

# Learning to utilize information presented over two sensory channels\*

JOSEPH HALPERN†

*University of Denver, Denver, Colorado 80210*

and

ALMA E. LANTZ

*Denver Research Institute, Denver, Colorado 80210*

In three separate experiments, Ss were provided with auditory, visual, or simultaneous auditory and visual information in a classification task. Difficulty of classification was manipulated by varying the stimulus exposure duration. Consistent bisensory facilitation effects were noted for later trials, with interference evident on earlier trials. Exposure duration influenced rate and not amount of learning, with bisensory performance being most affected by duration. A transfer paradigm was used in Experiment III, and little if any transfer was noted between unisensory and bisensory stimulus conditions. It was concluded that Ss were extracting the most salient bisensory stimulus components from the auditory and visual modes of information into a unidimensional information configuration.

Studies of intersensory processes are, for the most part, designed to determine whether, and/or how, organisms utilize information that is presented to more than one sensory modality. Typically, a bisensory task is used where stimulus information is provided to the visual and auditory modalities, either simultaneously or separated by some short interval. The design frequently involves three groups, one of which receives combined visual and auditory information, the other two groups experiencing either the auditory or visual stimuli only. An intersensory facilitation effect is noted when the performance of the bisensory group is superior to the performance of each of the other two groups. Intersensory interference refers to circumstances where bisensory performance is inferior to the performance of either of the unisensory groups. While both facilitation and interference have been obtained (as well as the absence of intersensory effects), the most typical result is facilitation, particularly in situations where the stimuli are difficult to detect or discriminate. Excellent reviews of the "intersensory" literature have been provided by Loveless, Brebner, and Hamilton (1970) and Nickerson (1973), with the latter being largely limited to experiments which emphasize a response time (RT) dependent measure.

With few exceptions, data on intersensory processes

come from experimental contexts which have used either a detection or discrimination paradigm along with short-duration stimulus exposures. Although there has not been a considerable effort directed toward theoretical development, two different, but not opposing, theories have been proposed. These are the energy summation hypothesis (e.g., Bernstein, 1970) and the preparation-enhancement model (e.g., Nickerson, 1973). Very briefly, the former attributes bisensory facilitation to the cross-modal summation of stimulus energy, while the latter treats one of the two sources of bisensory information as an accessory which serves primarily to "set," "orient," or prepare the S for the other source of information.

At a strictly intuitive level, it would appear that the ability to effectively utilize two sources of simultaneous or near-simultaneous bisensory information would develop as a result of some practice. To a considerable extent, this "practice" effect has been minimized by the use of detection paradigms and relatively simple, repetitive "signal" stimuli (Dornbush, 1971, and Nazzaro & Nazzaro, 1970, are exceptions). Nonetheless, in reviewing data provided by these detection paradigms, Loveless et al (1970) note a consistent "learning" effect. That is, bisensory facilitation is typically noted only after some initial period of bisensory interference. This would suggest that Ss require some time to learn how to process bimodal information. In the energy-summation and the preparation-enhancement models, which were developed primarily from RT data and short-duration stimulus exposure detection/discrimination paradigms, the role of learning is minimized.

\*This research was supported by the Engineering Psychology Division of the Office of Naval Research. Special thanks are due to the Computer Image Corporation for technical assistance and preparation of stimulus materials.

†Requests for reprints should be sent to Joseph Halpern, Department of Psychology, University of Denver, Denver, Colorado 80210.

In the experiments reported here, Ss were confronted with a classification task with unfamiliar dynamic stimulus information presented either auditorially, visually, or both for 2.5- or 5-sec durations. The purpose of the research was to demonstrate the role of learning in bisensory information processing, and to detail the effects of certain parameters of this learning.

### EXPERIMENT I

A variable that is known to exert a considerable influence on the course of learning in bisensory stimulus situations is task difficulty (Loveless et al, 1970). In the present experiment, task difficulty was operationalized as stimulus exposure duration. A shorter duration exposure (2.5 sec) was considered to result in a more difficult classification task than a 5-sec exposure. Ss were provided with bisensory or unisensory information at either 2.5- or 5-sec exposures.

#### Method

**Subjects.** The Ss were 90 University of Denver undergraduate volunteers, recruited from the Career Placements Office. They were paid \$2 per session and a bonus of 2c for each correct response. The 90 Ss were assigned randomly to one of six groups.

**Stimuli and Apparatus.** Stimuli for the classification task were developed from recordings of ambient sea noises (various whale sounds). From a single 30-min sonar recording of a whale, 72 samples, each 5 sec in length, were selected randomly. An identical selection procedure was used with two other whales, for a total of 216 different stimuli, 72 from each stimulus category. These were randomly combined to form a single audio tape. The tape was then used to provide acoustic inputs to an animation computer. The audio inputs were fed through a set of 16-bandpass filters and envelope detectors, which provided uniform amplitude response characteristics across an audio spectrum ranging from 47 Hz to 19.2 KHz. Each filter corresponded to one of 16 circles presented as a two-dimensional array arranged in a 4 by 4 matrix. Figure 1 is a typical frame from a 16-mm film recording of the visual display. The circles are generated by the animation computer. Each circle receives its animation input from one filter. The signals are ordered such that the lowest frequency channel animates the upper left circle, while the highest frequency channel animates the lower right-hand circle. As the acoustic signal excites the filters, the size of each circle varies in response to the output of its filter section. Changes in the diameter and the velocities with which these variations occur indicate the dynamic energy content of the acoustic signal over the 16 frequency bands. With no input to the filters, the circles appear at their initial diameters. When a signal is applied and the output from a filter begins to increase, the diameter of the corresponding circle visibly decreases, the rate of decrease varying with input intensities. These momentary changes in circle diameter appear to the observer as a dynamic pattern.

The above process was used to convert the auditory information into a visual display which was then recorded on 16-mm film according to the following specifications. Each trial began with the onset of a trial number (1 sec in duration) followed by a 3-sec blank interval, a 2.5- or 5-sec stimulus, a 5-sec blank interval, the next trial number, etc. The audio tape followed the same temporal pattern. Visual information was provided by a 16-mm film projector, and auditory information originated from a separate stereo tape deck and a pair of matched speakers.

**Design and Procedure.** The Ss were divided into six groups of 15 Ss each. Each group received visual information, with a stimulus duration of either 2.5 or 5 sec (Groups V<sub>2.5</sub> and V<sub>5</sub>, respectively),

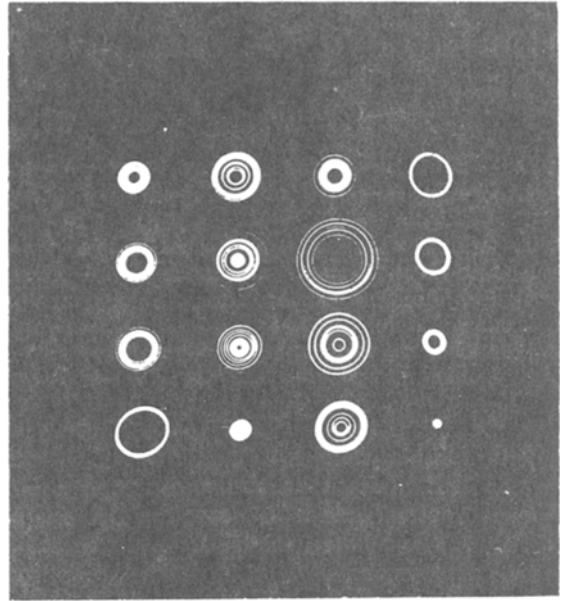


Fig. 1. Photograph of a typical frame of the visual display.

auditory information (Groups A<sub>2.5</sub> and A<sub>5</sub>), or combined visual and auditory information (Groups V+A<sub>2.5</sub> and V+A<sub>5</sub>). In bimodal conditions, audio and visual stimulus onsets and offsets were concurrent.

The Ss were run individually in a darkened room, and were instructed to make a judgment after each stimulus by depressing one of three telegraph keys labeled "I," "II," or "III." They were given a single practice exposure to a sample from each stimulus category, and no other information concerning any relationships between visual and auditory stimuli. Immediate feedback was given after each response. All groups received a total of 640 trials, 320 trials on each of 2 consecutive days. Day 2 trials were preceded by 25 "practice" exposures which required no response. The 640 trials were randomly constructed from the set of 216 stimuli such that no single stimulus appeared more than four times throughout the 2 days.

#### Results and Discussion

Figures 2 and 3 give the mean proportion of correct responses, in blocks of 40 trials, for the groups receiving 2.5- and 5-sec exposures, respectively. The figures show a rather marked bisensory facilitation effect at both levels of exposure duration for the later trial blocks. As expected, facilitation was dependent upon experience, since considerable interference was apparent during the earlier blocks of trials. In general, the figures show that while the performance of all groups improved with trials, the rate and amount of improvement was greater for groups experiencing bisensory information. These observations are substantiated by the results of analyses of variance. A separate analysis was done for the data in each figure. The analyses involved two independent variables, groups as between-S and trials as within-S variables, with mean proportion of correct responses per trial block defined as the dependent measure. The analyses yielded a significant main effect of trials for both figures,  $F(15,630) = 1.94, p < .01$  (2.5-sec

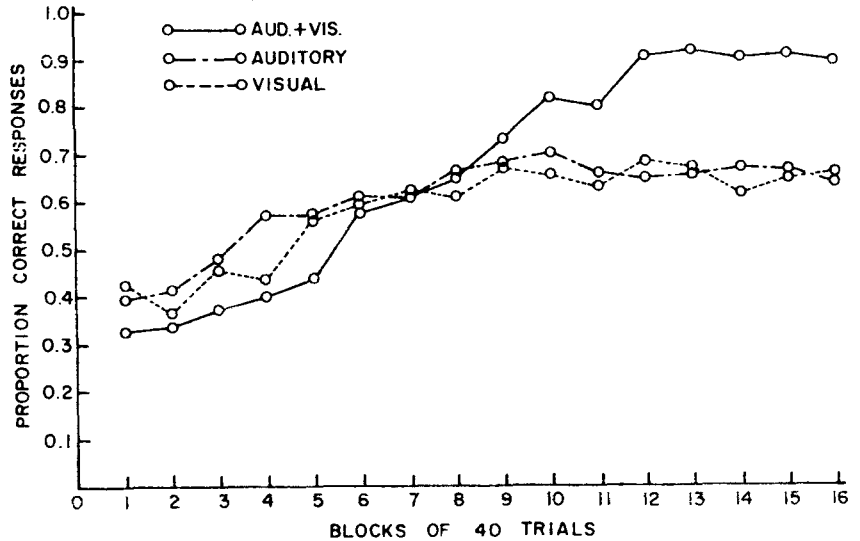


Fig. 2. Mean proportion of correct responses for each group in blocks of 40 trials (2.5-sec exposure, Experiment I).

exposure), and  $F(15,630) = 1.54, p < .05$  (5-sec exposure), as well as significant Groups by Trials interactions,  $F(30,630) = 2.47, p < .01$ , and  $F(30,630) = 1.97, p < .05$ , for Figs. 2 and 3, respectively.

Comparisons between Figs. 2 and 3 show that differences in performance as a function of exposure duration were not great.<sup>1</sup> In fact, the effect of "task difficulty" is apparent only on the earlier trials. Specifically, until about the fifth or sixth trial block, 2.5-sec exposures resulted in poorer performance than 5-sec exposures. These differences were not, however, significant, and had disappeared by the seventh trial block, which, interestingly, is approximately the same point at which bisensory and unisensory performance differences disappear. Finally, Group A+V<sub>5</sub> shows a facilitation effect somewhat earlier than does Group A+V<sub>2.5</sub>. In general, then, it appears as though exposure duration influences rate of learning and not amount, as there is little or no difference in

performance with task difficulty by the ninth trial block. It also appears as though shorter durations lengthen the time needed to learn to utilize bisensory information.

### EXPERIMENT II

As stated previously, much of the available research concerned with intersensory phenomena involves an RT dependent measure (e.g., Nickerson, 1973). Bisensory facilitation, as defined by RT, is not uncommon. Again, the vast majority of these experiments have employed detection/discrimination tasks with repetitive "signal" stimuli. Further, the two modes of information are frequently not contiguous in time, and one of the two is typically relegated to the role of a noninformative accessory. In the present experiments, there are no time lags between visual and auditory presentations, the stimuli are not repetitive, and the task requires classification.

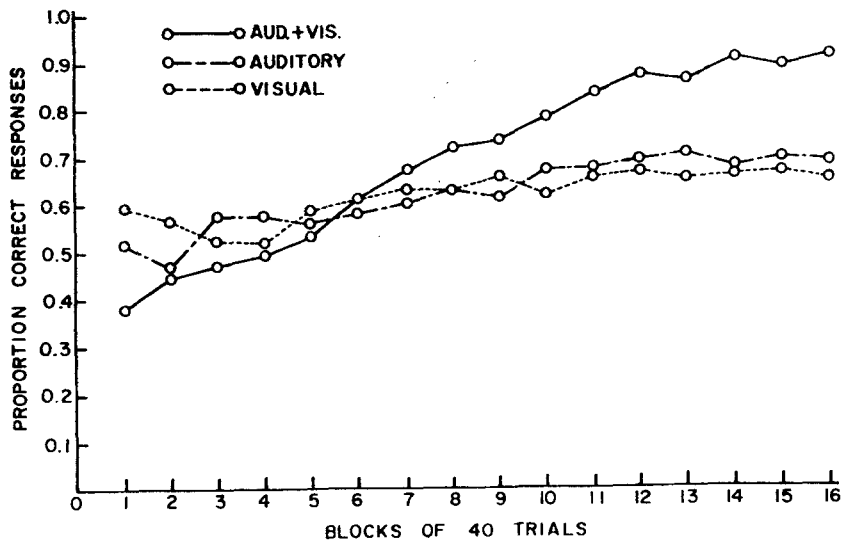


Fig. 3. Mean proportion of correct responses for each group in blocks of 40 trials (5-sec exposure, Experiment I).

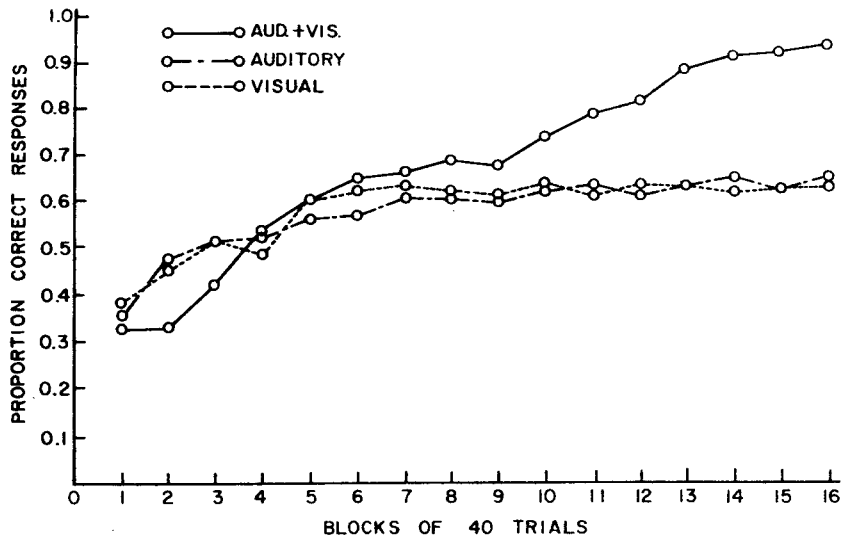


Fig. 4. Mean proportion of correct responses for each group in blocks of 40 trials (2.5-sec exposure, Experiment II).

Further, there is no evidence, i.e., from Experiment I, to suggest that either source serves an accessory function.

In spite of these differences, it is of some interest to determine whether or not similar intersensory effects would be obtained with an RT dependent measure as with correct choices. It should be noted that RT here is probably best considered a response time rather than a reaction time. Experiment II was an exact replication of Experiment I (90 Ss, six groups, etc.), except that a speed/accuracy incentive function was introduced. That is, Ss were informed that they could respond at any time after the termination of the stimulus but that more rapid correct responses would yield greater payoff. A correct response between 1 and 2 sec yielded an 8c bonus; between 2 and 3 sec, a 6c bonus, etc. There were no penalties for incorrect responses. It is of some importance to note that Ss in Experiment I did not typically respond immediately

at stimulus offset, preferring instead to wait some period.

### Results and Discussion

**Proportion Correct Responses.** Figures 4 and 5 give the mean proportion of correct responses, in blocks of 40 trials, for the 2.5- and 5-sec exposures, respectively. Inspection of the figures indicates that the speed/accuracy incentives had some effect on performance but that this was not very great. That is, Figs. 4 and 5 show essentially the same pattern of performance as do Figs. 2 and 3. Analyses of variance, identical to those done in Experiment I, yielded similar, but not identical, results. For the data in Fig. 4, there was a significant main effect of trials,  $F(15,630) = 1.87$ ,  $p < .01$ , and a significant Groups by Trials interaction,  $F(30,630) = 2.07$ ,  $p < .01$ . In Fig. 5, which gives the 5-sec exposure data, the analysis yielded a significant Groups by Trials

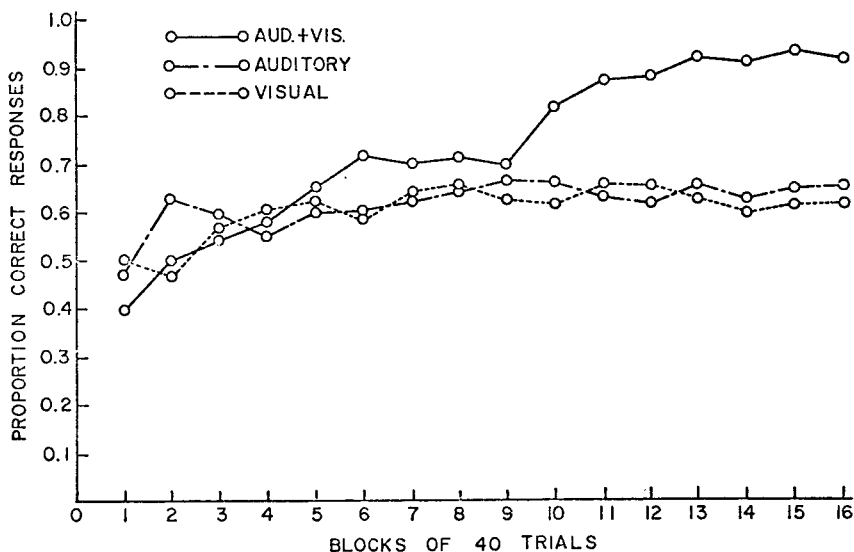


Fig. 5. Mean proportion of correct responses for each group in blocks of 40 trials (5-sec exposure, Experiment II).

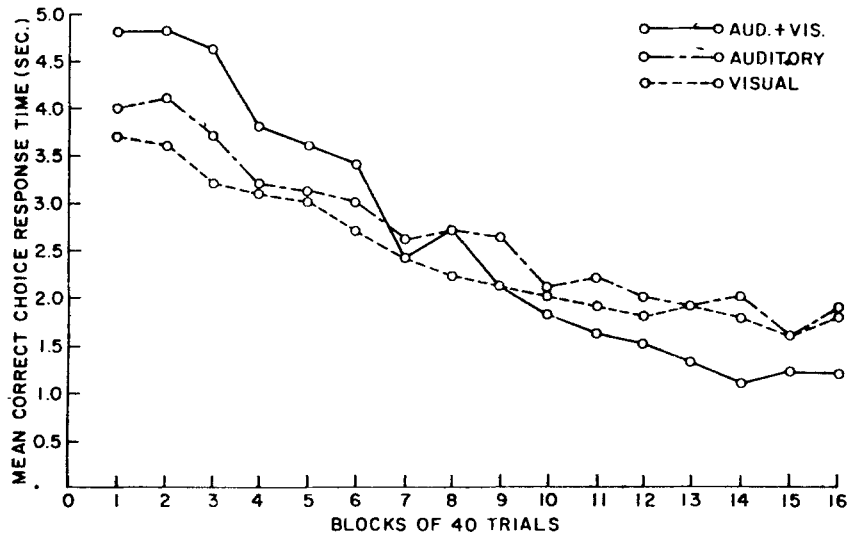


Fig. 6. Mean choice response time for correct responses in blocks of 40 trials for each group (2.5-sec exposure, Experiment II).

interaction,  $F(30,630) = 2.77, p < .01$ , and a significant main effect of groups,  $F(2,42) = 2.79, p < .05$ . Differences between trials were not significant,  $F(30,630) = 1.41, p > .05$ .

The figures show that emphasis on speed influenced 2.5- and 5-sec exposures somewhat differently. As in the previous experiment, there was a "cross-over," with earlier trials showing interference and later trials showing facilitation. The most apparent difference between the two experiments is the fact that bisensory facilitation occurs earlier with speed/accuracy incentives. This difference is more dramatic for the 5-sec exposures where a main effect of groups was obtained. In general, altering the task to a speeded classification situation improved the rate of learning for both exposure durations, with improvement being greater, but not significantly so, for the bisensory groups, and greatest (again not significant) for Group A+V5. Finally, speed instructions also

improved accuracy during the early stages of learning, especially for the bisensory groups, and particularly when exposures were relatively long.

**Choice Response Time.** The mean correct RTs, in blocks of 40 trials, for the 2.5- and 5-sec exposures are given in Figs. 6 and 7, respectively. Consider Fig. 7 first. Here, RT declines rapidly for all groups and stabilizes at less than 1 sec. It seems likely that decisions were being made during the stimulus presentation interval. Interestingly, Group A+V5 had a longer RT than the equivalent unisensory groups until about the fifth trial block, when the differences disappeared. The fifth trial block is when Group A+V5 surpassed the unisensory groups in proportion of correct responses. It is also noteworthy that RT decreased as the proportion of correct responses increased. There is, in effect then, "no speed/accuracy tradeoff," as Ss are achieving both speed and accuracy simultaneously. This conclusion

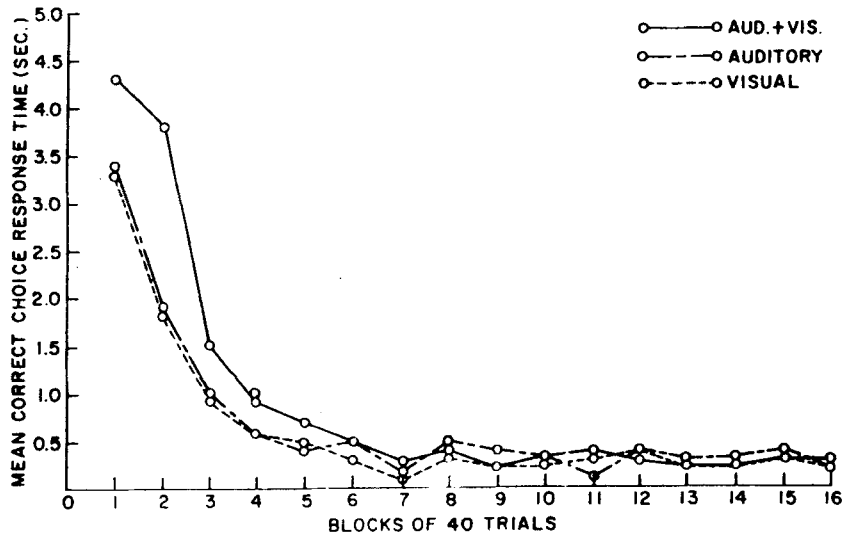


Fig. 7. Mean choice response time for correct responses in blocks of 40 trials for each group (5-sec exposure, Experiment II).

is, however, equivocal, since the RT measure is apparently confounded by the fact that a 5-sec exposure is too long.

The data for the shorter duration exposures are somewhat different, and more informative. First, across all trials and groups, RTs are longer here. Although RT decreases over time, the decrease is more gradual and performance is never quite as good as with the 5-sec exposures. Clearly, by the end of training, 5-sec-exposure RT is shorter than 2.5-sec RT. The combined effect of exposure duration on terminal RT performance was evaluated by computing mean RT over the last four trial blocks for all groups under each exposure. Two composite means, one for each exposure duration, were compared via a Cochran and Cox adjusted *t* (variance for the 5-sec exposures was low, while that for the 2.5-sec exposures was very high). The means were significantly different,  $t_{adj}(88) = 1.79, p < .05$ . Thus, at a point where choice performance for 2.5- and 5-sec exposures was roughly equal, RT was significantly lower for the latter groups. Finally, both figures show a pattern of early interference and later facilitation, although this pattern is less dramatic with RT than with choice performance.

**EXPERIMENT III**

In any theoretical description of bisensory information processing, it is essential to detail how the two different sources of stimulation are utilized. In an energy summation model, the visual and auditory information are assumed to combine according to some process function, e.g., additivity, integration, etc. In an enhancement-preparatory model, the two modes serve different, but related, functions. A transfer situation provides one way of establishing how multimodal information is utilized by the S in a task such as the one here. Data relating to multimodal transfer are limited to crossmodal tasks where, for example, Ss are shifted from visual to auditory stimulation, e.g., Holmgren, Arnouldt, and Manning

(1966) and Rubinstein and Gruenberg (1971). In the present experiment, Ss were exposed to either unisensory or bisensory conditions during a pretransfer phase (Phase I), and then shifted to the other condition for a posttransfer phase (Phase II).

**Method**

**Subjects.** The Ss were 120 University of Denver undergraduate volunteers recruited from the Career Placements Office; they were paid \$2 per session plus bonuses for correct responses.

**Apparatus and Stimuli.** The apparatus and stimuli were identical to those used in the previous experiments.

**Design and Procedure.** The design and procedures were the same as those used in Experiment I, except for the following. There were eight groups of Ss. Each group received either unisensory or bisensory information during Phase I. Phase I training occurred over a 2-day period, the first day consisting of 320 trials as in the two previous experiments. On the second day, Phase I training was continued for an additional 120 trials. The Ss were then immediately shifted to the other training condition for an additional 320 trials with no added instructions, e.g., if preshift training consisted of the unisensory conditions, then postshift was bisensory. The Ss in Group A→A+V<sub>5</sub> were given 5-sec exposures and the auditory Phase I condition, and then shifted to the bisensory condition, while those in Group V→A+V<sub>5</sub> were transferred from visual preshift to bisensory. Groups A+V→A<sub>5</sub> and A+V→V<sub>5</sub> received Phase I bisensory stimulation and transfer to auditory and visual Phase II conditions, respectively. The design was replicated for 2.5-sec exposures, yielding a total of eight groups.

**Results and Discussion**

Figures 8 and 9 give the mean proportion of correct responses, in blocks of 40 trials, for groups receiving the 2.5- and 5-sec exposures, respectively. Phase I performance, for all groups, was comparable to the performance over the first 11 blocks of trials of the like-treated groups of Experiments I and II. In both figures, a sudden drop in performance, immediately after the shift in stimulus conditions, is apparent, with mean proportion of correct responses deteriorating to near Trial Block 1 performance levels. This sudden deterioration in performance could have been due to several factors in addition to changes in stimuli, not the least of which was the fact that stimulus conditions were altered with no warning. Subsequent Phase II trials, however, show rapid

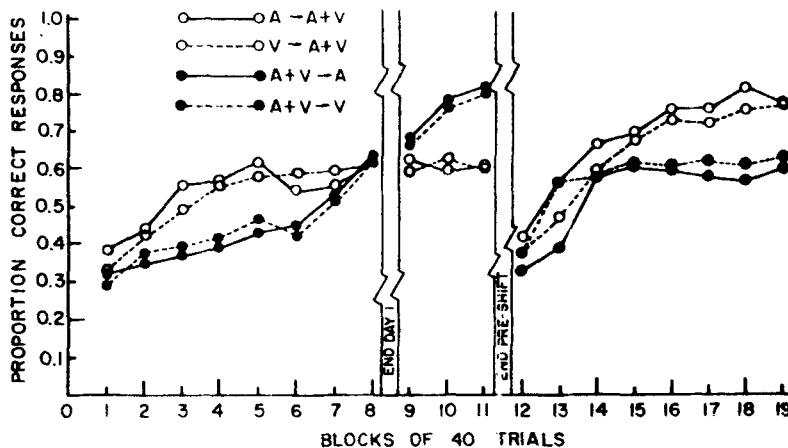
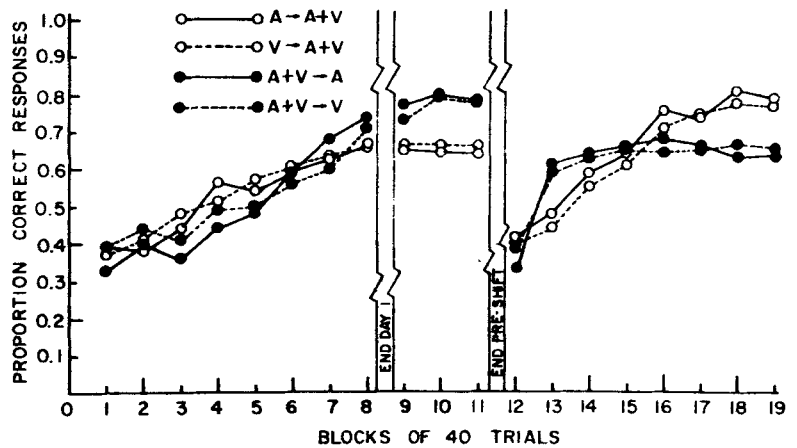


Fig. 8. Mean proportion of Phase I and Phase II correct responses for each group in blocks of 40 trials (2.5-sec exposure, Experiment III).

Fig. 9. Mean proportion of Phase I and Phase II correct responses for each group in blocks of 40 trials (5-sec exposure, Experiment III).



increases in performance for all groups with some differences in performance as a function of exposure duration.

There are two different, but related, considerations involved in the transfer phases, rate of learning and terminal performance levels. Consider the rate characteristic first, i.e., a comparison between Trial Blocks 1-8 and 12-19. In Fig. 8, there was some positive effect of Phase I bisensory training on Phase II unisensory performance, but the effect was slight and probably due to familiarity and warm-up effects. A somewhat more marked effect of Phase I unisensory training on Phase II bisensory performance is evident in the figure. In general, there seems to be some positive transfer from unisensory to bisensory conditions and little or no transfer from bisensory to unisensory stimulation.

A somewhat different pattern of results is evident for the longer exposures in Fig. 9. Here, Phase I bisensory information resulted in an almost instantaneous rise to unisensory asymptotic performance levels, i.e., by the second Phase II trial block. Partially as a result of this, bisensory facilitation does not appear until the fifth trial block. There is, however, some transfer, since facilitation did not occur until the eighth trial block of Phase I.

Comparisons between terminal Phase I and Phase II performance show clearly that prior experiences resulted in only minimal benefit. That is, considering the last two Phase I and Phase II trial blocks, it is apparent, for both exposure durations, that unisensory Phase I performance is approximately the same as unisensory Phase II performance, and that bisensory Phase I and Phase II performances are virtually identical. Thus, when Ss experience bisensory stimulation during Phase I, and are transferred to unisensory stimulation, their performance deteriorates to levels typical of those shown by Ss with equal amounts of only unisensory information. When Ss experience unisensory stimulation during Phase I, and are transferred to a bisensory Phase II condition, performance after 8 trial blocks reaches a level roughly equal to the

performance of Ss who had experienced 11 blocks of trials of bisensory stimuli.

While it would be possible to speculate on the reasons for the relatively small differences found in rates of learning and terminal performance as a function of the direction of shift (i.e., unisensory to bisensory or bisensory to unisensory) and exposure duration, such speculation is hardly warranted by the data. Specifically, the small amounts of positive transfer obtained might just as likely have been due to general procedural familiarity as to any characteristic of the stimulus complex. It seems most reasonable to conclude, from these data, that Ss are processing unisensory and bisensory information according to different rules and strategies. It appears as though Ss are combining the two sources of information that occur simultaneously. It may well be that, in the bisensory situation, Ss are developing what Garner (1972) has termed a composite stimulus where two "relevant" stimuli combine to form a new and different composite stimulus. This possibility is explored at length later.

### GENERAL DISCUSSION

The results of the experiments presented previously are entirely consistent in showing bisensory facilitation which develops after some initial period of bisensory interference. This process is readily apparent with a dependent measure based on correct choices, and somewhat less dramatic with an RT measure. Task difficulty, as defined by two different exposure durations, was shown to influence rate of learning, with the effect being more considerable for bisensory presentations. In general, although resulting in eventual superior performance, bisensory stimuli required additional processing time measured either in terms of number of trials or time to respond.

Prior to any discussion of the possible characteristics involved in bisensory processing, it might prove helpful to emphasize certain performance characteristics that are indicated by the data. First, performance with audio and visual unimodal

information was approximately equal. While this is not a routine finding (see, for example, Dornbush, 1971, and Rollins, 1972), unimodal performance differences seem to depend, to a considerable extent, on procedural and stimulus characteristics. In view of the apparent equivalence between audio and visual sources of information here, it seems reasonable to conclude that both supply relevant information. Second, subsequent to some initial interference, the duration depending upon certain experimental variables, a clear bisensory facilitation effect is apparent, the variables seemingly having little effect on whether or not facilitation will occur, but rather on when it will occur. Finally, there appears to be little or no transfer across processing modes, i.e., bisensory performance seems to profit little from previous unisensory experience, and unisensory processing seems to show little benefit of prior bisensory experience.

In view of the consistency of the above findings, one can begin to propose a general description of bisensory processing characteristics. This can perhaps be best accomplished by eliminating certain possibilities. First, in spite of the fact that the processing of bisensory information requires greater effort and results in initial relatively poor performance, Ss are not choosing to ignore one of the two sources of information. That this is the case is indicated by the consistent bisensory facilitation effects. Second, there is no indication that either of the two sources which comprise the bisensory stimulus is operating as an accessory. That is, if Ss were attending to one source of stimulation in preference to another, there would have been considerably more positive transfer than was obtained. Similarly, it seems unlikely that Ss, in the bisensory conditions, were utilizing the two sources of information as separate, since such an approach would probably also have resulted in greater transfer. In contrast, the pattern of results suggest that bisensory information is being treated as a new stimulus configuration somewhat different from its components. Whether this configuration is arrived at by a process of energy summation, information summation, or information integration, etc., or whether the process is additive or nonadditive is entirely unclear at this point. What is most likely, however, is that a combined auditory and visual stimulus, unlike either of the components, is being developed in the bisensory stimulus situation.

The clear implication is that a composite stimulus is being developed from the bisensory components.

Related notions have been expressed by Garner (1974) and Lockhead (1970), but under very different circumstances, i.e., with respect to integrality of dimensions and redundancy gains. While they disagree on several major issues, e.g., the necessity of dimensional integrality for redundancy gains, both address the issue of how composites of information are formed from separate informational dimensions. In view of the fact that Garner and Lockhead treat situations where redundancy is the critical variable, and we have no way of measuring the redundancy characteristics of the visual and auditory stimuli, little can be said to directly relate these results with their proposals. There does, however, appear to be some correspondence, albeit indirect, between the data reported here and the findings reported by Garner (1974).

#### REFERENCES

- BERNSTEIN, I. H. Can we see and hear at the same time? Intersensory facilitation of reaction time. In A. F. Sander (Ed.), *Symposium on attention and performance, III*. Amsterdam: North Holland, 1970.
- DORNBUSH, R. L. Delay of auditory input in "simultaneous" auditory and visual short-term memory. *Perception & Psychophysics*, 1971, **9**, 97-98.
- GARNER, W. R. *The processing of information and structure*. Potomac, Md: L.E.A., 1974.
- HOLMGREN, G. L., ARNOULT, M. D., & MANNING, W. H. Intermodal transfer in a paired-associates learning task. *Journal of Experimental Psychology*, 1966, **71**, 254-259.
- LOCKHEAD, G. R. Identification and the form of multidimensional discrimination space. *Journal of Experimental Psychology*, 1970, **85**, 1-10.
- LOVELESS, N. E., BREBNER, J., & HAMILTON, P. Bisensory presentation of information. *Psychological Bulletin*, 1970, **73**, 161-199.
- NAZZARO, J. R., & NAZZARO, J. N. Auditory versus visual learning of temporal patterns. *Journal of Experimental Psychology*, 1970, **84**, 477-478.
- NICKERSON, R. S. Intersensory facilitation of reaction time: Energy summation or preparation enhancement? *Psychological Review*, 1973, **80**, 489-509.
- ROLLINS, H. A. Serial position effects in simultaneous bisensory memory. *Journal of Experimental Psychology*, 1972, **94**, 162-167.
- RUBINSTEIN, L., & GRUENBERG, E. M. Intramodal and cross-modal sensory transfer of visual and auditory temporal patterns. *Perception & Psychophysics*, 1971, **9**, 385-390.

#### NOTE

1. Additional statistical analyses showed that the difference in performance as a function of exposure duration were not significant in this or the following experiment.

(Received for publication February 25, 1974;  
revision received May 28, 1974.)