INTRODUCTION

A variety of reports have demonstrated ways in which the human visual system dominates the spatial perception of both somatosensory and auditory stimuli when disparate bimodal stimuli are presented (see Welch 1999; Welch and Warren 1980). These interactions have been described as “visual capture” when the perception of one’s hand position is displaced toward the visually perceived location when the visual alignment is altered (Hay et al. 1965) and the “ventriloquism effect” when auditory stimuli are perceived to originate from the location of simultaneously presented visual stimuli (Howard and Templeton 1966). The influence of the visual system on the spatial location of auditory or somatosensory inputs is likely due to the high spatial acuity of the visual system compared with the other two sensory modalities. If it is in fact the case that the sensory system with the highest acuity dominates the percept of discordant bimodal stimuli, then under the right experimental conditions auditory or somatosensory stimuli should dominate the perception of visual stimuli.

There have been a few reports describing conditions where auditory stimuli influence the perception of visual stimuli. Recent examples include the finding that two visual stimuli moving toward each other are perceived to “bounce” if an auditory stimulus is presented simultaneous to the contact of the two visual objects (Sekuler et al. 1997; Watanabe and Shimono 2001). In addition, the presence of multiple auditory stimuli can increase the number of perceived visual stimuli (Shams et al. 2002). Despite these examples of auditory-visual interactions, there have been no studies that describe in detail the conditions under which auditory stimuli can influence clearly discriminable visual stimuli. For example, the extent of the interaction may only occur under conditions where the visual stimuli are difficult to discriminate, yet no interactions may occur under less ambiguous visual stimulus conditions. In this report psychometric functions of the perception of temporal rate were defined for auditory, visual, and combined auditory and visual stimuli. A variety of stimulus conditions were presented to determine the essential components in both space and time that are critical in generating bimodal interactions.

The spatial illusions generated by disparate visual and auditory or somatosensory stimuli can persist after relatively brief periods of exposure to spatially disparate bimodal stimuli (e.g., Canon 1970; Radeau and Bertelson 1974; Recanzone 1998; see also Welch 1999). However, there have been no demonstrations of a strong cross-modal aftereffect that is generated by a non-visual stimulus. One reasonable explanation of how these aftereffects occur is that multi-modal neurons in the brain shift their spatial preferences or alignments during the period of exposure to discrepant stimuli, and this shift endures in the absence of any corrective feedback. If this is a general neuronal mechanism, then the same types of aftereffects should be generated in the visual system following a period of exposure to discrepant visual and auditory stimuli under conditions that the auditory stimuli “capture” the percept of the visual stimuli. The second set of experiments of this report was designed to determine if such auditory aftereffects on visual perception could be induced.

METHODS

All procedures were in compliance with the policies and procedures contained in the Federal Policy for the Protection of Human Subjects and in the Declaration of Helsinki and were approved by the U.C. Davis Human Subjects Committee. All subjects performed these experiments with informed consent.

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
Stimuli and apparatus

Experiments were conducted in a double-walled sound booth (IAC, New York) measuring 2.4 \( \times \) 3.1 \( \times \) 2.0 m \((l \times w \times h)\) lined with echo-attenuating foam. Subjects sat in a chair facing a 9-cm dual-cone speaker placed 1.5 m from the center of the interaural axis. A red LED \((0.3^\circ \text{ visual angle})\) was placed 4\(^{\circ}\) above the speaker to serve as the fixation point. An additional red LED placed directly in front of the center of the speaker was used to present the visual stimuli. Audustic stimuli consisted of 1 kHz tones or broadband (Gaussian) noise of 111- to 142.8-ms duration (5-ms rise/fall) presented at 65 or 40 dB SPL (A-weighted). All stimulus generation and data collection were under the control of a PC using Tucker-Davis Technologies (Gainesville, FL) hardware and software.

Paradigm 1

In this experiment subjects performed a two-interval two-alternative forced-choice task designed to study the interactions of temporal rate perception across a range centered at 4.0 Hz. Four females and four males (Subjects A–H; aged 22–46 yr; median = 25 yr) participated in these experiments. All subjects were right-handed. Trials consisted of two sequences of four auditory stimulus pips and/or four flashes of a visual stimulus with an inter-sequence interval of approximately 1 s. The first sequence of four was always presented at 4 Hz \((125 \text{ ms ON, 125 ms OFF})\). The second sequence was presented at the same or a different temporal rate ranging from 3.5 Hz \((142.8 \text{ ms ON-OFF})\) to 4.5 Hz \((111 \text{ ms ON-OFF})\) in 0.2-Hz steps. On all trials where both visual and auditory stimuli were presented the two stimuli were aligned at the center of each sequence to maximize the amount of overlap in time between the two stimulus modalities and to minimize onset and offset latency cues. Subjects were instructed before a block of trials that they were to attend to only one stimulus modality and to ignore the other. They were further informed that the unattended stimulus (the distracter) would be the same frequency as the attended stimulus (the target) on some trials, but would be a different frequency on others, and therefore was unreliable. On each experimental day, subjects were first given a block of 18 practice trials with the largest temporal rate differences of the attended stimulus to familiarize the subject to the task and ensure that they could perform the easiest trials correctly. A block of 12 randomly interleaved trials for each trial type was then presented. Immediately following this block, the subject was instructed to now attend to the other stimulus modality and then performed a block of 18 practice trials immediately followed by a block of 12 randomly interleaved trials for each trial type. All four blocks of trials were completed within 50 min, and the order of which stimulus modality was attended was counterbalanced between subjects and between experimental days for each subject.

Each session consisted of four possible trial types. The first trial type (Fig. 1A) used only a single stimulus type (the target), which was the modality that the subject was instructed to attend. The second type of trials (Fig. 1B) presented both the target and the distracter stimuli in both sequences and at the same temporal rate. The third type (Fig. 1C) had the target stimulus presented at 4.0 Hz during both sequences, but the distracter was presented at either a higher or a lower temporal rate during the second sequence. The final trial type (Fig. 1D) was similar, except that in this case the distracter was presented at 4.0 Hz during both sequences and the target stimulus was presented at either a higher or a lower temporal rate during the second sequence.

Subjects initiated each trial by fixating a red LED and placing a three-position switch into the center position. After a delay of 1–2 s, the two stimulus sequences were presented. The subject was instructed to judge whether the attended stimulus during the second sequence was presented at a higher or lower temporal rate than the first sequence and to indicate this judgment by moving the switch to one of two possible positions. They were further instructed to always maintain fixation on the fixation LED throughout each trial.

A set of subjects was tested on this task on only one session, but were only required to attend to the visual stimulus. These subjects (Subjects I–N; 3 females aged 26, 29, and 31 yr and 3 males aged 20, 23, and 39 yr) were instructed to always attend to the visual stimulus and to always ignore the auditory stimulus when it was presented. They were also told that the auditory stimulus was unreliable. Subjects first performed a block of 18 practice trials with feedback: correct responses were signaled by an audible click from above and to the right; incorrect responses were signaled by no click. The next block of trials was the same as those described above for the attend-visual task, except that feedback was provided. On target-constant trials (Fig. 1C) the subjects were rewarded randomly (50% probability) regardless of their response. Subjects then performed a third block of 12 randomly interleaved trials for each stimulus type but no feedback was provided during this block.

Paradigm 2

This task used only a single sequence in which both auditory and visual stimuli were presented on all trials. Subjects A–H were instructed to fixate a red LED and move the switch to the center position to initiate a trial. Four auditory pips and four light flashes were presented from a location 4\(^{\circ}\) below the fixation point. The subjects’ task was to indicate if the two sequences were aligned temporally.
(“same”) or if they were not aligned with each other at any time during the period of the entire sequence (“different”). Trials with correct same responses had the two stimuli presented at the same temporal rates of 3.0, 4.0, and 5.4 Hz. The other trial types (different) consisted of visual stimuli presented at 4.0 Hz and auditory stimuli presented at 3.0–5.4 Hz in steps of 0.2 Hz. Each session consisted of a block of 20 practice trials followed by a block of 12 trials of each trial type presented in randomly interleaved fashion (180 trials).

Paradigm 3

The third task was used during the aftereffect experiments. The subjects first performed the same/different task of paradigm 2. This session was immediately followed by a “training” paradigm. Subjects were instructed to fixate the same fixation LED and move the switch to the center position to initiate a trial. A series of four tone pip and light flash sequences were then presented with an inter-sequence interval of approximately 1 s. One to seven sequences were presented before the visual stimulus was presented at a lower luminance. Subjects were instructed to maintain fixation and to move the switch in either direction when they detected the flashing light to be dimmer. Subjects had to make the response before the fourth flash of the sequence was completed. Feedback was provided by a click stimulus on correct trials. This response contingency ensured that the subjects remained alert and attended the visual stimulus throughout the session. Eleven different stimulus sequences were presented 60 times on randomly interleaved trials (approximately 20 min). Four types of training sessions were used. In the first type (simultaneous) the auditory and visual stimuli were presented simultaneously across a temporal rate range of 3.0 to 5.0 Hz. In the second type (auditory faster) auditory stimuli were presented 0.6 Hz faster than the visual stimuli (auditory: 3.6–5.6 Hz, visual 3.0–5.0 Hz). In the third type (auditory slower) the auditory stimuli were presented 0.4 Hz slower than the visual stimuli (auditory: 3.0–5.0 Hz, visual 3.4–5.4 Hz). In the fourth type (control) simultaneous, auditory faster, and auditory slower stimuli were presented on randomly interleaved trials. In this control case each stimulus type was presented only 20 times in each session. Once the training trials were completed, subjects immediately performed the same/different task again in the absence of feedback.

Data analysis

Psychometric functions were fitted using a maximum likelihood ratio with sigmoid functions (Quick 1974) of the form

\[ Y = A - (A - \min) \times \exp(-fd)\beta \]  

where \( Y \) is the probability that the subject will respond “higher,” \( A \) is the asymptote of the function, \( \min \) is the minimum of the function, \( f \) is the temporal rate of the stimulus, \( \delta \) is the mid-point between the maximum and minimum values, and \( \beta \) is the steepness of the function. The maximum likelihood function provided the values of \( A, \min, \delta, \) and \( \beta \). Comparison of all sessions excluding trials where the auditory stimulus was presented at 4.0 Hz during the second sequence revealed that the fit with these functions was significant at the 0.01 level in all but one of over 500 functions tested.

From these functions, the temporal rate corresponding to the 50% “higher” response was taken as the response bias. The term “\( \beta \)” was taken as the slope of each function. The bandwidth of the function was measured as the frequency range between the 25 and 75% level of higher responses. These values were computed using a paired, two-tailed \( t \)-test to determine if there were consistent differences between different trial types across subjects.

To evaluate stimulus interactions, functions were fit individually for each trial type and compared with the single function fit using the data from both trial types being compared. The difference in log likelihood from these two fits approximates a \( \chi^2 \) distribution, and instances where the difference between the fits from each set of trials and the combined trials corresponded to a \( P \) value of \(<0.05\) were interpreted to indicate a significant difference between the two conditions (Hoel et al. 1971).

RESULTS

The first objective was to determine how well subjects could discriminate temporal rate changes in each modality. Figure 2 shows the results from two representative subjects, as well as the mean responses across all eight subjects. These plots were taken from the single stimulus trials from each block of 12 interleaved trials, and therefore, are based on the two nonpractice blocks that were presented in succession on the same day. Across subjects, the ability to discriminate changes in the 1-kHz tone temporal rate (solid squares) from a 4.0-Hz comparison was better than their ability to discriminate the same changes in visual temporal rate (open squares). Figure 2A shows the results from the subject whose performance was closest to the mean of all subjects. Subjects could generally discriminate the largest differences in temporal rate of the visual stimuli (3.5 and 4.5 Hz), but showed poor performance at the smaller changes in temporal rate (3.9 and 4.1 Hz). In contrast, for auditory stimuli performance was perfect at the two largest changes in temporal rate (3.5, 3.7, 4.3, and 4.5 Hz) and usually well above chance for the two smallest temporal rate changes. A notable exception to this general trend is shown for subject B (Fig. 2B), who showed the worst visual temporal rate discrimination performance of all subjects tested, but nonetheless showed the same auditory temporal rate acuity as the other subjects.

Figure 2C shows the mean (±SE) across all subjects. For auditory stimuli, all subjects showed perfect performance except for the two changes in temporal rate nearest the 4.0-Hz standard that was tested. In contrast, the mean response for the visual task never showed perfect performance at any temporal rate tested. Inspection of these curves indicates that these subjects were much better at perceiving differences in auditory temporal rate than differences in visual temporal rate.

To quantify these impressions, each psychometric function was fitted with a sigmoid function and was tested to determine whether one function or two better described the results. Two functions better described the data than one for seven of eight subjects (\( P < 0.01 \)), indicating that auditory temporal rate discrimination was consistently better than visual temporal rate discrimination. Analysis of response bias, defined as the frequency corresponding to the 50% point on the \( y \)-axis, showed no statistically significant differences across subjects (mean: 3.995 ± 0.01 and 3.975 ± 0.02 Hz for auditory and visual tasks, respectively; paired 2-tailed \( t \)-test; \( P > 0.05 \)). However there was a significant difference in the slope of these functions across subjects (mean: 55.34 ± 5.61 vs. 22.99 ± 4.92 Hz for auditory and visual tasks, respectively; \( P < 0.01 \)). This was further reflected by the bandwidth of the psychometric functions being over three times greater for visual stimuli (0.42 ± 0.21 Hz) than for auditory stimuli (0.13 ± 0.07 Hz), which was a statistically significant difference (paired, 2-tailed \( t \)-test; \( P < 0.01 \)). Thus the ability to discriminate auditory temporal rate is clearly superior to the ability to discriminate the same temporal rates in the visual modality.

The next question addressed was how the presence of an
auditory stimulus influences the perception of visual temporal rate. Subjects were instructed to attend to only the visual stimulus and were informed that, when an auditory stimulus was presented, it could be either at the same rate or at a different rate than the visual stimulus, and therefore, should be ignored. The same temporal rates tested on the single modality trials (Fig. 2) were presented on these trials. If there was no influence of the auditory stimulus on their performance at this visual temporal rate discrimination task, the data on the bimodal trials should be indistinguishable from the visual-alone trials shown in Fig. 2. If the auditory stimuli dominated the performance, these data should be indistinguishable from the auditory-alone trials in Fig. 2. Examples of the results from bimodal trials, where the temporal rates of the auditory and visual stimuli were identical but the subjects were instructed to attend to the visual stimuli and ignore the auditory stimuli, are shown in Fig. 3 for the same two representative subjects shown in Fig. 2, as well as the mean across subjects. In these graphs, the dashed gray line (vis pred) shows the predicted data if the subjects were not influenced by the presence of the auditory stimulus (taken as the visual-alone trials of Fig. 2). The solid gray line (aud pred) shows the predicted data if the subjects based their responses on only the auditory stimulus (taken as the auditory-alone trials of Fig. 2). The data from the bimodal trials are shown as open circles. In each individual subject, the data fell more closely to the auditory prediction than to the visual prediction, indicating a clear influence of the auditory stimulus on their performance on the visual temporal task. In many cases there was a dramatic improvement in performance (e.g., subject B, Fig. 3B). Pooling the data across subjects showed that the visual temporal rate discrimination performance was improved at all temporal rates tested (Fig. 3C).

Functions for the bimodal trials were not significantly different from the auditory-alone trials in seven of eight subjects. When the bimodal trials were compared with the visual-alone trials, five of eight subjects showed a significant difference between the two functions. However, when the data were pooled across subjects, the bimodal and visual-alone trials were better fitted with two functions instead of one ($P < 0.01$). Response bias, function slopes, and response bandwidth were not significantly different between bimodal and auditory-alone trials (bimodal trials: $3.985 \pm 0.01$, $53.03 \pm 6.14$, and $0.16 \pm 0.08$, respectively; 2-tailed paired $t$-test; all $P > 0.05$). However, both the slope and the bandwidth measures were significantly different on the bimodal trials compared with the visual-alone trials ($P < 0.05$). These results indicate that the presence of the auditory stimulus can “capture” the perception of the temporal rate of the visual stimulus when both stimuli are presented at the same temporal rate.

A more striking example of this interaction is shown by the trials where the subjects were attending to the visual stimulus but the visual stimulus did not change in temporal rate, whereas the auditory (distracter) stimulus did. If there was no influence of the auditory stimulus on visual rate perception, the percentage of trials that the subjects responded higher should be near 50% regardless of the auditory temporal rate presented. If there was a strong influence, then the performance should approximate that of the auditory-alone trials. Figure 4 shows the results from these trial types for the same two subjects and pooled across all eight subjects. The dashed line shows the prediction if the subjects followed the visual stimulus (chance performance) and the solid line shows the prediction if the subjects followed the auditory stimulus (auditory-alone trials of Fig. 2). Data from these trials are shown as open triangles.

**FIG. 2.** Temporal rate discrimination during single-stimulus trials. A and B: results from the single stimulus trials in 2 sessions in 2 representative subjects. The percentage of trials the subjects responded “higher” is plotted as a function of the temporal rate presented during the second sequence. Solid squares show the results from auditory-alone trials and open squares show the results from visual-alone trials. Lines show the fitted functions. C: mean ($\pm$SE) responses pooled across all 8 subjects. In all cases, auditory temporal rate discrimination was better than visual temporal rate discrimination. Points where error bars are not shown were smaller than the symbol size.
In all cases, the subjects responded in a manner that was consistent with the auditory stimulus and not the visual stimulus. Comparisons between the response bias (4.02 ± 0.02), slopes (49.00 ± 5.10), and bandwidth measures (0.17 ± 0.06) were not significantly different between these target-constant trials when the subjects were attending to the visual stimuli and the auditory-alone trials (all \( P > 0.05 \)) but were different from the visual-alone trials (all \( P < 0.05 \)). This was further supported by no significant difference between one and two fitted functions comparing the target (visual) constant trials to the auditory-alone trials in all subjects tested. These results show a striking ability of the auditory stimulus to influence the percept of the temporal rate of visual stimuli.

The final comparison was on trials in which the auditory stimulus was presented at 4.0 Hz during both stimulus sequences while the subjects attended to the visual stimulus (distracter-constant trials). If there was no influence of the auditory stimulus on visual temporal rate perception, the results from these trials should be indistinguishable from the visual-alone trials. However, if there is an influence, the performance should be near chance as there is no difference in the auditory temporal rate. The data from these trials were consistent with the preceding two figures: there was a dramatic influence on the visual temporal rate discrimination performance by the auditory stimulus (Fig. 5). For all subjects, the data on the distracter-constant trials were more similar to that predicted by the auditory stimulus (solid line, chance performance) than for the visual stimulus (dashed line, visual-alone trials of Fig. 2), with most subjects performing at near chance levels for all visual temporal rates presented. It was not possible to fit a statistically significant function to the data from two subjects. Data from the remaining six subjects were pooled to reveal a mean response bias of 4.02 ± 0.05, which was not significantly different from either the visual-alone or the auditory-alone trials. The mean slope was 6.98 ± 6.45, which was significantly different from both auditory-alone and visual-alone trials (\( P < 0.05 \)). In only two cases were the 25 and 75% higher response measures within 3.5 to 4.5 Hz, giving bandwidth measures of 0.71 and 0.55 in those two cases. All other subjects had bandwidths that were unmeasurable. It was possible to fit a sigmoid function to the mean data, which was significantly different from both the visual-alone and the auditory-alone functions (\( P > 0.05 \)). This dramatic effect on visual temporal rate perception with a constant auditory stimulus is particularly surprising given that this effect was reliably seen even for temporal rates where all subjects had very good performance, such as 4.5 Hz (visual alone mean: 93.1%; auditory constant mean: 68.1%). Thus even though the subjects were presented with salient differences in the temporal rate of visual stimuli, in the presence of an unchanging auditory stimulus they were much less likely to perceive this difference.

The preceding analysis showed that auditory stimuli had a pronounced influence on the performance on this visual temporal rate discrimination task across subjects. The next ques-

![FIG. 3. Bimodal trials when subjects attended to the visual stimulus. A and B: results from the same 2 subjects shown in Fig. 2. The solid line shows the results to auditory-alone trials (aud pred) and the dashed line shows the results from visual-alone trials (vis pred), reproduced from Fig. 2. Open circles (bimodal) show the results from the trials where both stimuli were presented at the same temporal rate during the second sequence and the subjects were attending to the visual stimulus. If there were no interactions, the circles should fall on the dashed line. C: mean responses pooled across all 8 subjects, errors bars (SE) are smaller than the symbol size. The inset at the bottom shows a schematized representative trial where the temporal rate increased for both stimuli. In every case, visual temporal rate discrimination was improved by the presence of the auditory stimulus presented at the same temporal rate.](1082/G. H. Recanzone)
tion was whether visual stimuli had a similar effect on auditory temporal rate perception. Analysis of the same classes of trials taken from blocks when the subjects were attending to the auditory stimuli showed that visual stimuli had essentially no influence on auditory temporal rate perception using the parameters employed in this study. Figure 6 summarizes these results by showing the mean responses across all subjects in a manner similar to previous figures. There were no significant differences in the bias, slope, or bandwidth of the fitted functions (all $P > 0.05$) between bimodal trials and auditory-alone trials (Fig. 6A) or on distracter-constant trials where the visual stimulus was presented at 4.0 Hz during both stimulus sequences and the subjects attended to the auditory stimuli (Fig. 6B). There were significant differences in the bias, slope, and bandwidth of the fitted functions for the target-constant trials, where the auditory stimulus was presented at 4.0 Hz during both stimulus periods, as subjects performed near chance (Fig. 6C). This is expected if the visual stimulus had no influence on auditory temporal rate perception.

In summary, these stimulus interactions indicate that normal human subjects are better able to discriminate auditory temporal rates than visual temporal rates. Further, the presence of an auditory distracter can profoundly influence the perception of visual temporal rates. However, the visual stimuli had no measurable influence on the perception of auditory temporal rates.

One question that arises from this strong influence of auditory stimuli on visual temporal rate perception is whether subjects were really ignoring the auditory stimulus. All subjects reported that they were ignoring the auditory stimuli to the best of their ability during the attend-visual trials, but they all also noted that the auditory task was considerably easier than the visual task. It could be argued that this knowledge of how much easier the task seemed in the auditory realm may have led the subjects to nevertheless attend to the auditory stimulus during the visual task. To test this possibility, the same experiment was performed on six additional naïve subjects that had not participated in any auditory temporal rate experiments. Additionally, these subjects were given feedback during the training block, as well as during the next block of 12 trials/stimulus. It was reasoned that this feedback would instruct the subjects that the auditory stimulus was unreliable. On visual-constant trials, subjects were randomly rewarded (50% probability) regardless of their response. Thus on the bimodal trials the auditory stimulus gave correct information, but on the distracter-constant trials the auditory stimulus provided incorrect (no) information. On the target-constant trials, following the auditory stimulus would provide rewards (“correct”) on half the trials and no reward on the other half for each auditory temporal rate presented. It was reasoned that this strategy should provide sufficient feedback of the unreliability of the auditory stimulus and help the subjects focus on only the visual stimulus.

The results from these subjects during the session without feedback are shown in Fig. 7. These naïve subjects showed the same influence of the auditory stimulus on their visual temporal rate performance as was seen in the practiced subjects. For the target-constant trials, even though the subjects were instructed that the auditory stimulus was unreliable and were...
given feedback, they still performed in the same direction as the auditory stimulus (Fig. 7A). Similarly, the auditory-constant trials resulted in a decreased performance on the visual task (Fig. 7B). These results are essentially the same as for the other eight subjects shown in Figs. 4C and 5C, respectively. Thus, even in subjects that never were required to attend to the auditory stimulus, the same auditory influences on their visual temporal rate discrimination abilities were observed.

**Influences of auditory bandwidth, auditory intensity, and spatial disparity**

All of the previous analyses were based on visual and 1-kHz tone auditory stimuli being presented at a location along the midline and 4° below the fixation point at reasonably high auditory intensities (65 dB SPL). One question that arises from these results is what stimulus parameters are important for the auditory stimuli to influence the perception of visual stimuli. Similar experiments on the ventriloquism effect have noted that both the spatial and the temporal disparity between auditory and visual stimuli strongly influence the ability of visual stimuli to capture the spatial location of auditory stimuli (e.g., Slutsky and Recanzone 2001). Thus it would not be surprising if these influences of the auditory stimuli on visual temporal rate discrimination were diminished or eliminated if the two stimulus modalities could be clearly differentiated into two distinct “objects.” To investigate how the spatial disparity between the two stimuli, as well as the intensity of the auditory stimuli, may influence these perceptions, a subset of four subjects (A–D) was tested in several additional sessions. In these experiments, the spatial locations of the auditory and visual stimuli were manipulated in three ways. In the first condition, both auditory and visual stimuli were positioned at 8° eccentricity to the left or right (and 4° below the fixation point). In the second condition, the visual stimulus was positioned 8° to the left and the auditory stimulus 8° to the right and vice versa. In the third condition the visual stimulus was positioned at 90° to the left and the auditory stimulus was positioned at 90° to the right. Each of these configurations was tested using auditory stimuli at 65 and 40 dB SPL. In all cases, there was no difference within subjects between conditions where the target was presented to the left or right of the midline. Further, the data from these sessions were not qualitatively different from those where both stimulus modalities were presented on the midline.

To verify these impressions, all eight subjects (A–H) were tested under five different stimulus conditions. In all cases, there was no difference in the results from these experiments, summarized in Fig. 8, from those previously illustrated. In each panel of Fig. 8 the mean (solid black line) and ±SE (dashed black lines) show the fitted functions for the condition described above (65 dB SPL on the midline; 0° Loud). Each colored line shows the function fitted to the mean across subjects for a different experimental condition. The first was

---

**FIG. 5.** Results from the distracter-constant trials while subjects attended to the visual stimulus. Conventions and subjects are the same as in Fig. 3, with the data from these trials shown as open diamonds. If the auditory stimulus did not influence visual temporal rate perception, the data should fall along the dashed line. In every case, subjects performed near chance levels as predicted if the auditory stimulus has a strong influence on visual temporal rate perception (solid line), even at visual temporal rates that they could clearly discriminate during the visual-alone condition.
FIG. 6. Visual stimuli do not influence auditory temporal rate discrimination. Each panel shows the mean results from the different trial types when the same 8 subjects were instructed to attend to the auditory stimulus. All plots follow the conventions of Figs. 3–5, with data for the different trial types shown as filled symbols. Bimodal trials (A) and distracter-constant trials (B) were not different from the auditory-alone trials. However, there was a large difference in performance during target-constant trials (C), when the attended auditory stimulus did not change in temporal rate but the visual stimulus did. In all cases, the data fall close to the solid line, indicating that there was no influence of the visual stimulus on auditory temporal rate discrimination.

FIG. 7. Results from naive subjects. Mean responses to visual-alone trials from 6 naive subjects that never attended to the auditory stimuli and were given a previous block of trials with feedback are shown as open squares and dashed lines. A: results from target-constant trials (triangles) which should be at chance performance if the auditory stimulus had no influence on visual temporal rate perception (compare with Fig. 4C). B: results from distracter-constant trials (diamonds) which should follow the dashed line if there was no auditory influence (compare with Fig. 5C). Some symbols are shown offset for clarity. Thus even naive subjects that have been explicitly trained that the auditory stimulus was unreliable still showed the auditory capture effects.

with both the auditory and the visual stimuli presented along the midline (0° disparity) at 40 dB SPL (0° Quiet; gray line). The next two used the visual stimulus at 8° to the left and the auditory stimulus at 90° to the right (98° disparity) both at 65 dB SPL (98° Loud; red line) and 40 dB SPL (98° Quiet; green line). The last two used the same spatial configuration and intensities, but broadband noise was used as the auditory stimulus instead of a 1-kHz tone (98° Loud Ns; blue line and 98° Quiet Ns; pink line). Figure 8A shows the results from the
auditory-alone trials. Qualitative analysis showed that there were no statistically significant differences between any of these conditions, as evidenced by most of the fitted functions being within ±SE and highly overlapping with the loud midline condition. This was even more apparent for the visual-alone stimuli (Fig. 8B), where the visual stimulus was either 4° below fixation or 4° below and 8° to the left. Again, there were no significant differences across subjects under these different spatial conditions. Figure 8, C and D, shows that the auditory stimulus was equally able to influence the perception of visual temporal rate under these different spatial and intensity conditions. Despite the fact that all of the stimuli presented with 98° disparity are clearly localized to different spatial locations, there were still the same robust interactions where the auditory stimulus could alter the perception of visual temporal rate. This was true for both the target-constant condition (Fig. 8C) and distracter-constant condition (Fig. 8D), as well as the bimodal condition (not shown). In no cases were the functions generated under these spatial and intensity conditions significantly different from those derived when the two stimuli were presented at the midline and the 1-kHz tone stimulus was presented at 65 dB SPL (all \( P > 0.05 \)). In summary, there was no evidence that the ability of the auditory stimulus to influence the perception of temporal rate in the visual domain was reduced under these different spatial, intensity, or auditory spectral bandwidth conditions.

Influences on temporal rate disparity

Given that these effects are so robust across different stimulus parameters, one question raised is what stimulus parameters are necessary to eliminate these effects. To address this issue, two experiments were done in which the temporal rate of the distracter was very different from the temporal rate of the target. The first experiment used a 1-s duration tone or light during both sequences as the distracter stimulus. On these sessions, only two trial types were presented: the single-stimulus and the distracter-constant. Figure 9 shows the mean results from the 0° loud condition when subjects were instructed to attend to the visual stimulus. In this case, the auditory stimulus has no effect on the ability to perform the visual task. This was true across subjects, where the functions for the visual-alone and distracter-constant trials were not significantly different from each other in any of the eight subjects tested (all \( P > 0.05 \)). Further, the bias, slopes, and bandwidth values were not significantly different from each other (all \( P > 0.05 \)). The same result was observed when the subjects were attending to the auditory stimulus while the visual stimulus was presented uninterrupted for 1 s during both sequences, as expected from the previous results (Fig. 9B). These control experiments were also performed with the stimuli presented from the midline at 40 dB SPL, as well as under the 98° disparity condition at both auditory stimulus intensities.
(65 and 40 dB SPL) for both 1-kHz tone and broadband noise stimuli. In all of these experiments, the results were indistinguishable from those shown in Fig. 9, A and B. Thus it is not the case that the presence of an auditory stimulus will always influence visual temporal rate perception. If that were the case, one would expect to see that the subjects would consistently perceive the visual stimulus to be at a lower frequency than it actually was (i.e., closer to the 1 s tone auditory stimulus, nominally 0.5 Hz).

The second control experiment used a higher distracter temporal rate (8.0 Hz) presented during both sequences within a trial. The results across subjects performing the visual task are shown in Fig. 9C. Again, there was no influence of the higher rate auditory distracter on visual temporal rate perception. This was true across subjects as well as within individual subjects. These results indicate that, with large differences in temporal rates, the influence of auditory stimuli on visual temporal rate perception is eliminated.

Given the results of these experiments, it seems clear that there is a dependency of the temporal rate of stimuli for this effect to be generated. It is also still possible that subjects were simply unable to ignore the auditory stimulus on at least some of the more difficult visual trials, even those that had never performed the auditory temporal rate discrimination task. To address these issues, a paradigm was used in which subjects were explicitly instructed to attend to both the auditory and the visual stimuli, and to report whether they perceived the two to be presented at the same or different temporal rates. In this task (Paradigm 2; see METHODS), on all trials only a single sequence of four tone bursts and light flashes were presented. Fifteen different trial types were presented on randomly interleaved trials throughout the session. On three trial types, visual and auditory stimuli were presented at the same rate (3.0, 4.0, and 5.4 Hz). For the remaining trial types, the visual stimuli were presented at 4.0 Hz and the auditory stimuli were presented at rates of 3.0 to 5.4 Hz in 0.2-Hz steps.

Two different stimulus configurations were used in separate sessions. In the first, the auditory and visual stimuli were aligned by the center of the two sequences to maximize the amount of temporal overlap, as was done in preceding experiments. The results of this experiment across all eight subjects (A–H) are shown in Fig. 10A. The insets above the plot show schematized trials, with auditory stimuli presented at a slower rate than the visual stimuli shown to the left side of the plot, and auditory stimuli presented at a faster rate than visual stimuli shown to the right side of the plot. The shaded circles show trials where the visual and auditory stimuli were identical, and the correct response was same (0% different). The open circles show trials where the visual stimulus was presented at 4.0 Hz and the auditory stimulus was presented at a different temporal rate spanning a range of 3.0 to 5.4 Hz. The shaded bar at the bottom of the graph represents the range of frequencies tested in the experiments described above (3.5 to 4.5 Hz; Figs. 2–9). To determine which trial types the subjects perceived to be the same as when both stimuli were presented at 4.0 Hz, statistical analysis (2-tailed t-test) compared the responses when both stimuli were presented at 4.0 Hz against
shown as asterisks above each data point in Fig. 10: single asterisks denote $P < 0.05$; double asterisks denote $P < 0.01$.

This analysis revealed that subjects were able to correctly identify when the two stimuli were presented at the same temporal rates on approximately 85% of the trials (shaded circles; all $P > 0.05$). However, when the auditory and visual stimuli were at different, but similar, temporal rates, subjects often indicated that the two stimuli were presented simultaneously. This was the case even for temporal rates that subjects could reliably perceive as different from 4.0 Hz (e.g., 4.6 Hz) in the previously described target-alone conditions for both auditory stimuli (mean: 100% for 4.5 Hz) and visual stimuli (mean: 93.1 ± 2.8% for 4.5 Hz). Thus even though subjects were explicitly instructed to attend to both stimuli to determine if they occurred at the same or different temporal rates, and these same subjects could reliably discriminate 4.0 from 4.6 Hz for both auditory and visual stimuli, they still consistently perceived these two stimuli to occur at the same temporal rate. This indicates that the influence of auditory stimuli on visual stimuli cannot simply be accounted for by the subject subconsciously attending to one stimulus instead of the other in spite of the instructions. Rather, these results show that the ability of the auditory stimulus to capture the temporal information of the visual stimulus is very pronounced.

Inspection of the curve in Fig. 10A indicates that the influence of auditory stimuli is greater for higher frequencies compared with lower frequencies. This asymmetry was present whether the data were plotted as ratios (difference in temporal rate divided by either the auditory or the visual temporal rate) or on a logarithmic axis (not shown). It is possible that this is due to the way that the two stimulus modalities were configured. To maximize the amount of temporal overlap between the two stimuli, they were aligned at the center each sequence (shown schematically above the plot and in Fig. 1). This results in the auditory stimulus leading the visual stimulus at stimulus onset when the auditory stimulus is presented at a lower temporal rate (left), and visual stimuli leading when the auditory stimulus is presented at a higher temporal rate (right). Previous studies have shown that human subjects are better able to discriminate temporal disparities if auditory stimuli lead visual stimuli resulting in a similar asymmetry (see Slutzky and Recanzone 2001). Thus the asymmetry in the function could be due to this difference in temporal disparity discrimination if subjects were basing their responses on the disparity of the first pip and flash in the sequence of four. To test this possibility, the same subjects were run under the same conditions except that the auditory and visual stimuli were aligned at the stimulus onset. In this case, when the auditory stimulus was presented at a lower rate than the visual stimulus, the visual stimuli always lead after the first tone pip and light flash. In contrast, when the auditory stimulus was presented at a higher rate, the auditory stimuli always lead the visual stimuli after the first tone pip and light flash (inset, Fig. 10B). Thus if the asymmetry of the graph was due to the onset disparity, it should be mirror-reversed across most stimulus configurations. This is likely due to the fact that there was much less temporal overlap between the two stimuli, particularly during the third and fourth elements of each sequence. In cases where there was a large difference in temporal rates, one stimulus modality would continue to be

![Graph](image-url)
presented well after the other stimulus modality had stopped. However, the general shape of the function, particularly the asymmetry, is the same under the two stimulus conditions. This indicates that the subjects were not simply cueing off the differences in the onsets of the individual pips and light flashes.

In summary, these experiments indicate that the ability of the auditory stimulus to influence the visual stimulus is dependent on the temporal disparity between the two stimuli. Further, faster auditory stimuli are better able to capture visual stimuli than are slower auditory stimuli. This may be because visual temporal rate processing deteriorates as the stimuli are presented faster, but this has yet to be explicitly tested.

**Auditory-induced aftereffects of visual temporal rate perception**

One interesting feature of spatial auditory-visual interactions is that they can be long-lasting. This has been termed the ventriloquism aftereffect (Canon 1970; Radeau and Bertleson 1974; Recanzone 1998). In this case, presentation of a consistent temporal disparity between auditory and visual stimuli, for example by wearing prism spectacles, induces a shift in the perception of auditory space in the direction of the visual stimulus. The ventriloquism aftereffect will generally occur only when the ventriloquism effect is present during the training period and can be interpreted as an enduring change in the representation of auditory space following a period of exposure to spatially disparate auditory and visual stimulation. Similar types of aftereffects have not been demonstrated in the temporal domain, or under conditions in which the auditory stimulus has been shown to influence the percept of visual stimuli. The next series of experiments was designed to determine if a similar aftereffect in the temporal domain could be demonstrated.

Subjects were asked to report whether combined visual and auditory stimuli were presented at the same or different temporal rates, as shown in Fig. 10A. Following this session, subjects were asked to perform a slightly different task. In this version, combined auditory and visual stimuli were presented as before, but as one to seven sequences with an approximately 1-s inter-stimulus interval. The subjects were asked to attend to the visual stimulus and to indicate with the movement of a switch when they perceived a dimming of the visual stimulus (Paradigm 3, see METHODS). They were further instructed that they must make the response before the fourth light flash of the sequence to be counted as a “hit.” Feedback was given after each trial and the entire block took approximately 20 min to complete. During these trials, the auditory stimulus was simultaneously presented under four different conditions (the visual stimulus was paired with faster, slower, the same, or randomly interleaved auditory temporal rates, see METHODS). Immediately following this training session, the subjects performed the same/different task once again.

If the training period could produce a long-lasting shift in the perception of visual temporal rate, then the responses during the posttraining sessions should be uniformly shifted in a predictable direction. For example, in the auditory-faster condition, the training period should shift the perception of visual temporal rate to be higher than it actually is. Thus, in the posttraining case, a 4.0-Hz visual temporal rate would be perceived as being higher than 4.0 Hz. This would cause the subjects to be more likely to perceive the auditory and visual stimuli to be the same when the auditory stimulus was presented at a higher temporal rate than the visual stimulus. Similarly, subjects would be less likely to perceive that the auditory and visual stimuli were presented at the same rate when the auditory stimulus was presented at a lower temporal rate than the visual stimulus. This would result in a rightward shift in the function for the posttraining case compared with the pretraining case.

The results of such an experiment are shown for a representative subject in Fig. 11A. In the pretraining condition, there was a characteristic asymmetry in the response function similar to that shown across all subjects in Fig. 10A. However, after training when the auditory stimulus was presented at a higher temporal rate, the expected shift in the curve to the right occurred. This was true across all four subjects tested and the mean data are shown in Fig. 11B. In all cases, there was a clear

![A Single Subject](http://jn.physiology.org/)

**FIG. 11.** Visual temporal rate aftereffect. Results from the same/different task before (circles) and after (squares) approximately 20 min of exposure to auditory stimuli presented at a rate 0.6 Hz faster than visual stimuli. Conventions as in Fig. 10 with shaded symbols corresponding to trials where the auditory and visual stimuli were presented at the same temporal rate. A: results from a single subject. B: mean results from 4 subjects (±SE). Error bars are only shown in 1 direction for clarity. There is a clear shift to the right of the posttraining functions, indicating that after training these subjects perceived visual temporal rates to be faster than before training.
shift of the curves to the right on the order of approximately 0.2 Hz, which is about one-third of the disparity during the training session. This was true for stimulus pairs where the auditory stimulus was slower than 4.0 Hz as well as faster than 4.0 Hz. Thus, after a relatively brief exposure to these temporally disparate stimuli, there was a measurable lasting effect on temporal rate perception.

These data are summarized across subjects as difference functions (post-pre) in Fig. 12A, where the black line shows the mean difference, and the colored lines show the difference functions for each of the four subjects. In all cases there was an increase in the percentage called different when the auditory stimulus was presented at a slower temporal rate and a decrease in the percentage called different when the auditory stimulus was presented at a higher temporal rate than the visual stimulus. A perceptual shift when the training period consisted of auditory stimuli presented at lower temporal rates (0.4 Hz slower) than the visual stimuli was not as clear. In this case, two subjects (red and blue lines) showed a similar, but leftward, shift in their functions, whereas the other two did not. This is evidenced by the flatter difference functions (Fig. 12B) for the auditory-slower condition. In the zero disparity training condition, there were no significant differences between the pre- and posttraining conditions, although two subjects showed a trend toward better performance across all stimulus locations (not shown). Finally, under the control condition, there was no difference between the pre- and posttraining functions for any subject (Fig. 12C). In summary, it was possible to demonstrate an aftereffect when auditory stimuli were presented at a faster temporal rate than visual stimuli during the training period, but this aftereffect was not reliably present when the auditory stimuli were presented at slower temporal rates during the training period.

**DISCUSSION**

The results of this study indicate that there are clear influences on visual temporal rate perception by simultaneously presented auditory distracters. These influences are not heavily dependent on the spatial alignment of the two stimuli, or on the auditory stimulus intensity or bandwidths tested. However, they are dependent on the disparity in the temporal rate between the two stimuli. These effects can also persist after a period as brief as 20 min in which auditory and visual stimuli are presented at a consistent temporal rate disparity within the region of interactions.

One possible interpretation of these results is that the subjects were attending to the auditory stimuli despite instructions to ignore them. It is unlikely that this was the case across all trials and subjects given their reports that they were attempting to follow the instructions to the best of their ability. Further, there was very little individual variability between subjects. However, it is possible that subjects may have responded to the change in the auditory distractor on trials where they were uncertain of the temporal rate of the visual stimulus. This seems unlikely given the distracter (auditory) constant trials when the subjects were performing the unambiguous trials of the visual task. In this case, even though most subjects could reliably detect a change in visual temporal rate between 4.0 and 4.5 Hz, they performed much worse when the auditory stimulus was simultaneously presented at 4.0 Hz (Fig. 5). In addi-

![FIG. 12. Difference functions across different aftereffect training paradigms. The difference between the post- and preresponses (percentage called “different”) for each subject are shown as a different colored line, with the mean shown as a solid black line. A: difference functions for the training with auditory stimuli 0.6 Hz faster than the visual stimuli. There is a significant increase in the percentage of trials subjects responded different when the auditory stimulus was presented at a slower temporal rate, and a significant decrease when the auditory stimulus was presented at a faster temporal rate (compare with Fig. 10B). B: same analysis for training periods where the auditory stimulus was presented at 0.4 Hz slower than the visual stimulus. Two subjects showed the mirror-reversed effect compared with the auditory faster condition, but the other 2 did not. C: same analysis when the training could be 0.6 Hz faster, 0.4 Hz slower, or at the same rate (0 Hz difference). In this case, all subjects showed no consistent difference across temporal rates between the pre- and posttraining conditions.](http://jn.physiology.org/lookup/doi/10.1152/jn.00201.2003)
stimuli, the same effects were noted (Fig. 7). Further, even under conditions in which subjects were explicitly instructed to attend to both stimuli, the auditory capture effects were still robust (Fig. 10). This was true even under stimulus conditions, such as 4.6 Hz, where the subjects could easily discriminate both auditory and visual stimuli as being different from 4.0 Hz. However, under these conditions they were essentially unable to tell that a 4.0-Hz visual stimulus was different from a 4.6-Hz auditory stimulus. Finally, the finding that there was an after-effect indicates that the subjects did experience a shift in their visual temporal rate perception. If they had simply been using the change in the auditory temporal rate to perform the task, there should have been no difference between the pre- and posttraining sessions, as the stimuli were identical between these two conditions. Therefore it seems safe to conclude that the auditory system can have a pronounced influence on visual temporal rate perception in normal human subjects.

A second factor to consider is whether the subjects were cueing on differences in the duration (and/or gaps) of the tone pips or light flashes, or on differences in the onsets of stimuli between the two modalities. For most of the experiments (Figs. 2–8) the four pips and light flashes were aligned by the center of the sequence to minimize the onset differences. However, one cannot change the temporal rate without also changing either the duration or the gap, so this will always confound in this type of experiment. Nonetheless, it seems unlikely that subjects could use this cue exclusively, largely because the experiments using paradigm 2 showed similar influences of the auditory stimuli on visual temporal rate perception even when the stimuli were aligned at the onset (Fig. 10B). For example, the difference in onsets for the fourth pip and flash of the sequences with the visual stimulus presented at 4.0 Hz and the auditory stimulus at 4.6 Hz is almost 100 ms, but subjects perceived this as different on only about 20% of the trials. Further, thresholds of subjects asked to discriminate different onsets of auditory and visual stimuli using a 200-ms-duration stimulus were on the order of 100 ms, or half the stimulus duration (Slutsky and Recanzone 2001). Thresholds would presumably be less for these shorter duration stimuli (125 ms). Thus it seems that while duration and onset difference cues are available to the subject they are unable to use them in performing this task.

An interesting feature of these interactions is that the effects are more robust for faster temporal rates than slower temporal rates. It is unclear how differences between auditory and visual discrimination would manifest themselves at temporal rates different from a 4.0-Hz standard. This standard was chosen as it is clearly discriminable in both modalities and the temporal overlap between stimuli is relatively high over the 1-s stimulus sequence duration. This difference between faster and slower temporal rates was also reflected in the asymmetry of the functions when auditory and visual stimuli were directly compared with each other (Fig. 10). This asymmetry was not simply due to the relative difference in frequency, as the same asymmetry was noted when differences were plotted as ratios of either the auditory or the visual temporal rate. This asymmetry was also in contrast to that observed when subjects were asked to judge whether a single auditory and visual stimulus was temporally aligned (Slutsky and Recanzone 2001) in that there was not a linear offset between the two sides of the curve. This may be due to nonlinear differences between the ability to discriminate visual temporal rates near this comparison frequency (4.0 Hz) under these experimental conditions. If this is true, then the interaction between auditory and visual stimuli is predicted to be greater at higher temporal rates as the visual task becomes more difficult.

The interactions of auditory and visual stimuli described in this study are consistent with previous reports using different experimental designs. Auditory “driving” (Knox 1945) has been studied previously where subjects are asked to match the frequency of a visual “flicker” stimulus to that of an auditory “flutter” stimulus, and the reverse (e.g., Gebhard and Mowbray 1959). These experiments showed that the auditory stimulus could have a strong influence on the perception of the visual flicker, but the visual stimulus had no influence on the perception of auditory flutter in the two subjects they tested. Similarly, auditory driving is also independent of stimulus intensity (Olgivie 1956a,b) as are magnitude estimations of the temporal rate of auditory, visual, and combined auditory-visual stimuli (Welch et al. 1986) and auditory intensity has been shown to influence the perception of visual stimulus intensity (Stein et al. 1996). Finally, there was no difference in the subjective reports that the auditory stimulus was driving the percept of visual stimuli even when the auditory stimulus was displaced by 50° from the foveated visual stimulus (Regan and Spekreijse 1977). More recently, Shams et al. (2000, 2002) have shown that single light flashes are perceived as two flashes if two or more auditory stimuli are simultaneously presented, a phenomenon they term “sound-induced illusory flashing.” They also found that presenting a single auditory stimulus (7-ms duration, 23 ms before the first light flash) with up to four light flashes (17-ms duration, 57 ms ISI) did not influence the perception of the number of light flashes (Shams et al. 2002). This is a considerably different paradigm than the distracter-constant trials used in the present report and may be more analogous to the 3.0-Hz auditory and 4.0-Hz visual temporal rates used in Paradigm 2 (Fig. 10). The results of the present study extend these previous findings to show that auditory system has strong influences on the ability to discriminate visual temporal rate even under conditions when subjects are explicitly instructed to ignore the auditory stimulus, that these influences are dependent on the temporal rate disparity between the two stimuli, and that temporally based aftereffects can be generated in the visual system by auditory stimulation. This study found the same quantitatively defined auditory-visual interactions across different spatial locations, intensities, and spectral content. This finding was extended to several different spatial, intensity, and spectral differences on the assumption that the interactions would decrease as the subjects could reliably perceive the two stimuli as being separate objects. However, the auditory-visual interactions were not dampened even when using broadband noise stimuli and spatial separations of 98° that crossed the midline, despite the fact that these stimuli were clearly perceived as originating from different locations. Further, attenuating the auditory stimulus to a quiet, but still above threshold, level did not reduce the interactions. This is in agreement with a previous study on temporal rate perception (Welch et al. 1986) but is in contrast to the spatial interactions, the ventriloquism effect, which is dependent on the spatial disparity of the auditory and visual stimuli (Jack and Thurlow 1973; Slutsky and Recanzone 2001). Thus only the disparity in temporal rate between the two

J Neurophysiol Vol 89 February 2003 www.jn.org
stimuli appeared to limit the interactions, indicating that the auditory system has a nonspatial and nonspectral influence on visual temporal rate perception.

The finding that auditory stimuli have a clear influence on visual temporal rate perception is consistent with the modality appropriateness hypothesis (Welch and Warren 1980). This hypothesis posits that the sensory system that has the greatest acuity for the stimulus parameter being discriminated will dominate the perception of bimodal stimuli. For spatial tasks, the visual system has greater acuity than the auditory system and therefore the ventriloquism effect results. In this temporal rate discrimination task, the auditory system has greater acuity (Fig. 2) and therefore dominates the perception of the visual stimulus. This “dominance” is not necessarily complete, as previous studies using different methods indicate that for spatial interactions, the visual stimulus can account for approximately 80% of the localization of auditory stimuli (see Welch 1999; Welch and Warren 1980). The two-alternative forced-choice paradigms, the frequency range tested, and analysis methods used in this study would probably not be able to differentiate between similar magnitudes of these effects (e.g., 80 vs 100%). Thus, rigorous quantification of the magnitude of these interactions, particularly as a function of the comparison frequency, awaits further study.

How exactly these auditory and visual interactions are accomplished by the CNS is currently unclear. One possibility is that multi-modal neurons combine the information between the two modalities to generate the percept of a single object. Candidate structures based on studies in macaque monkeys would include regions that respond to both auditory and visual stimuli within the parietal lobe (e.g., see Andersen 1997; Cohen et al. 2002; Linden et al. 1999), the superior temporal sulcus (e.g., Bruce et al. 1981; see Cusick 1997), or frontal lobes (e.g., Graziano et al. 1999; Russo and Bruce 1994; see Graziano and Gross 1998). Recent studies in the auditory cortex of the macaque monkey suggest that there are different streams of processing “what” and “where” information (Kaas and Hackett 2000; Rauschecker 1998; Rauschecker and Tian 2000) that integrate with similar what and where information from the visual system in the frontal cortex. The lack of spatial, spectral, and intensity dependence on this effect would indicate that this integration more likely occurs within more rostral regions of the superior temporal gyrus and sulcus as well as the frontal cortex than in the parietal lobe. However, a recent functional imaging study has revealed that the prefrontal cortex and insula, as well as the posterior parietal cortex, show increased activity levels to temporally disparate auditory and visual stimuli (Bushara et al. 2001). It is also possible that feedback from these multi-modal areas would result in changes in neuronal responses in unimodal cortical areas, potentially as low as the primary visual and auditory cortices. However, no changes in visual-evoked responses were noted between conditions in which there was “auditory driving” and under conditions where the same visual temporal rates were presented, but the auditory stimulus did not drive the percept of the visual temporal rate (Regan and Spekreijse 1977).

Regardless of the neural substrate of these effects, they can also be long-lasting as demonstrated by the aftereffect (Figs. 11 and 12). The most pronounced and reliable aftereffect occurred when the 20-min training period paired auditory stimuli presented at a temporal rate 0.6 Hz faster than the visual stimuli. This value was chosen because most subjects responded that a 4.6-Hz auditory stimulus was the same as a 4.0-Hz visual stimulus (Fig. 10). During the training period, all subjects reported that they perceived the two stimuli to be presented simultaneously. The aftereffect was statistically reliable, but represented a shift of only approximately 0.2 Hz. This is generally less than the spatially based ventriloquism aftereffect (e.g., Canon 1970; Radeau and Bertelson 1974; Recanzone 1998; Welch 1999). The temporal aftereffect was also not reliably observed when the training period consisted of auditory stimuli presented at a lower temporal rate. The temporal rate was chosen as 0.4 Hz slower because pilot experiments using a difference of 0.6 Hz resulted in no effect and all subjects reported that they could tell that the auditory and visual stimuli were presented at different rates. Figure 11 shows that most subjects could discriminate an auditory temporal rate 0.6 Hz slower than a visual temporal rate on about 75% of the trials. The difference of 0.4 Hz was chosen because it was only detected approximately 40% of the time (in contrast, auditory rates 0.6 Hz faster than visual rates were detected as different on only approximately 15% of the trials). If one assumes that an aftereffect is only 30–40% effective, it is unlikely that these methods could detect a 0.1-Hz shift.

This study demonstrates a visual aftereffect that is induced by auditory stimuli and also demonstrates temporally based aftereffects. These results, as with the ventriloquism aftereffect, indicate that a lasting change in the neural representation of both space and time can be generated over a relatively short period of exposure to bimodal disparities. This further implies that the bimodal and multi-modal representations of real-world objects are extremely dynamic, presumably to account for real-world differences in both the physical attributes and the processing capabilities of each sensory system. In the example of temporal rate, the temporal disparity between the arrival of the visual and auditory stimuli will depend on the distance to the observer. It appears that the nervous system uses the temporal information that is accurately encoded by the auditory stimulus to drive the perception from the less accurately encoded visual events. For spatial information the opposite occurs, where the accurately encoded visual information is used to drive the perception of the less accurately encoded auditory events. Thus taking advantage of the strengths of information encoding by each stimulus modality may be a general mechanism by which multi-sensory features of single objects are bound into a single representation. Illusions such as those described in this study, and the corresponding aftereffects, may provide a useful tool in elucidating the neuronal mechanisms of multisensory integration.

The author thanks T. M. Woods, K. H. Britten, M. L. Sutter, and L. A. Krubitzer for comments on earlier versions of this report, and K. H. Britten and H. Heuer for assistance with the data analysis.

This work was funded by National Eye Institute Grant 1 R01 EY-013458 and in part by Core Grant EY-12576.

REFERENCES


AUDITORY INFLUENCES ON VISUAL PERCEPTION


