

Multimodal signal detection: Independent decisions vs. integration

ROBERT M. MULLIGAN and MARILYN L. SHAW
Rutgers—The State University, New Brunswick, New Jersey 08903

The pooling of information from simultaneous, spatially congruent auditory and visual stimuli is examined in a signal detection task. The paradigm used permits discrimination among a number of models of the decision mechanisms involved in processing multiple component stimuli. Parameter-free predictions are presented for the weighted integration model and for three versions of the independent decisions model. The data support an independent decisions model of the bimodal detection process in which attention is shared equally between modalities.

The manner in which observers combine information from multiple sources has been a traditional topic of information processing research. Green and Swets (1966), for example, have discussed results of a number of experiments investigating the pooling of information from multiple observations, including those using multiple component signals, those employing multiple observer teams, and those using repeated stimulus presentations. The present paper examines this problem in a signal detection experiment using simultaneous inputs to two sensory modalities as the multiple sources of information.

Several formal models have been developed in an attempt to describe the decision processes involved in pooling information from several sources. Two major classes of models have emerged, *independent decisions* and *integration*. According to the independent decisions model, a separate detection decision is made about each source of information and the results of these decisions are then combined to make a judgment. For example, in a signal detection paradigm, subjects report the presence of a signal whenever there are a sufficient number of positive detection decisions.

In the *integration model*, the observer sums the sensory input values at a predecisional stage of processing and then reports the presence of a signal if this aggregate value exceeds its criterion. Thus, the final decision is based on the sum of the multiple observations rather than on an evaluation of several independent decisions.

A good deal of data has been reported, from both

This paper is based on part of a dissertation submitted by the first author to Rutgers—The State University in partial fulfillment of the requirements for the PhD degree. The research was supported by NSF Grant BNS 77-26296 to the second author and by a Rutgers University Research Council Grant (1977-78) to H. R. Schiffman, whose generous contribution of resources is gratefully acknowledged.

visual and auditory information processing experiments, in support of each of these two models. However, much of these data are from studies that have considered only one of the alternative models; for example, Burns (1979), Gardner (1973), Rumelhart (1970), Shaw and Shaw (1977), and Shiffrin and Geisler (1973) fit their data to an independent model, while Kinchla (1977) and Kinchla and Collyer (1974) found evidence for an integration model. From the relatively few studies in which data have been compared with both models (e.g., Green, McKey, & Licklider, 1959; Green & Swets, 1966; Wickelgren, 1967), some evidence has been generated for each. Shaw (Note 1) has recently reported strong evidence in favor of independent decisions based on experiments in which the multiple-component stimuli consisted of letters presented at two locations. In two experiments, she tested predictions of an integration model and three versions of the independent decisions model. The results supported independent decisions, with some subjects performing according to a version in which attention is shared equally between locations on each trial and others according to a version in which, on a given trial, attention is devoted exclusively to one of the locations. All of the models described by Shaw, and their predictions, are presented in greater detail below.

Relatively few experiments have examined the pooling of information across sensory modalities. The design used in these studies is fairly consistent. Signals are presented in three types of trials: unimodal auditory, unimodal visual, and bimodal trials on which auditory and visual signals of comparable detectability are presented simultaneously.

Although the designs are similar and all the studies have shown a gain in detectability of bimodal relative to unimodal signals, the models to which the data have been fit and the conclusions drawn show little agreement. Eijkman and Vendrik (1965) and Fidell (1970) fit their data to integration models in which

processing on auditory and visual channels is independent. Brown and Hopkins (1967) reported data from a bimodal detection task which supported a type of independent decisions model, although, as Morton (1967) has pointed out, the model they derived lacks generality. Finally, Craig, Colquhoun, and Corcoran (1976) obtained results supporting neither an integration model nor either of two threshold-independent decisions models. Their use of a vigilance-type task (a priori signal probability was only .05) and the averaging of data over 18 subjects, however, make Craig et al.'s results difficult to evaluate and compare with those of other investigations.

To summarize, although there is agreement that auditory and visual processing is not correlated in a bimodal detection task, these few studies have not been in agreement as to which model best describes the decision processes involved in combining information from the two modalities. None of these previous studies has compared continuous versions of the independent decisions and integration models in the same experiment. Using the paradigm described below, these models will be tested in the present study.

Before formally stating the independent decisions and integration models to be tested here, it will be useful to outline briefly the paradigm to be used. The task employed is a variant of a general paradigm described by Shaw (Note 1). All signals are presented against a background of independent auditory and visual noise. On each trial, the stimulus, S_i , consists of either a signal-plus-noise pattern or noise only. These two possibilities, combined with the two modalities, provide for four different stimulus patterns. Accordingly, the subscript "i" can take on the value "a," "v," "b," or "n" for unimodal auditory, unimodal visual, bimodal, and noise-only trials, respectively.

The dependent variable, P_i , represents the conditional probability of a "no" response given the stimulus pattern S_i . Thus, P_v is the probability of a "no" response given that a visual signal has been presented. The probability of a "no" response is used in order to simplify the expression of predictions.

THE FORMAL MODELS

In the continuous view of the detection process, each stimulus presentation is assumed to evoke a sensory response that can be represented by a random variable. In the present context, two random variables, X_{aj} and X_{vj} , which refer to the internal representation of auditory and visual modalities, respectively, will be designated. The second subscript, j , can take on a value of "0," indicating a noise-only trial, or "1," indicating a signal plus noise trial. For example, the pair X_{a0} and X_{v1} represent the values of the random variables for the auditory and visual modalities on a

unimodal visual trial. For each random variable there are two probability distributions, one for signal plus noise trials and one for noise-only trials. In the models described below, the only assumptions made regarding these four distributions is that they have finite means and variances.

The Independent Decisions Model

In a multimodal detection task, this model states that an independent decision is made for each modality about the presence of a signal. In each case, the decision will be positive if the value of the random variable, X_{ij} , exceeds the decision criterion, β_i , for that modality. The overall decision rule is inclusive-disjunctive; that is, if either decision is positive, the subject responds "yes" and if both are negative, the response is "no."

Three different versions of the independent decisions model have been developed (Shaw, Note 1): the sharing model, the Type 1 mixture model, and the Type 2 mixture model. Each of these three versions generates a set of testable predictions.

The sharing model. In this version of the independent decisions model, the subject is seen as dividing or sharing attention between modalities on each trial. The proportion of attention assigned to each modality and the decision criteria, β_a and β_v , are assumed to remain constant across trials. Thus, P_i , the conditional probability of a "no" response given S_i , is simply the product of the probabilities that each of the random variables will not exceed its corresponding criterion:

$$P_i = P(X_{aj} < \beta_a)P(X_{vj} < \beta_v). \quad (1)$$

Two predictive equations can be derived from the model. First, it can be shown that (Shaw, Note 1):

$$P_a + P_v < P_n + P_b. \quad (2)$$

Furthermore, after logarithmic transformation, the following equality can be derived:

$$\ln P_a + \ln P_v = \ln P_n + \ln P_b. \quad (3)$$

The mixture models. The second version of the independent decisions model holds that processing is characterized by a mixture of separate rules for allocating attention to each modality. On a given trial, the subject attends primarily to one modality or the other in a probabilistic fashion. The criteria are assumed to differ as a function of which modality is given primary attention. The parameters of X_{aj} and X_{vj} may be the same or different. Overall performance is a weighted mixture of that which would be expected if each separate rule was used exclusively. Two types of mixture model will be considered below.

The Type 1 mixture model. In this model (hereafter called mixture 1), information on each trial is obtained from only one modality. If the probability that attention is directed to the auditory modality equals α , and the corresponding probability for visual modality equals $(1 - \alpha)$, then P_i for this model is:

$$P_i = \alpha P(X_{aj} < \beta_a) + (1 - \alpha) P(X_{vj} < \beta_v). \quad (4)$$

In this model, the unattended modality has no influence on the response probabilities.

From Equation 4, it can be shown that:

$$P_a + P_v = P_n + P_b. \quad (5)$$

The Type 2 mixture model. This model differs from the Type 1 mixture model in that some information from the unattended modality is used in the detection decision. Attention is directed primarily to one modality with probability α and to the other modality with probability $1 - \alpha$, such that the conditional probability of a "no" response is now:

$$P_i = \alpha P(X_{aj} < \beta_a) P(X_{vj} < \beta'_v) + (1 - \alpha) P(X_{aj} < \beta'_a) P(X_{vj} < \beta_v), \quad (6)$$

where β_i is the criterion for a modality when attention is directed primarily to that modality and β'_i is the criterion when attention is directed primarily to the other modality.

The inequality in Equation 2 also holds for the Type 2 mixture model. This model differs from the sharing model, however, in that the logarithms of the response probabilities do not sum as in Equation 2; that is, for the Type 2 mixture model:

$$\ln P_a + \ln P_v \neq \ln P_n + \ln P_b. \quad (7)$$

As has been noted previously (Shaw, Note 1), the predictions of all three independent decisions models are free of any effects of attention on the parameters of X_{aj} and X_{vj} or on their corresponding decision criteria. Also, these predictions hold only when subjects are required to report signals from both modalities.

If the subjects are instructed to attend to only one modality, all three models make the same predictions.

The three independent decisions models and their predictions in terms of P_i are summarized in Table 1. The characteristics and predictions of the weighted integration model, to be discussed next, are also given.

The Weighted Integration Model

This version of the integration model, introduced by Kinchla and Collyer (1974), allows for the possibility that information from separate sources might be differentially weighted in the integration process. Translated into the multimodal detection situation, the model holds that the detection decision is not based directly on the values of the internal random variables generated in each modality. Rather, these values are summed to produce the decision random variable, Y_i :

$$Y_i = wX_{aj} + (1 - w)X_{vj}. \quad (8)$$

The internal representations in each modality are assigned weights, "w" and "1 - w" in the two modality cases. The subject is assumed to respond "no" whenever Y falls below some criterion value, so that:

$$P_i = P(Y_i < \beta). \quad (9)$$

If the probability distributions of the internal random variables are assumed to be Gaussian, as they have been by Kinchla (1977; Kinchla & Collyer, 1974), then the weighted integration model predicts additivity of the transformed response probabilities. The transformation called for by this assumption is the inverse Gaussian, commonly known as the z score. Thus, the integration model makes the following prediction regarding response probabilities:

$$z_a + z_v = z_n + z_b, \quad (10)$$

where z is the inverse Gaussian or z transformation of the probability of a "no" response, P_i . It should be noted that Equation 10 holds only when the variances of the signal and noise distributions are equal.

Table 1
Contrasting Predictions for Probability of a "No" Response

Model	Equation	Prediction			
		1	2	3	4
Sharing Model	$P_i = P(X_{aj} < \beta_a) P(X_{vj} < \beta_v)$	N	Y	Y	N
Type 1 Mixture Model	$P_i = \alpha P(X_{aj} < \beta_a) + (1 - \alpha) P(X_{vj} < \beta_v)$	Y	N	N	N
Type 2 Mixture Model	$P_i = \alpha P(X_{aj} < \beta_a) P(X_{vj} < \beta'_v) + (1 - \alpha) P(X_{aj} < \beta'_a) P(X_{vj} < \beta_v)$	N	N	Y	N
Weighted Integration	$P_i = P(wX_{aj} + (1 - w)X_{vj} < \beta)$	N	N	N	Y

Note—N = no; Y = yes. Prediction 1: $P_a + P_v = P_n + P_b$. Prediction 2: $\ln P_a + \ln P_v = \ln P_n + \ln P_b$. Prediction 3: $P_a + P_v < P_n + P_b$. Prediction 4: $z_a + z_v = z_n + z_b$.

TESTING MODEL PREDICTIONS

Predictions of the four models are summarized in Table 1. Three of the models predict additivity of response probabilities, each under a different transformation: logarithmic for the sharing model, identity (no transformation) for the Type 1 mixture model, and the inverse Gaussian for the weighted integration model. These predictions are graphically illustrated in Figure 1.

The Type 2 mixture model is different from the others in that it does not predict additivity under any transformation. This model will be supported if, in the obtained data, the sums of the logarithms are not equal and if $P_a + P_v < P_n + P_b$.

By generating predictions of the models at a number of accuracy levels, Shaw (Note 1) has shown that the models are most readily discriminated at high accuracy levels, that is, when P_n is high (.85 to .95) and P_b is low (.05 to .15). Accuracy levels within

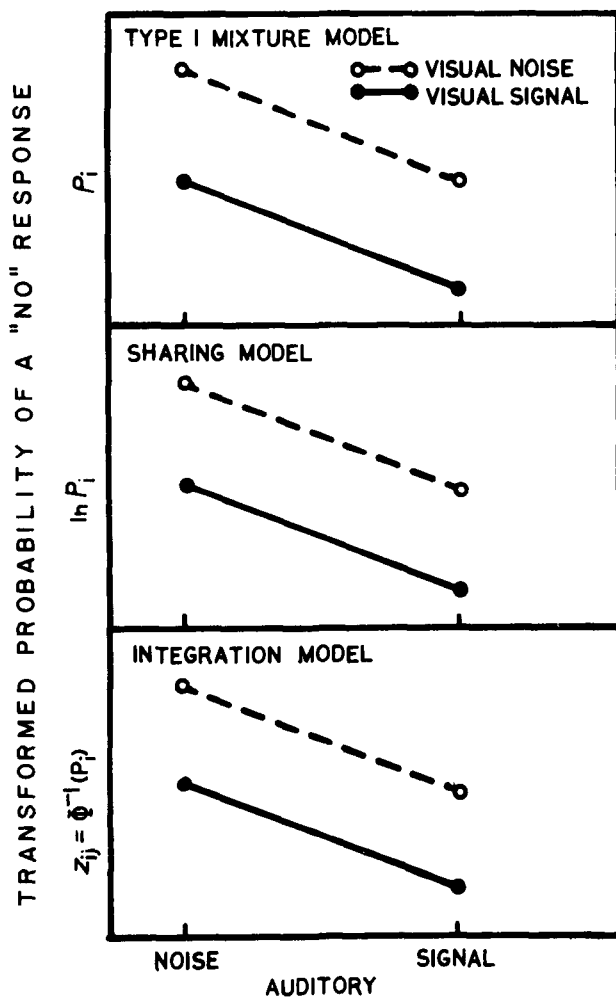


Figure 1. Additivity of transformed response probabilities predicted by three models.

this range were achieved in the present experiment by adjusting stimulus intensity.

EXPERIMENT

This experiment was designed to test the model predictions described above in a multimodal detection situation. In selecting the type of stimuli to be used, consideration was given to creating a bimodal signal condition in which optimal combinations of auditory and visual information could be realized. One potentially important factor, not considered in previous studies of bimodal effects, is the spatial relationship of the component stimuli. Previous studies of this sort have presented the auditory signal over earphones so that this signal's spatial location was not typical. This is in contrast to the visual signal which has been presented under more typical conditions, that is, at a specific spatial locus at some distance from the observer. In the present study, auditory signals were presented via a loudspeaker placed adjacent to the visual stimulus so that the two shared roughly the same spatial locus. An auditory frequency of 500 Hz was chosen to facilitate localization.

In addition, these spatially congruent stimuli were peripherally (40 deg azimuth) rather than centrally located. The rationale for this decision involved the fact that, for the visual system at least, the sensory mechanisms that respond maximally to peripherally located stimuli are specialized for detection, as contrasted with those sensitive to centrally located (foveal) stimuli, which are specialized for pattern recognition (e.g., see Schneider, 1969; Weiskrantz, Warrington, Sanders, & Marshall, 1974). Since this peripherally sensitive system is maximally sensitive to changes in stimulus energy, an intermittent stimulus, consisting of a string of four 125-msec bursts, was used rather than a single burst of constant intensity.

In terms of the models in question, it might be argued that these spatially coincident bimodal stimuli would favor integration of information or facilitate correlated processing of auditory signals. This type argument, however, depends upon having fairly specific knowledge of the neural mechanisms involved in combining sensory information. Although there is some evidence from sensory physiology for the existence of bimodal neural mechanisms at the single-cell level (e.g., Gordon, 1972), the relative contribution of these cells to stimulus detection is undetermined. We can say, at least, that these spatially specific stimuli are more typical of those that might be encountered routinely than are the stimuli that have been used in other studies.

Method

Subjects. Two male and two female Rutgers University under-

graduates served as subjects. All had normal or near-normal acuity in each eye. Their auditory thresholds for .125-8.0 kHz were within 10 dB of audiometric zero in each ear (International Standards Organization, 1964). Also, none of the subjects had any apparent visual or auditory deficit or asymmetry which might impair perception of stimuli presented in the spatial periphery. Each was paid \$2.50/h plus a variable, small incentive bonus for correct responses.

Apparatus. Pure tone signals, generated by an audio oscillator (Hewlett-Packard, 200AB), were shaped by an electronic switch (Grason-Stadler, 829D) and interval timer to provide a "signal" that consisted of four 125-msec bursts of a 500-Hz tone (250-msec repetition period, 50% duty cycle, 10-msec rise-decay time). The pulsed signals were then directed through an amplifier (Advent, 301), attenuator (Hewlett-Packard, 350D), transformer, and a mixing network in which they were added to a 20-kHz low-pass noise (Grason-Stadler, 455C) before being delivered through a KLH Model 12.5 loudspeaker.

The audio signal was split at the amplifier and directed to a voice-operated relay. Relay closure triggered the lamp timer and driver unit of a Scientific Prototype Model 800F two-channel tachistoscope, permitting simultaneous onset of light and sound signals. The tachistoscope was programmed so that one channel remained on continuously (providing "visual noise") and the second channel was driven remotely by the audio signals. The timer of this later channel was set at 125 msec so that the light and tone pulses would be equal in duration. Separate switches in the audio and lamp circuits permitted the experimenter to arrange unimodal or bimodal signal presentation before each trial.

With the exception of the transducers, all of these components were contained in a soundproof control room. The lamps and loudspeakers—the bimodal stimulus array—were placed in the observer's room, an adjacent 2.5 × 2.1 m soundproof chamber (Industrial Acoustics Corporation). The lamps were mounted in a 26 × 26 × 20 cm box with a 10 × 8 cm opening at the front. A piece of white Plexiglas covered the opening, creating an area of uniform luminance which subtended a visual angle of 2.8 deg at the position of the observer, 2 m away.

The loudspeaker was mounted behind a 10.2-cm-diam hole in a 15 × 15 × 15 cm enclosure as described by Gardner (1968). The lightbox was placed on top of the speaker, which was mounted on a shelf projecting from a corner of the observer's room. Adjustable shelf brackets allowed the bimodal stimulus array to be adjusted vertically to approximate ear-eye level for each observer. A 28 × 34 cm piece of black plastic foam speaker grill (2.5 cm thick) covered the front surface of the speaker and lightbox, giving the illuminated area a grainy, textured appearance.

A small red warning-fixation light was mounted on the wall of the booth directly opposite the observer's chair. With the observer orienting toward and fixating on this light, the stimulus array was located at 40 deg azimuth. Timing of the fixation light and other events in the trial sequence was controlled by a series of internal timers (Hunter, 111C).

Procedure. The sequence of trial events is illustrated in Figure 2. Each trial began with the onset of the warning-fixation light located at 0 deg azimuth. One second later, the 1-sec observation period began. On signal trials, the observation period contained a string of four 125-msec signal bursts. Offset of the warning light signaled the beginning of a 5-sec period during which observers made their verbal responses and, on practice trials, received feedback.

The observer's task was to indicate, by responding verbally, the presence or absence of a signal during the observation period. The observer's responses were transmitted by an intercom and recorded by the experimenter. During the practice sessions and warm-up trials, partial feedback ("correct" or "incorrect") was provided via the intercom after each trial. For all sessions, feedback, in the form of summary statistics, was available at the end of each block of 100 trials.

Three types of signals were presented: an *auditory* trial (S_a) that

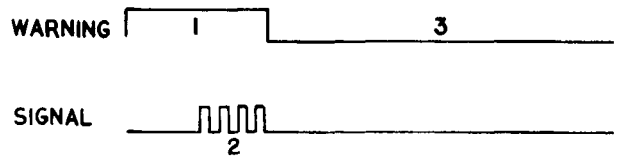


Figure 2. Temporal sequence of trial events: (1) warning/fixation interval, (2) signal interval, (3) response/feedback interval.

consisted of four 125-msec pulses of the 500-Hz sinusoid embedded in a white noise background; a *visual* trial (S_n) that consisted of four 125-msec pulses of a lamp (driven by one channel of the tachistoscope) and a constant background light (provided by a lamp driven by the second channel); and a *bimodal* trial (S_b), in which both the light and the sound signals were presented simultaneously.

The auditory and visual noise sources were on continuously throughout each block of trials. The audio white noise was maintained for all subjects at 58 dBA measured at the position of the observer's head (Brüel & Kjaer, Type 2203 sound-level meter). The background luminance level of the 10-cm² transilluminated surface was approximately 1.2 log fL for all observers (Salford Electrical Instruments exposure photometer). Measured intensity of the signal plus noise for the four observers was 58-59 dBA for the audio stimulus and 1.2-1.3 log fL for the light stimulus. Because our primary concern was with having observers achieve a desired level of performance rather than with determining sensitivity in any absolute sense, precise measurement of stimulus energy levels was not crucial. The chosen stimulus intensities were periodically monitored, however, to assure their reliability.

Each 60-90-min session contained three to five blocks of 100 trials. Except for the first three practice blocks, trials were always presented heteromodally; that is, the three trial types were randomly mixed within the block. The probability of signal occurrence was .6, with S_a , S_v , and S_b trials being equiprobable. Noise-only trials (S_n) occurred with a probability of .4.

Observers were informed of the a priori stimulus probability schedule. Each observer was seated in a chair directly facing the red warning-fixation light and was instructed to orient his or her head toward and fix his or her gaze upon this light while focusing attention on the stimulus array in the periphery. The importance of maintaining head orientation and eye fixation was stressed repeatedly and was visually monitored by the experimenter. On each trial, the observer was required to report whether or not a signal had occurred and whether it was a light, a tone, or both.

After screening for normal vision and hearing, each observer served in four practice sessions (at least 1,200 trials) before the experimental data were collected. During the first three practice sessions, signal energy in each modality was gradually decreased until performance approached the desired level. When necessary, bonus money was allocated according to a payoff matrix in order to shift the observer's criterion.

Data was collected in two subsequent sessions, each consisting of three blocks of 100 trials for each of the three observers (600 total trials), and four blocks of 100 trials for the fourth (800 total trials). Each session began with 50 warm-up trials of gradually increasing signal intensity and included an additional 20 warm-up trials at the beginning of the second and third (and fourth) blocks.

Statistical analysis. The response probability data were analyzed using statistics, described by Shaw (Note 1), which test the predictions of the models described earlier. In the case of the Type 1 mixture model, which predicts additivity of response probabilities under the identity transformation, the appropriate test of the null hypothesis is:

$$z = \frac{\hat{P}_n + \hat{P}_b - \hat{P}_a - \hat{P}_v}{S_{\text{total}}}, \quad (11)$$

where \hat{P}_i is the sample estimate of the true value of P_i and S_{total} is the estimated standard deviation of the numerator of the right side of Equation 11.

For the *sharing model*, which predicts additivity under the logarithmic transformation, the appropriate statistic is:

$$z = \frac{\ln \hat{P}_n + \ln \hat{P}_b - \ln \hat{P}_a - \ln \hat{P}_v}{S_{\text{total}}}, \quad (12)$$

where S_{total} is the standard deviation of the numerator of the right side of Equation 12.

The *weighted integration model* predicts additivity of the response probabilities under the z (inverse Gaussian) transformation. The appropriate statistic in this case is:

$$z = \frac{\hat{z}_n + \hat{z}_b - \hat{z}_a - \hat{z}_v}{\left(\sum_i \frac{\hat{P}_i(1 - \hat{P}_i)}{N_i[\text{ord}(z_i)]^2} \right)^{1/2}}.$$

For a detailed discussion of the derivation of these statistics see Shaw (Note 1).

Results

The data and summary statistics for the four observers are presented in Table 2. A z score whose absolute value exceeds 1.96 indicates rejection of the model in question at $p < .05$.

Data from Observers 2, 3, and 4 firmly support the sharing version of the independent decisions model. Subject 1's data also rejected the null hypothesis for the sharing model ($p < .05$) but did not distinguish between the sharing model and the integration model.

DISCUSSION

The results strongly favor the independent decisions sharing model as a description of how information from simultaneous auditory and visual inputs is combined in a detection task. Data from three of four subjects soundly rejected the other three models considered. The fourth subject's data, although somewhat less incisive, also supported the sharing model.

Table 2
Summary Statistics for Three Models

Sub- ject	Probability of a "No" Response				z Score		
	P_b	P_a	P_v	P_n	Shar- ing	Mix- ture 1	Inte- gration
1	.08	.38	.37	.90	1.90*	-3.29†	1.98**
2	.08	.26	.35	.90	.58	-5.04††	3.14†
3	.06	.30	.22	.91	.56	-7.03††	4.00††
4	.15	.42	.34	.92	.17	-5.04††	4.46††

Note—A significant z score for a given model is evidence against that model. Rejection of all three models tested above would indicate possible support of the Mixture 2 model. For probabilities, b = signal in both, a = auditory signal, v = visual signal, and n = no signal. * $p < .10$. ** $p < .05$. † $p < .01$. †† $p < .001$.

The conclusions of Eijkman and Vendrik (1965) and Fidell (1970), that performance in a bimodal detection task is best described by an integration model, are not supported by our data. When accuracy levels are sufficiently high, as they were in the present study, to allow a comparison of the independent decisions and integration models, the independent decisions model is shown to be a better predictor of performance. The generality of this result is, of course, limited somewhat by the rather specific stimulus conditions of the present experiment. Although we have no evidence suggesting that the results are specific to these particular conditions (in fact, a similar experiment using a single 100-msec stimulus rather than a string of four pulses yielded the same results), replication with stimuli at 0 deg azimuth or at different spatial loci would be a worthwhile endeavor.

The results complement those from Shaw's two visual search experiments related earlier (Shaw, Note 1), suggesting that similar decision processes are used to combine information from the component stimuli in these two situations. Shaw's results indicated that some observers performed according to the predictions of the sharing model and others performed according to those of the Type 1 mixture model. The fact that all observers in the present study performed in accordance with the same model may indicate that attention and decision processes are more uniform with these particular stimuli than in the multiple location task. Data from several additional subjects would be needed to verify this. In any case, the results do demonstrate that the paradigm and theoretical framework presented by Shaw provide a useful means of investigating other situations involving multiple sources of information.

Evidence favoring the independent decisions model of pooling information from multiple sources has emerged from other paradigms as well. One example comes from the spatial frequency view of visual pattern perception (e.g., Blakemore & Campbell, 1969), in which the various spatial frequency channels can be viewed as a set of independent information sources. The manner in which information from component spatial frequencies is combined in a detection task has been examined by Graham (1977) and by Sachs, Nachmias, and Robson (1971). The independent decisions model, referred to as the *probability summation model* in the spatial frequency literature (Graham, 1977), has generally provided a better account of two-interval forced-choice data than has the integration model.

Other evidence for independent decisions as a strategy for combining multiple sources of information comes from a decision-making study employing more complex stimuli (Shaw & Cantor, Note 2). Subjects were asked to decide on the basis of the values of two sources of information—a weather

prediction and a prediction of hotel availability—whether or not to visit a particular city. In a rough parallel to a signal detection paradigm, the authors found that subjects combined information from the two sources according to the independent decisions model. One goal of our subsequent research will be to investigate, by examining other situations in which information from several sources is pooled, the apparent pervasiveness of the independent decisions strategy.

Our discussion of the various decision models presented in this paper has assumed that the probability distributions of the internal random variables are independent. It is also conceivable that the random variables generated by each modality are correlated. The effects of this possibility on predictions made by the weighted integration model and the sharing model have been considered previously (Shaw, Note 1). It can easily be shown that the predictions made by the weighted integration model are unchanged by the degree of correlation between auditory and visual signals, assuming that the signal and noise distributions have equal variances. If these variances are unequal, the predictions of additivity of the transformed response probabilities will not hold.

Correlation between the internal random variables does change the additivity predictions of the independent decisions sharing model. If it is assumed that the joint distribution of the random variables X_{aj} and X_{vj} is bivariate normal, then, with nonzero correlation between the variables, it can be shown that the response probabilities are additive under the identity transformation. The results, in this case, would be the same as those predicted by the Type 1 mixture model.

The effect of correlated signals on the predictions of the mixture models is more intuitive. With the Type 1 mixture model, because one modality is being completely ignored, a correlation between modalities will have no effect on the response probabilities. For the Type 2 mixture model, the effect of nonindependence will be a function of how much information from the “unattended” modality is utilized. The predictions will tend to resemble those of the Type 1 mixture model, however, since, as attention to the secondary modality increases, the Type 2 mixture model increasingly resembles the sharing model, which, with correlated signals, yields predictions identical to those of Type 1 mixture model.

Data from the present experiment resembles neither that predicted by the weighted integration model nor that predicted by the Type 1 mixture model. The possibility that the random variables for the auditory and visual signals are correlated, therefore, is unlikely. Although there is substantial evidence that visual and auditory sensory inputs are combined at several points

in the nervous system, the present data indicate that, in a bimodal detection experiment, these two modalities act as independent sources of information. This result was obtained in the present experiment in spite of the spatial coincidence of the component stimuli.

REFERENCE NOTES

1. Shaw, M. L. *Attending to multiple sources of information*. Manuscript accepted pending revision.
2. Shaw, M. L., & Cantor, N. Personal communication, 1980.

REFERENCES

- BLAKEMORE, C., & CAMPBELL, F. W. On the existence of neurons in the human visual system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology*, 1969, **203**, 237-260.
- BROWN, A. S., & HOPKINS, H. D. Interaction of the auditory and visual sensory modalities. *Journal of the Acoustical Society of America*, 1967, **41**, 1-6.
- BURNS, D. A dual-task analysis of detection accuracy for the case of high target-distractor similarity: Further evidence for independent processing. *Perception & Psychophysics*, 1979, **25**, 185-196.
- CRAIG, A., COLQUHOUN, W. P., & CORCORAN, D. W. J. Combining evidence presented simultaneously to the eye and the ear: A comparison of some predictive models. *Perception & Psychophysics*, 1976, **19**, 473-484.
- EIJKMAN, E., & VENDRIK, A. J. H. Can a sensory system be specified by its internal noise? *Journal of the Acoustical Society of America*, 1965, **37**, 1102-1109.
- FIDELL, S. Sensory function in multimodal signal detection. *Journal of the Acoustical Society of America*, 1970, **47**, 1009-1015.
- GARDNER, M. Lateral localization of 0°- or near 0°-oriented speech signals in anechoic space. *Journal of the Acoustical Society of America*, 1968, **44**, 797-802.
- GARDNER, G. T. Evidence for independent parallel channels in tachistoscopic perception. *Cognitive Psychology*, 1973, **4**, 130-155.
- GORDON, B. The superior colliculus of the brain. *Scientific American*, 1972, **227**, 72-82.
- GRAHAM, N. Visual detection of aperiodic spatial stimuli by probability summation among narrowband channels. *Vision Research*, 1977, **17**, 637-652.
- GREEN, D. M., MCKEY, M. J., & LICKLIDER, J. C. R. Detection of a pulsed sinusoid in noise as a function of frequency. *Journal of the Acoustical Society of America*, 1959, **31**, 1446-1451.
- GREEN, D. M., & SWETS, J. A. *Signal detection theory and psychophysics*. New York: Wiley, 1966.
- KINCHLA, R. A. The role of structural redundancy in the perception of visual targets. *Perception & Psychophysics*, 1977, **22**, 19-30.
- KINCHLA, R. A., & COLLYER, C. E. Detecting a target letter in briefly presented arrays. A confidence rating analysis in terms of a weighted additive effects model. *Perception & Psychophysics*, 1974, **16**, 117-122.
- MORTON, J. Comments on interaction of the auditory and visual sensory modalities. *Journal of the Acoustical Society of America*, 1967, **42**, 1142-1143.
- RUMELHART, D. E. A multicomponent theory of the perception of briefly exposed visual displays. *Journal of Mathematical Psychology*, 1970, **7**, 191-218.
- SACHS, M. B., NACHMIAS, J., & ROBSON, J. G. Spatial frequency channels in human vision. *Journal of the Optical Society of America*, 1971, **61**, 1176-1186.

- SCHNEIDER, G. E. Two visual systems. *Science*, 1969, **163**, 895-902.
- SHAW, M. L., & SHAW, P. Optimal allocation of cognitive resources to spatial locations. *Journal of Experimental Psychology: Human Perception and Performance*, 1977, **3**, 201-211.
- SHIFFRIN, R. M., & GEISLER, W. S. Visual recognition in a theory of information processing. In R. L. Solso (Ed.), *Contemporary issues in cognitive psychology: The Loyola Symposium*. New York: Wiley, 1973.
- WICKELGREN, W. A. Strength theories of disjunctive visual detection. *Perception & Psychophysics*, 1967, **2**, 331-337.
- WEISKRANTZ, L., WARRINGTON, E. K., SANDERS, M. D., & MARSHALL, J. Visual capacity in the hemianopic field following a restricted occipital ablation. *Brain*, 1974, **97**, 709-728.

(Received for publication April 3, 1980;
revision accepted June 20, 1980.)