greatest of all the conditions. The asynchronous Condition 2 was expected to be of intermediate compellingness, and indeed the threshold was somewhat lower than that of Condition 1. The remaining three conditions were still lower and not apparently different from one another in degree of compellingness.

Thus, the experiment demonstrated that variations in compellingness—created in an a priori manner, but along the lines suggested by Welch and Warren (1980) to be important in affecting the unity assumption in discrepancy work—have predictable effects on the spatial discriminability of auditory and visual signals.

DISCUSSION

The primary points of discussion have been made earlier but bear repeating here. In contrast with the results found by Pick et al. (1969), the present study found the sum of V(A) and A(V) bias effects to be approximately 100% in the high-compellingness/

single-event condition, suggesting that a unified perceptual experience occurred as it does in the highly compelling visual-proprioceptive finger-viewing situation. Thus, it is very likely that previous evidence suggesting an apparent qualitative difference between the visual-auditory and visual-proprioceptive discrepancy situations was a result of different degrees of compellingness. This new finding revives the possibility that the visual-auditory and visual-proprioceptive situations might, after all, fit within a single overall framework, as initially proposed by Pick et al. Furthermore, the strong V(A) expected but not obtained by Pick et al. in their experiment was found in the current study, as was a significant A(V) effect, paralleling the P(V) effects found in highly compelling visual-proprioceptive conflict situations.

It is also useful to compare the present work with the program of research on visual-auditory effects reported by Radeau and Bertelson (1977). In their Experiment 3, they report the frequency of subjects' reports of the spatial correspondence of visual and auditory stimuli that were, in fact, spatially separated by 20 deg. Reports of spatial correspondence were called "fusion" reports. The relationship of the fusion phenomenon to the intersensory bias phenomena is of interest. Fusion is that situation in which the sources, although actually discrepant, are perceived as the same. This perceptual situation would appear to be the same as that which we propose produces a high sum of intersensory bias effects and no perceived discrepancy—that is, a high unity-assumption situation like our single-event/high-compellingness condition. Interfering in a significant way with the conditions that produce a high unity assumption, such as we did by introducing dual-event instructions and/or lower degrees of compellingness, should lead to the loss of fusion, or, in our terminology, a sum of bias effects that is significantly less than 100%. Radeau and Bertelson, indeed, found the frequency of fusion reports to decrease significantly as the temporal synchrony of the visual and auditory events was violated and as the meaningfulness of the visual-auditory stimulus relationship was decreased. Our "fusion" (100% sum of bias effects), similarly, was disrupted by loss of temporal synchrony of visual and auditory stimuli and/or by the introduction of dual-event instructions.

We regard fusion, then, as the perceptual result of the situation in which the conditions lead the subject to make a strong unity assumption and to perceive in accord with that assumption—that is, a single event. The determination of the magnitude of bias effects would seem to offer a more fine-grained approach to intersensory interaction than does the assessment of fusion, since fusion is assessed simply as present or not present, whereas bias effects can be evaluated over a wide range of graded strength.

REFERENCES

Canon, L. K. Intermodality inconsistency of input and directed attention as determinants of the nature of adaptation. *Journal of Experimental Psychology*, 1970, **84**, 141-147.
Hay, J. C., Pick, H. L., Jr., & Ikeda, K. Visual capture produced by prism spectacles. *Psychonomic Science*, 1965, **2**, 215-216.
Howard, I. P., & Templeton, W. B. *Human spatial orientation.* New York: Wiley, 1966.
Jack, C. E., & Thurlow, W. R. Effects of degree of visual association and angle of displacement on the "ventriloquism" effect. *Perceptual and Motor Skills*, 1973, **37**, 967-979.
Jackson, C. V. Visual factors in auditory localization. *Quarterly Journal of Experimental Psychology*, 1953, **5**, 52-65.
Lederman, S. J. Auditory texture perception. *Perception*, 1979, **8**, 93-103.
Pick, H. L., Jr., Warren, D. H., & Hay, J. C. Sensory conflict in judgments of spatial direction. *Perception & Psychophysics*, 1969, **6**, 203-205.
Radeau, M., & Bertelson, P. Adaptation to auditory-visual discordance and ventriloquism in semirealistic situations. *Perception & Psychophysics*, 1977, **22**, 137-146.
Warren, D. H. Spatial localization under conflict conditions: Is there a single explanation? *Perception*, 1979, **8**, 323-337.
Welch, R. B., & Warren, D. H. Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, 1980, **88**, 638-667.
Young, P. T. Auditory localization with acoustical transposition of the ears. *Journal of Experimental Psychology*, 1928, **11**, 399-429.