REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE 7 MAY 97

3. REPORT TYPE AND DATES COVERED

4. TITLE AND SUBTITLE
SPATIAL DISPARITY EFFECTS ON REACTION TIMES TO DUAL AUDITORY AND VISUAL STIMULI

5. FUNDING NUMBERS

6. AUTHOR(S)
LAWRENCE KENT HARRINGTON

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
UNIVERSITY OF MISSOURI-ST LOUIS

8. PERFORMING ORGANIZATION REPORT NUMBER
97-030

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
DEPARTMENT OF THE AIR FORCE
AFIT/CI
2950 P STREET
WRIGHT-PATTERSON AFB OH 45433-7765

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
DISTRIBUTION STATEMENT A
Approved for public release
Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

19970512 138

14. SUBJECT TERMS

15. NUMBER OF PAGES 49

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT

18. SECURITY CLASSIFICATION OF THIS PAGE

19. SECURITY CLASSIFICATION OF ABSTRACT

20. LIMITATION OF ABSTRACT

Spatial Disparity Effects on Reaction Times to Dual Auditory and Visual Stimuli

Lawrence Kent Harrington

Masters Thesis
University of Missouri-St. Louis
School of Optometry

Spatial Disparity Effects on Reaction Times to Dual Auditory and Visual Stimuli

Abstract

Saccadic reaction times to spatially and temporally coincident auditory and visual targets are much shorter than the reaction times to either of the individual targets alone. Because of the magnitude of this facilitation, neural summation (convergence of the auditory and visual signals in the neural pathway) is considered to be at least partially responsible for the reduction. In this study, which introduced spatial disparity into the dual (auditory and visual) stimulus presentation, a gradual increase in reaction time with increasing disparity was found. Evidence for multisensory convergence, also known as neural summation, was found over a wide range of disparities. In one of my three subjects evidence for inhibition was found at the largest disparity tested. I discuss the similarities between my findings and the known characteristics of neurons found in the middle to deep layers of the superior colliculus.

Introduction

We have highly developed senses which allow us to observe, interpret and interact with our external environment. These senses are complimentary in that they work with different and largely independent stimuli. We can see things we can't hear (snow falling), we can hear things we can't see (noises in the dark) and when we observe an event with multiple senses we can combine sensory information to create a more accurate percept of the world. When we both see and hear a person talking, we can combine the visual information (facial expressions, mannerisms) with auditory information (words, voice inflections) to better interpret their communication. However, if two or more senses are to provide complementary information about a single event we must be able to link or coordinate our senses temporally and spatially. Failure to coordinate our senses creates the illusion of two separate events instead of one. An illustration of bisensory illusion occurs when the difference between the speed of light and the speed of sound causes one event, the environmental discharge of static electricity, to be perceived as two events, lightning and thunder.

Several of our senses, vision, hearing, and touch, not only require temporal coordination but spatial coordination also. In the example of the person talking, we need to be able to see the person talking and hear the person talking in the same location. This ability of our central nervous system to coordinate both the temporal and spatial characteristics of different stimuli into a single, unified awareness is essential for accurately interpreting our environment.

Gradually the picture of how the multiple senses are integrated is being pieced together. Tract tracing methods, single cell recording techniques, and psychophysical testing have contributed to our current knowledge. Tract tracing techniques are useful in delineating areas of the central nervous system where multisensory convergence is likely to occur. Then single cell recordings can demonstrate whether there is multisensory convergence on to individual neurons. In a study of multisensory neurons in the superior colliculus of cats, single cell recordings were used to map both auditory and visual receptive fields (Meredith and Stein, 1996). These auditory and visual receptive fields were found to be spatially coincident with an average receptive field overlap of 86%.

While tract tracing and single cell recordings help delineate the neural pathways, psychophysical studies are required to describe the behavioral consequences of multisensory convergence. For example, a variation of threshold testing has been used to show that two near threshold stimuli from different sensory systems can be combined, resulting in a strong suprathreshold response (Stein et al, 1989). Another useful psychophysical technique in studying multisensory interactions is reaction time in combined or dual stimuli trials. Woodworth and Schlosberg(1954) proposed that reaction times could be broken down into two components: an irreducible minimum and a reducible margin. The irreducible minimum was the optimal response time, under ideal conditions, which could not be improved. The irreducible minimum is based on the minimum time it takes for all processing involved in making a given response to a particular stimulus, including receptor stimulation, conduction to and from processing centers, and muscle contraction.

Theoretically, evidence for a intersensory interaction would present if, in a dual stimuli trial, a reaction time was measured which was less than the irreducible minimum for either of the individual stimuli. If intersensory interaction exists, the most probable location for the interaction is in the conduction to and from processing centers. This multisensory interaction or convergence in the neural pathways is also known as neural summation or coactivation. While establishing an irreducible minimum for any type of stimulus maybe difficult, a similar argument can be made using cumulative probability density functions (CDF), which plots time versus probability that a given RT will be less than the specified time. Because the CDF argument is central to my experiment a short discussion of dual reactions and CDFs is indicated.

Originally it was thought that the reaction time (RT) to dual stimuli, for example light and sound, was simply the shortest of the two individual stimulus reaction times, RT(l) or RT(s) (Woodworth and Schlosberg 1954, Raab 1962). This assumption is accurate as long as there is a significant difference between RT(l) and RT(S).

$$\mathrm{RT}(l+s)=\mathrm{RT}(l) \quad \text{if } \mathrm{RT}(l) < \mathrm{RT}(S)$$
$$\mathrm{RT}(l+s)=\mathrm{RT}(s) \quad \text{if } \mathrm{RT}(s) < \mathrm{RT}(l)$$

However, when RT(l) and RT(s) are approximately equal RT(l+s) is significantly shorter than RT(l) or RT(s) (Hershenson 1962).

$$RT(l+s) < RT(l) \text{ or } RT(s) \text{ if } \quad RT(l) \sim RT(s)$$

Raab refers to this special case, RT(l) ~ RT(s), as physiological synchrony, because the neural events resulting from the visual and auditory stimuli are contemporaneous. Experimentally physiological synchrony can be created by using concurrent stimuli with similar mean RTs, or by taking stimuli with different mean RTs and staggering the stimulus presentation. In this case the difference between the mean RTs for the two stimuli could be used to give the slower stimulus a head start (Hughes et al 1994, Hershenson 1962).

Under physiologically synchronous conditions, the RT(s+l) can be triggered by RT(s) or RT(l). In this case a slow RT(s) may be "over taken" by a faster RT(l) or, alternatively, a slow RT(l) maybe over taken by a faster RT(s). In other words, the auditory and visual stimuli create a race between their respective sensory systems with the winner triggering the response. The reduced RT(l+s) compared to RT(s) and RT(l), resulting from this race is refereed to as "statistical facilitation" (Raab 1962). These observations have been incorporated into a "race model" of multisensory RT (Kornblum 1973, Meijers and Eijkman 1977, Miller 1986).

The above discussion implies that in the processing race, the auditory RT and the visual RT are independent. However, it is possible that the separate channels compete for neural resources and that RT(l) and RT(s) are not independent (Duncan 1980). A series of race models can be created to account for different levels of dependence between RT(l) and RT(s) (Meijers and Eijkman 1977). When extreme values for positive and negative dependence are used, a boundary for reaction times consistent with the race models can be established. Violations of this boundary provide psychophysical evidence that the different sensory channels converge into a common sensory pathway.

The cumulative distribution function (CDF) plots time versus probability that a given RT will be less than the specified time. The CDF starts at zero for times less than the irreducible minimum and approaches one for longer times.

In the case of dual stimuli, the upper limit for the CDF, consistent with the race models, would be the summation of the two individual stimulus CDFs. Miller states the above probability concept for race models in equation form:

$$\begin{aligned} P(RT < t\backslash S1 \text{ and } S2) &= P(RT < t\backslash S1) + P(RT < t\backslash S2) \\ &\quad - P[(RT < t\backslash S1) \text{ and } (RT < t\backslash S2)] \end{aligned}$$

Where $P(RT < t\backslash S1 \text{ and } S2)$ equals the dual stimuli CDF, $P(RT < t\backslash S1)$ and $P(RT < t\backslash S2)$ are the individual stimulus CDFs, and the last term reflects the correlation between the two individual stimulus RTs. Since P[(RT < t\S1) and (RT < t\S2)] must be greater than or equal to zero, it follows that

P(RT < t\S1 and S2) < P(RT < t\S1) + P(RT < t\S2)

Two comments should be made about this probability summation limit for race models. One, probability summation is an extreme limit which is unlikely to be approached under race conditions.

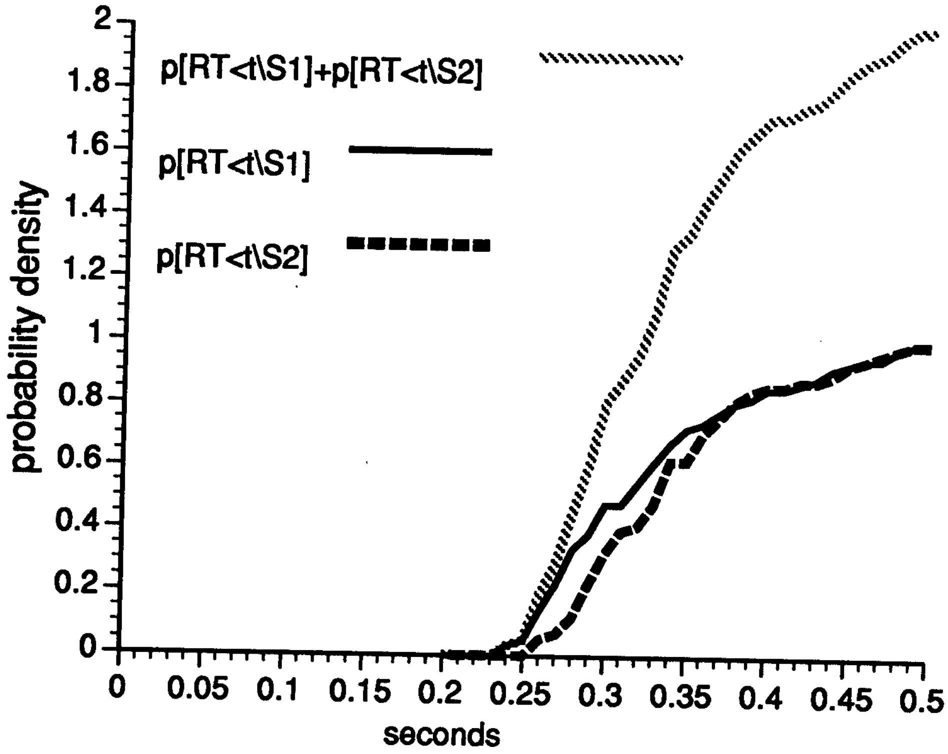

Figure 1: Summation of cumulative probability density function (CDF) for two different stimuli. p[RT<t\S1] and p[RT<t\S2] are the individual stimulus CDFs. p[RT<t\S1]+p[RT<t\S2] is the sum of the individual stimulus CDFs and therefore approaches a probability of 2. The summation CDF serves as the boundary for dual stimulus reaction times constant with race models. Violations of this boundary are evidence of neural summation.

Therefore <u>any</u> violation of probability summation argues strongly for coactivation. Second, probability summation approaches a maximum of two instead of one so that any violation of probability summation would have to occur at very short reaction times (Miller 1982). Single stimulus reaction times vary with many factors including stimulus intensity, environment, subject age, practice and response requested. Hughes et al(1994) looked for violations of the race model using manual undirected (button push), manual directed (joy stick) and saccadic responses. While all three types of responses demonstrated violations of the race model, the violations were the most robust for saccades. The introduction of a second stimulus creates the additional factors which can influence reaction times. These additional factors include temporal and spatial disparity. It is well known that because of differences in the speed of light and sound, and because of differences in processing time in the auditory and visual systems, the nervous system must have a way of combining asynchronous stimuli. Temporal disparity is an intriguing issue and in the case of dual auditory and visual stimuli it has received some attention (Miller, 1986). The issue of spatial disparity on the other hand has been largely overlooked. Frens et al (1995) briefly addressed this issue in a recent study of auditory and visual interactions. They found that the reduced latency for saccadic reaction times to spatially and temporally aligned targets was gradually lost as the spatial or temporal alignment changed.

The purpose of this study was to more thoroughly determine how spatial disparity effects saccadic reaction times to dual, auditory and visual, stimuli. In addition I sought to find out how spatially disparate the stimuli could be while maintaining evidence for neural summation. I had the long term goal, once I had demonstrated the legitimacy of technique, of mapping fields of multisensory neural summation. These behaviorally determined neural summation fields along with the receptive fields determined by single cell recordings in non-human species should lead to a better understanding of how the CNS combines information from the various sensory modalities into a behaviorally relevant unified perception of our environment.

Methods

Subjects

Three volunteers with self reported normal auditory function served as subjects. All subjects had received recent eye exams which included visual acuity and ocular motility tests. One of the subjects was the experimenter and two of the three subjects were informed as to the purpose of the study. Each subject was introduced to the experimental technique, and screened for the ability to complete the task, by performing 300 practice trials, divided evenly over three days. Subjects were asked to participate in the experiment with full knowledge that they could withdraw at any time. All subjects were required to sign informed consent forms.

Stimuli

Visual targets consisted of 3 mm diameter amber light-emitting diodes (LEDs), with an intensity of 200 cd/m^2 and duration of 1 sec. Thirty three millimeter diameter speakers were used to present a white noise auditory target of approximately 54 dB, with 1 sec duration. These targets were displayed in a dimly lit room with a black background and a baseline auditory noise level of less than 40 dB. Presentation of the targets was controlled by a personal computer which cued a custom built stimulus generator. The speakers were attached to an arc segment frame of radius 1.14 meter. LEDs were mounted in a 4 mm tube which held the LED 1 cm in front of the arc segment, allowing the speakers to be placed at various locations behind the LEDs. Subjects were seated in a comfortable chair with their head supported by a head rest to minimize head movement. The chair was placed so that the subject's face was at the center of radius of the arc holding the targets.

Experimental Procedure

Experiment one

The purpose of the first experiment was to verify that my technique was able to obtain a dual stimuli CDF greater than probability summation. I measured saccadic reaction times to auditory, visual and dual stimuli. One pair of visual and auditory targets was placed horizontally 20°, to the left and right of the central fixation LED. Subjects were asked to make prompt and accurate saccades to the targets. Eye movements were recorded by electrooculography (EOG), using silver-silver chloride electrodes attached near the outer canthi of each eye. Eye position output was filtered and amplified using a Grass Model 7P122F amplifier. Data were digitized at 200 *Hz* with a Mac Adios II 12 bit A/D converter and stored for subsequent analysis.

The RT was defined as the interval between stimulus presentation and the initiation of the saccade to the stimulus. Thirty trials for each stimulus type were run for a total of 90 trials per session. Before each session the subjects were asked to make a series of saccades to the 20° target location to document the amplitude of EOG signal and to verify that the signal to noise ratio exceeded my requirements. Usually subjects would sit for only one recording session per day, with each session lasting approximately thirty minutes. Each recording session was divided in half by a short (~5min) break. To minimize order effects, an ABCCBA testing strategy was used: During the first half of the session, 15 LED-only trials were followed by 15 speaker-only trials , and finally 15 dual (auditory and visual) trials; in the second half, the blocks of stimuli were presented in reverse order (dual, speaker-only, LED-only). CDFs were then generated for the three stimulus conditions.

Experiment two

Experiment two was similar to experiment one except that spatial disparity was introduced as an independent variable. The visual targets were located at 20° as in experiment one, while the auditory targets were moved from the visual target to create the disparity. On successive days the speakers were moved centrally, usually in 5° increments. I continued this strategy until the speaker had crossed the midline and reached a point opposite the LED. Difficulty in placing the speakers at center fixation initially kept us from running trials at 20° disparity and therefore trials were completed at 17.5° and 22.5° instead. Later a small adjustment to the apparatus allowed the testing of subjects one and three with a speaker at center fixation. These trials are discussed later as a special case, where the speaker provided no directional information.

Subjects were told to make saccades to the 20° LED position for all dual stimuli trials. On the speaker-only trials the speakers were placed at 20° for all recording sessions. At small disparities experiment two was similar to experiment one and no additional practice trials were run. When the speaker approached the center the subjects reported increased difficulty performing the dual stimulus task and were given one session of 100 additional practice trials before proceeding. At each disparity level 90 trials were run (30 LED, 30 speaker, 30 dual) in ABCCBA order.

Data Analysis

Eye movements were analyzed on a trial-by-trial basis using an interactive computer program to search for saccades. Eye position records were differentiated using a two-point central difference algorithm to yield an eye velocity signal. All trials were scanned visually by the experimenter and a suitable velocity threshold criterion was used to detect saccades. The computer then calculated and produced files for latency, amplitude and duration of saccades. These values were used to cull any saccades that were unlikely to be in response to the targets. Saccades with a latency less than .15 seconds or more than .5 seconds were infrequent and widely separated from the majority of reaction times (Figure 2). In addition, visual inspection of the eye position plots for reaction times shorter and longer than these values showed that they were inappropriate in direction or magnitude.

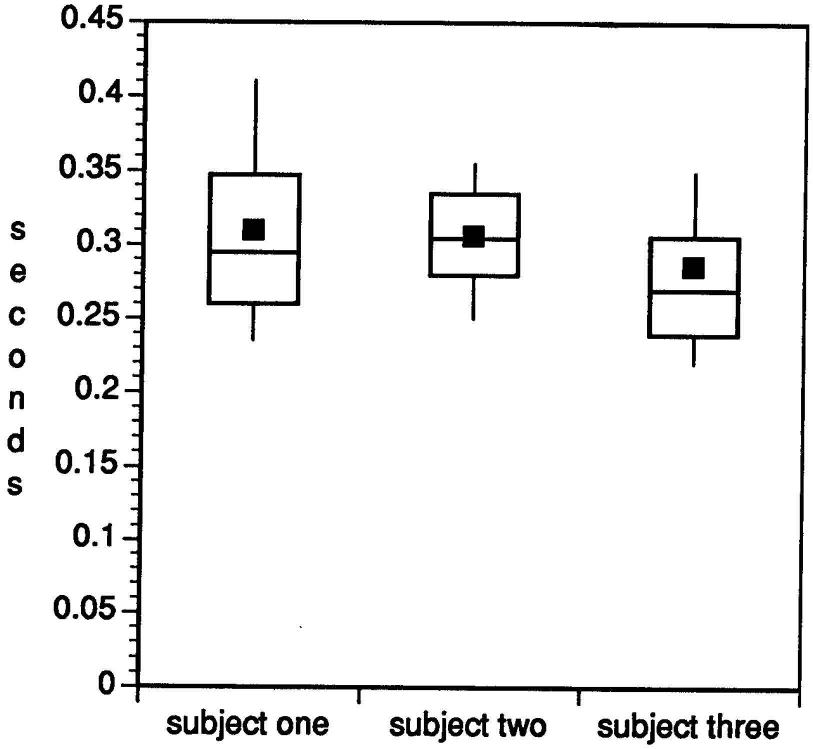

Figure 2: box plots of unfiltered reaction times from experiment one. The black square represents the mean. The box encloses the 25th to the 75th percentile and is divided by a horizontal line at the 50th percentile. The vertical lines extending from each box stop at the 10th and 90th percentile. For each subject I combined data obtained in the LED-only, speaker-only and dual stimulus conditions. Although subjects differed in mean values and variability note that reaction times of more than .5 seconds and less than .15 seconds were rare in each subject.

If more than one saccade was recorded per trial, the latency of the saccade with the most reasonable amplitude and duration was retained. On rare occasions when several saccades from a trial seemed reasonable in terms of latency, amplitude and duration, the entire trial was deleted from analysis.

Latency measures were summed into histograms with a 10-ms bin width to compare the different distributions. Cumulative distribution functions (CDFs) were calculated for the dual stimulus condition and compared to the sum of the speaker-only and LED-only CDFs. An analysis of variance with subsequent Fisher's test was performed on the reaction time distributions for each subject, with the two tailed significance level set at .05.

Results

Spatially aligned visual and auditory targets

After three hundred practice trials all three subjects had easily learned to detect the stimuli and make the appropriate saccades. Their adeptness at the task and the required directional decision kept anticipations to a minimum. For each subject anticipatory saccades were made on less than 1% of all trials. Inappropriately long reaction times were also infrequent for each subject (< 2% of trials).

	subject one	subject two	subject three
mean LED RT	0.336	0.325	0.307
std dev LED RT	0.058	0.046	0.044
mean spkr RT	0.322	0.315	0.27
std dev spkr RT	0.066	0.033	0.043
mean dual RT	0.274	0.277	0.247
std dev dual RT	0.054	0.031	0.033

Table I

All three subjects were similar in that the speaker-only reaction times were shorter than the LED-only reaction times, however there was a large overlap in the reaction time distributions. The overlap was sufficient to consider the stimuli physiologically synchronous and, as expected in this condition, the dual stimulus reaction times were shorter than the individual stimulus reaction times, for all three subjects(Table 1). This facilitation of the dual stimulus reaction time could have been statistical in nature (a result of the sampling from the overlapping LED-only and speaker-only distributions), or it could have resulted from

Figure 3: subject one, (upper plot) Cumulative distribution functions for the dual stimuli condition and the sum of the LED and speaker-only conditions. (lower plot) Difference between the two CDFs (dual minus sum). Violations of the statistical facilitation boundary are shown in two ways: In the top graph, when the dual probability curve is located to the left of the sum probability curve, and in the lower graphs, by the difference plot crossing above zero. There is a large violation between .2 and .29 seconds with a peak around .24 seconds

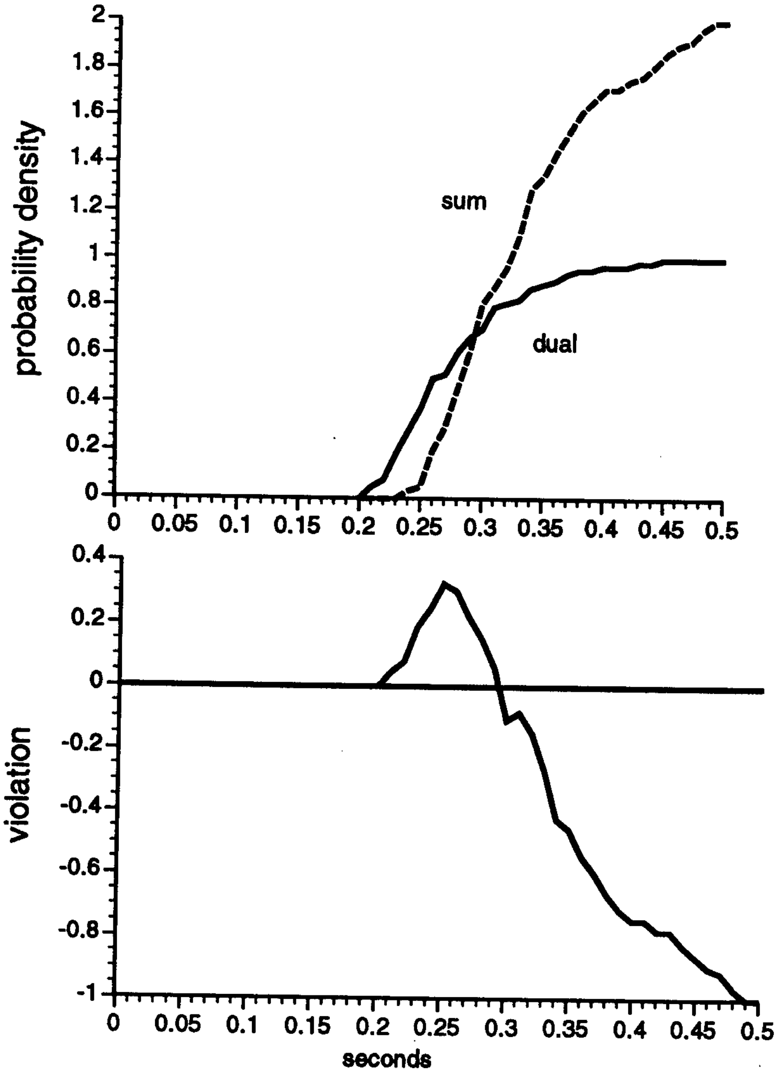

Figure 2

subject one
no disparity in dual
stimulus trials

Figure 4: subject two, (upper plot) Cumulative distribution functions for the dual stimuli condition and the sum of the LED and speaker-only conditions. (lower plot) Difference between the two CDFs (dual minus sum). Violations of the statistical facilitation boundary are shown in two ways: In the top graph, when the dual probability curve is located to the left of the sum probability curve, and in the lower graphs, by the difference plot crossing above zero. There is a large violation that occurs between .25 and .32 seconds with a peak around .29 seconds

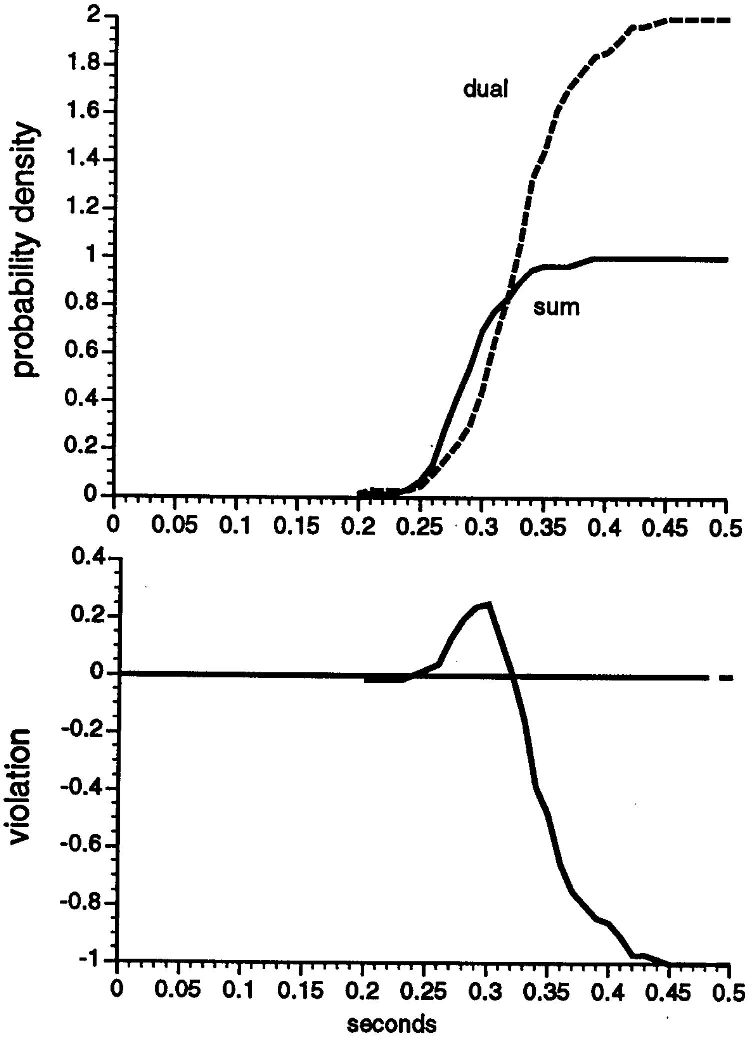

Figure 4

subject two
no spatial disparity in dual
stimulus trials

Figure 5: subject three, (upper plot) Cumulative distribution functions for the dual stimuli condition and the sum of the LED and speaker-only conditions. (lower plot) Difference between the two CDFs (dual minus sum). Violations of the statistical facilitation boundary are shown in two ways: In the top graph, when the dual probability curve is located to the left of the sum probability curve, and in the lower graphs, by the difference plot crossing above zero. There is a large violation that occurs between .2 and .27 seconds with a peak around .24 seconds

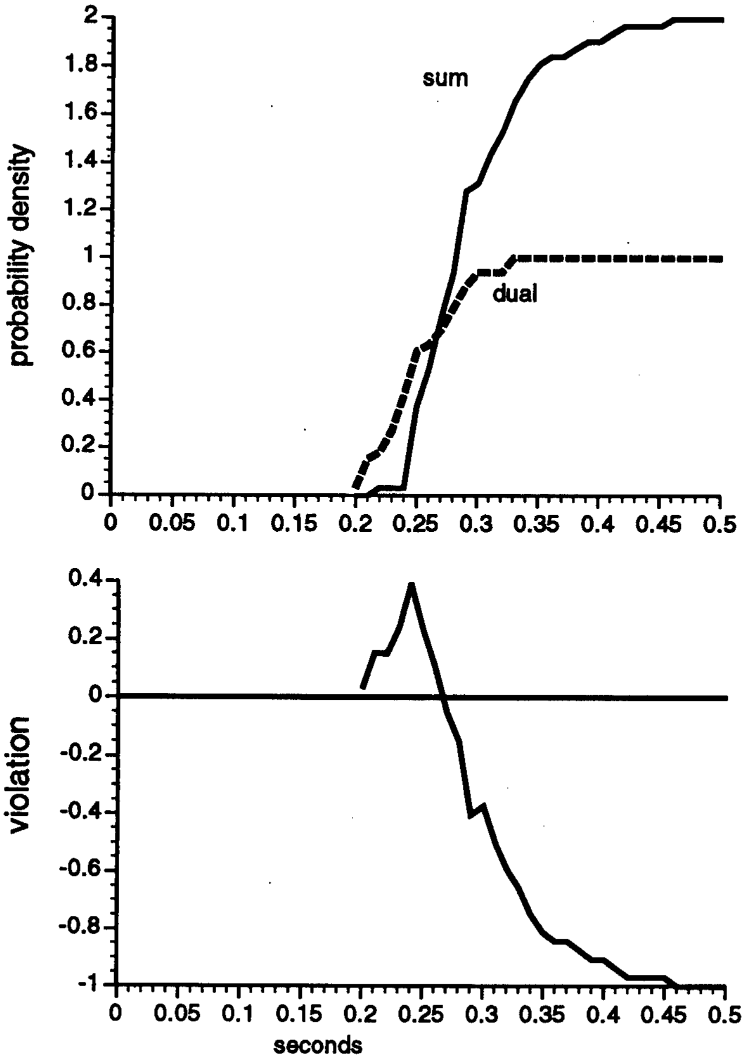

Figure 5

subject three
no disparity in dual
stimulus trials

summation in the neural pathways. To check for evidence of neural summation the CDFs were inspected.

Figures 3,4 and 5 compare the dual stimulus CDF to the sum of the individual stimulus CDF for the three subjects. In each figure the upper plot shows the CDFs and the bottom plot shows the difference between the two CDFs (dual minus sum). Violations of the statistical facilitation boundary are shown in two ways: In the top graph, when the dual probability curve crosses to the left of the probability curve for the sum, and in the lower graphs, when the difference plot is greater than zero. Vigorous violations, indicative of neural summation, were found for all three subjects.

In addition to the visual inspection of the CDF an analysis of variance was performed on the reaction times for each stimulus condition. In each subject the F value was significant, with a probability $\leq .05$. Subsequently a Fisher's Protected Least Significant Difference test was performed on the data for the auditory vs. dual conditions. The results of this test showed that the dual stimuli reaction times were significantly shorter than the reaction times for the speaker-only condition. P values were <.0001 for subject 1, .0002 for subject 2 and .0236 for subject 3.

Non-aligned targets

In Experiment two, spatial disparity between the LEDs and the speakers was introduced as an independent variable. On successive days, the speakers were moved centrally, in 5^o increments starting from the 20^o LED positions. As in experiment one, the subject's responses were reliable and prompt.

Practice effects were apparent in my data as reaction times continued to drop throughout the experiment (Figure 6). The persistent practice effect prevented the combining of data across levels of disparity and resulted in a loss of information about the shape (normal, unimodal, multimodal) of the reaction time distributions. Histograms of the reaction times at individual levels of disparity were uninterpretable, in regard to modality, as they were constructed from a relatively small number of trials.

For each level of disparity, I generated cumulative distribution functions and computed the difference between distribution graphs. These plots showed a gradual decrease in violation with increasing disparity. The disparity range over which violations were demonstrated was quite large, exceeding 25^o in all three subjects. Each subject showed violations at disparities that crossed the midline(i.e., when speakers and the LEDs were on opposite sides of the fixation target). Figures 7 and 8 are the difference graphs for subjects one and two, demonstrating the gradual loss of violation. At small amounts of disparity the difference in

Figure 6: Box plots of reaction times for different stimulus conditions and across different disparities for subject 2. The box encloses the 25th to the 75th percentile and is divided by a horizontal line at the 50th percentile. The vertical lines extending from each box stop at the 10th and 90th percentile. Reaction times for each level of disparity were taken in order, starting under aligned conditions (0 disparity) and proceeding to 40° disparity. The gradual reduction in latency over disparities and therefore time, demonstrate a persistent practice effect throughout the experiment.

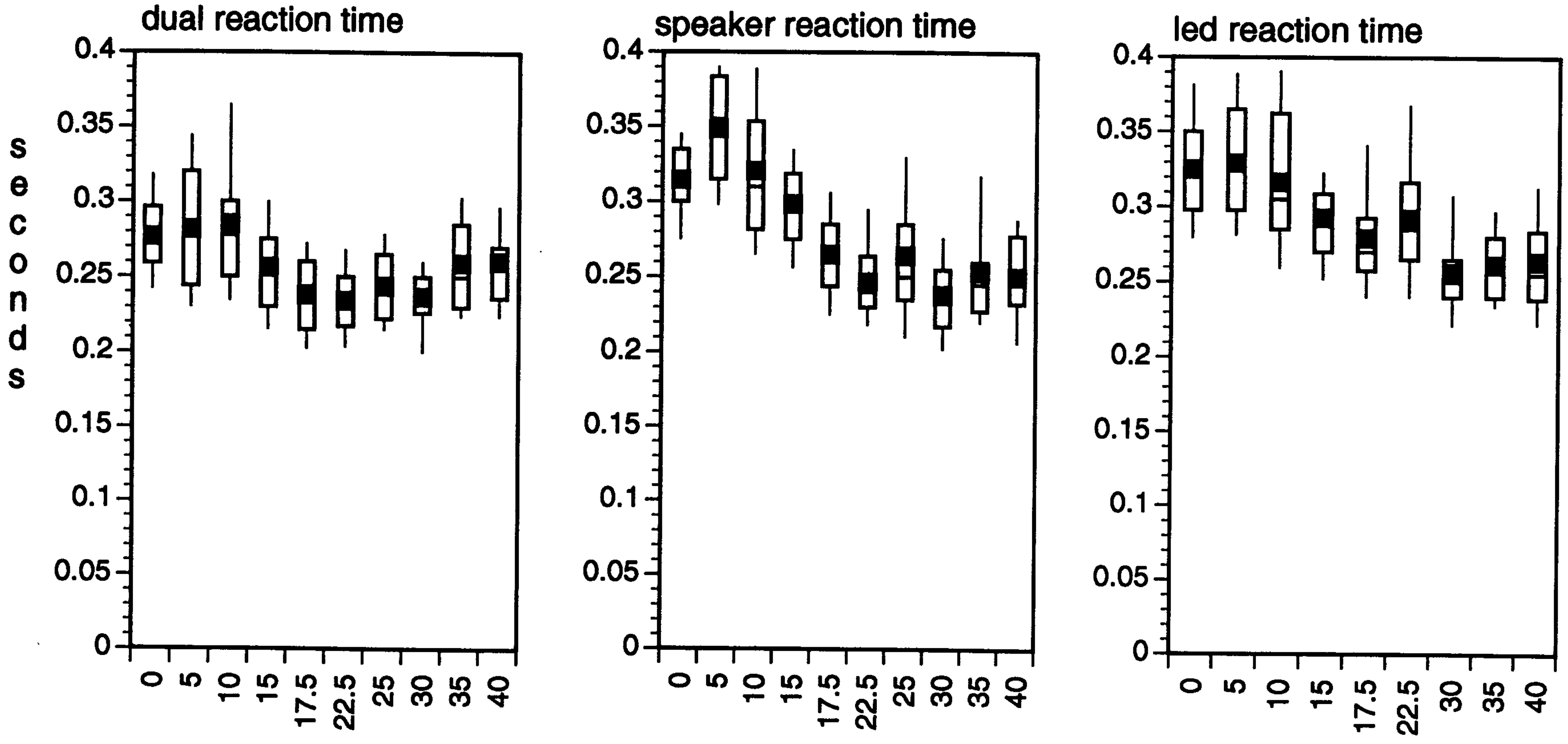
dual reaction time
speaker reaction time
led reaction time
seconds
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
5
10
15
17.5
22.5
25
30
35
40
disparity

Figure 6

Figures 7: Difference plots for subject one across disparities. Violations (evidence of neural summation) are demonstrated when the difference curves greater than zero. The violation is large in the aligned condition, is small to moderate between 10° and 22.5°, and is lost between 25° and 30° of disparity.

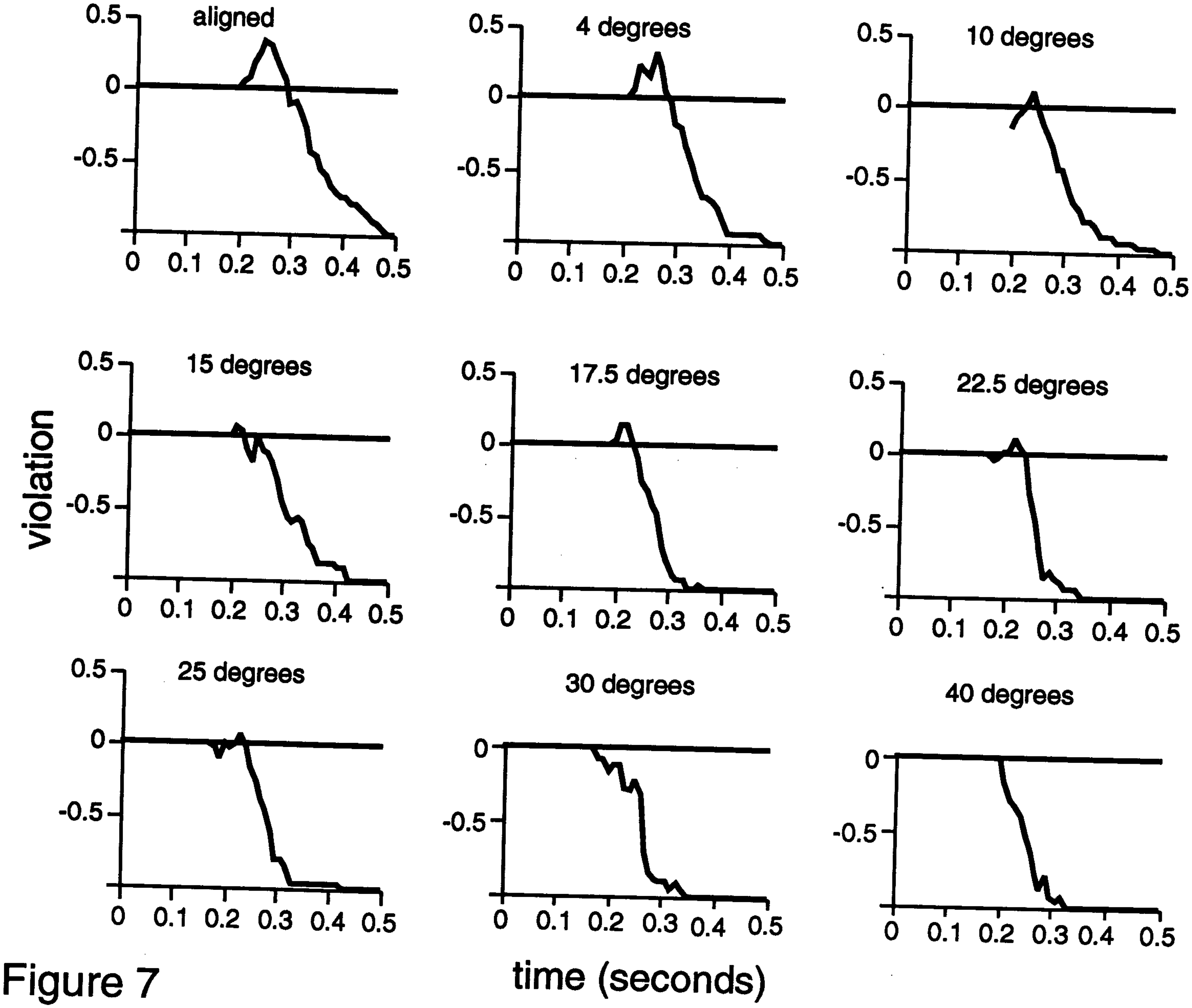
aligned
4 degrees
10 degrees
15 degrees
17.5 degrees
22.5 degrees
25 degrees
30 degrees
40 degrees
0.5
0
-0.5
0 0.1 0.2 0.3 0.4 0.5
violation
time (seconds)

Figure 7

Figures 8: Difference plots for subject two across disparities. Violations (evidence of neural summation) are demonstrated when the difference curve is greater than zero. The violation is large in the aligned condition and becomes gradually smaller with increasing disparity. Evidence for neural summation is lost between 30° and 35° of disparity. Note that more robust violations are seen at intermediate disparities (10°-17.5°) in subject two than in subject one.

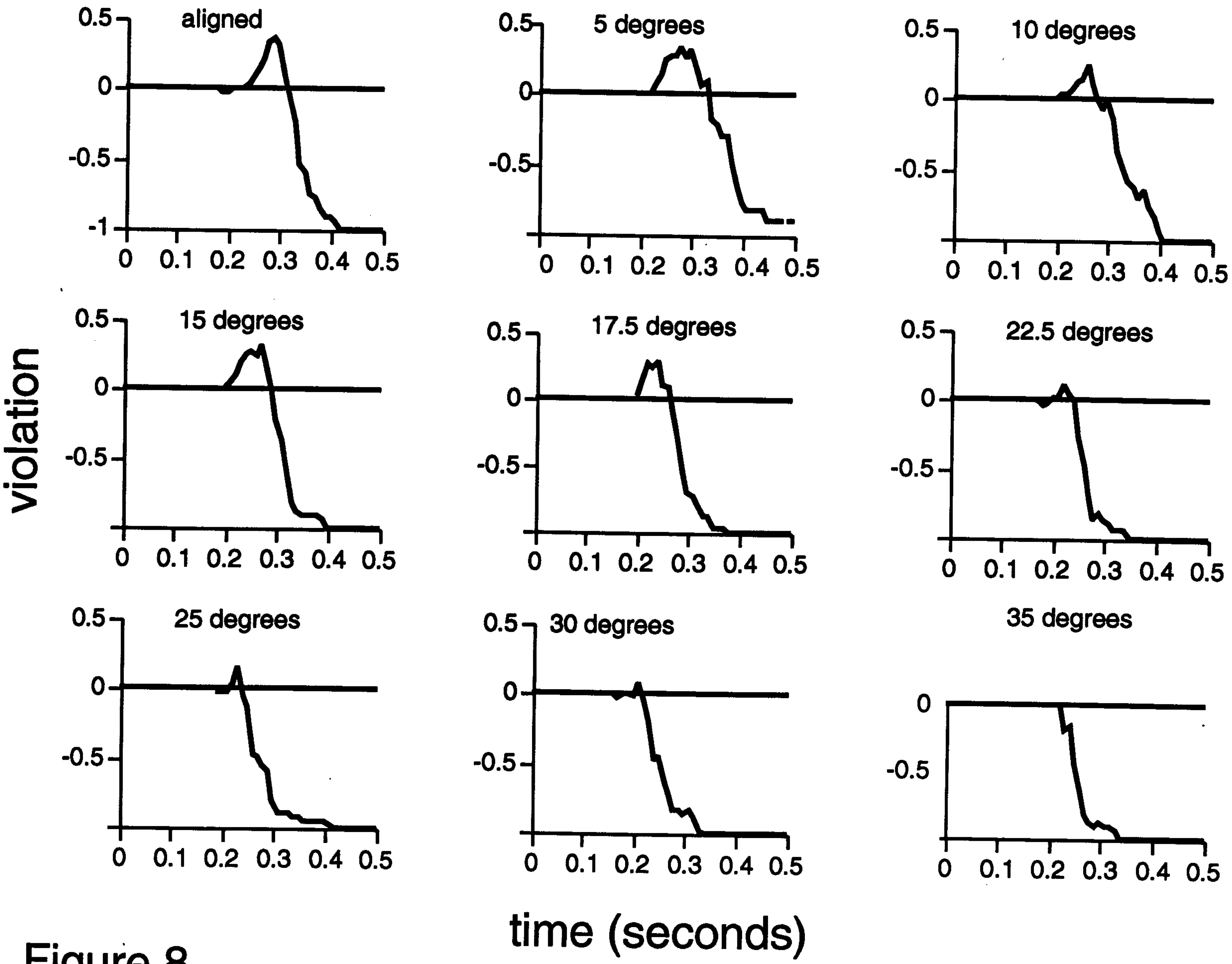
violation
aligned
5 degrees
10 degrees
15 degrees
17.5 degrees
22.5 degrees
25 degrees
30 degrees
35 degrees
0.5
0
-0.5
-1
0 0.1 0.2 0.3 0.4 0.5
time (seconds)

Figure 8

probability plot peaks at a value around .4 and the violation has a breadth of almost a tenth of a second. At the larger disparities the peak barely rises above zero and the breadth is reduced to one or two hundredths of a second. Eventual the peak is lost and the plot quickly falls from zero to minus one. The rate of this drop suggested that the spatially disparate auditory non-target maybe inhibiting the reaction to the visual target. This type of inhibition of response at larger disparities has been shown in single cell recording from multisensory neurons. Therefore, an analysis of variance with subsequent Fishers Protected Least Significant Difference test was performed for each subject at all of the disparities where violations did not occur. Evidence for inhibition was found for subject one at 40^{o} of disparity, probability of .004.

Subjects one and three were also tested under a special condition where, for dual stimulus trials, the speaker was placed directly behind the central fixation LED. In this case the speaker provided no directional information. Both subjects demonstrated moderate to large violations under this special condition. This is illustrated for subject three in Figure 9.

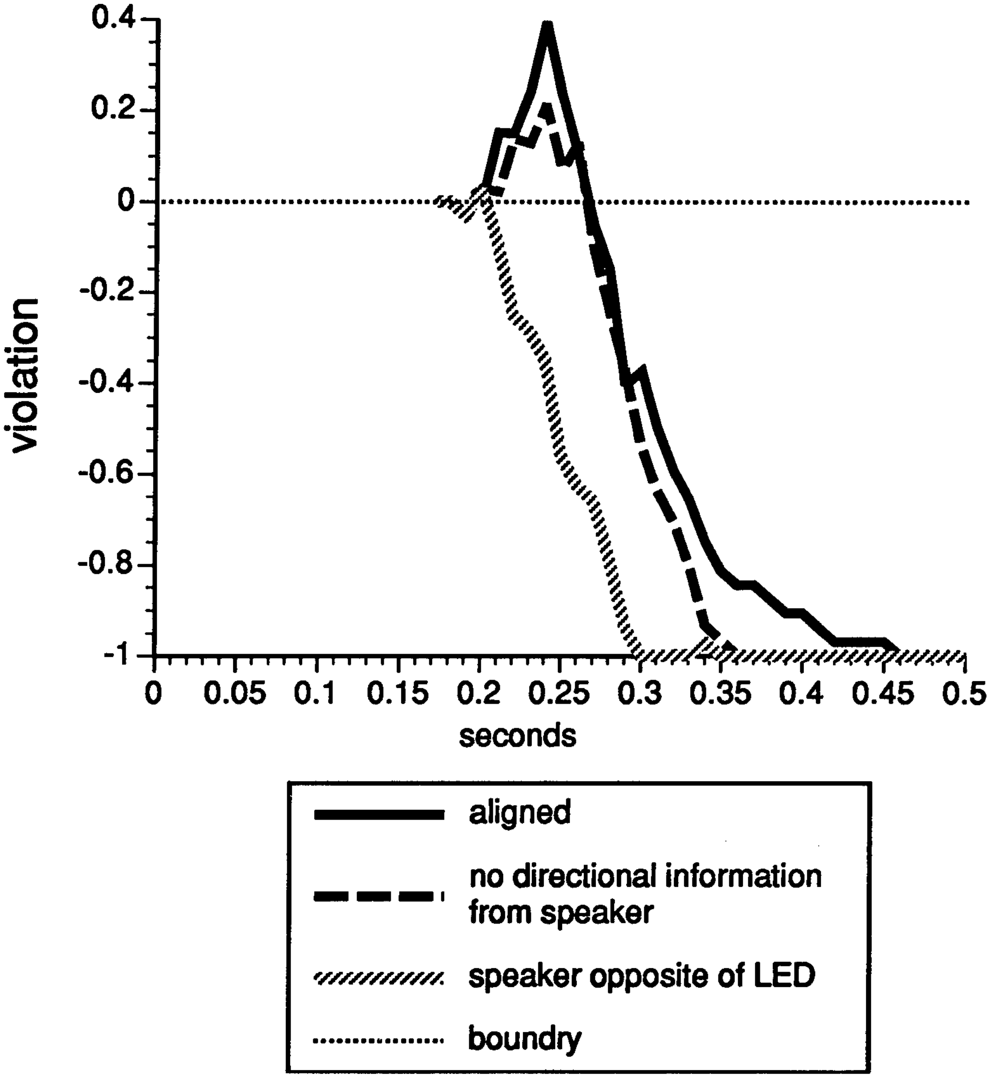

Figure 9: The special condition for dual stimuli presentations, subject 3. A large violation was obtained in the aligned stimulus condition, a moderate violation was seen in the special condition in which the speaker did not provide directional information and no violation was found when the speaker was located in the opposite hemifield.

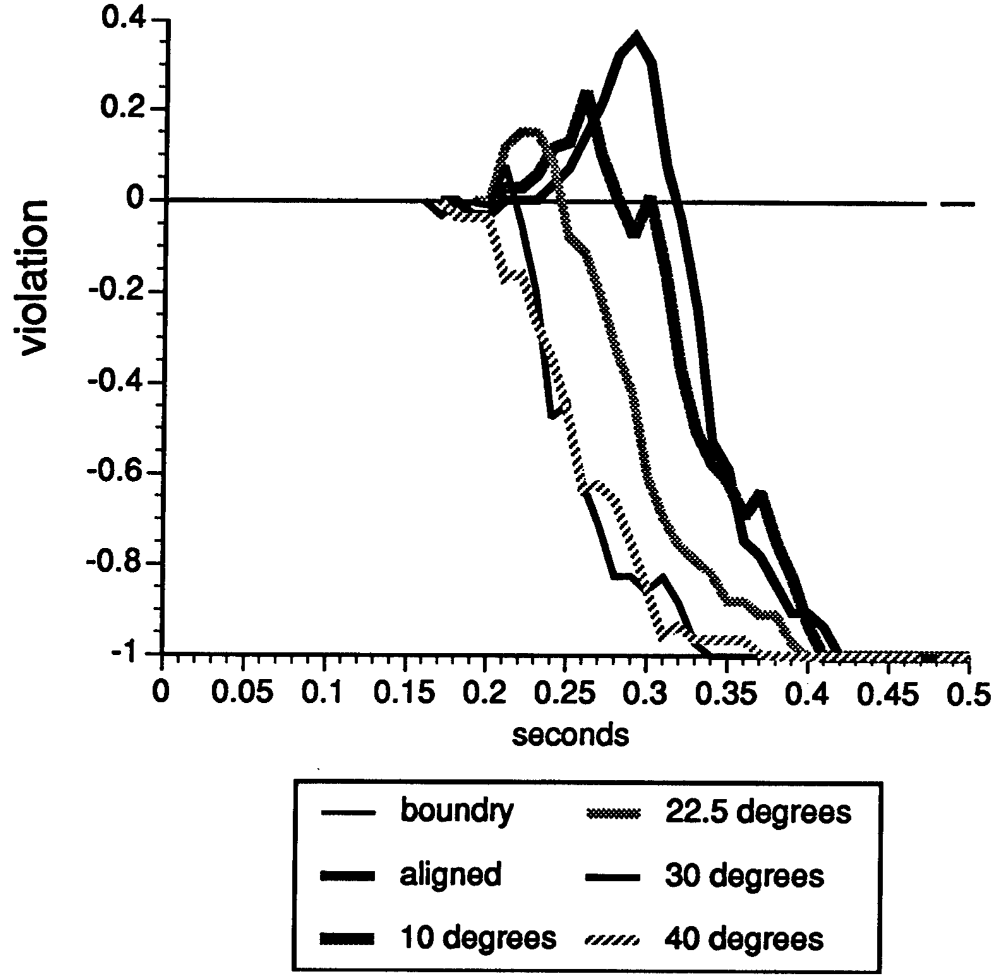

Figure 10: Comparison of difference curves at 5 disparities for subject two, note the shift of the peak violation to shorter reaction times with increasing disparity. This trend was not found for subjects one and three.

Figure 9 also shows that the peak of the violation for the aligned and the no directional information (20^o disparity) conditions occurred at very similar latencies. For both subjects one and three peak violations occurred between .22 and .24 seconds across disparities. However, in subject two with increasing disparity the peak of the violation moved to shorter latencies(figure 10). This pattern could be a consequence of how the violation boundary is constructed. Summation of two separate cumulative distributions to form the boundary forces violations to occur at short reaction times(Miller 1982). Alternatively the migration of the violation peak could be due to the progressive reduction of reaction times over the experiment.

Discussion

Reaction times to spatially aligned dual stimuli targets have been used as evidence of neural convergence, also known as neural summation, in the different sensory systems. Evidence of neural summation is found when there are violations of the race inequality or statistical facilitation boundary(Miller, 1986). In this study on saccadic reaction times to dual (visual and auditory) stimuli found evidence for neural summation over a large range of spatial disparities. In addition it was felt that there was a gradual loss of neural summation with increasing disparity. These findings are consistent with and may be a result of the multisensory convergence that is known to occur in the deep layers of the superior colliculus. My discussion will first address some experimental design issues and then will proceed with possible interpretations of the results.

Technical Issues

When disparity was introduced in the dual stimuli condition, the subjects had to be instructed about which stimulus was the saccade target. There were two possibilities: The subject could be instructed to make a saccade to whichever target they perceived first, or they could be told to make a saccade to the position of one of the targets(LED or speaker). The choice of response is significant because, for both the visual and auditory targets , saccadic reaction time varies with the eccentricity of the target.

The disadvantage in allowing the subject make a saccade to the first stimulus perceived is that it introduces two uncontrolled variables: saccade size and a decision regarding which saccade target is detected first. Permitting these variables to remain uncontrolled might confound the study and risked making the results uninterpretable. In addition allowing a wide range of saccade size creates measurement difficulties. Smaller saccades are difficult to differentiate from noise in EOG measurements, and the accuracy of a saccade cannot be judged if the target location is undefined.

However, instructing the subjects to look at the LED position also had a minor disadvantage in that the subjects knew the size, but not the direction, of the saccade even before the presentation of the target. This was not considered a problem for two reasons. First, the dual stimulus reaction times were used in comparison to unimodal reaction times measured under similar conditions. Second, the known saccade size did not result in a significant number of anticipatory saccades. The most reasonable instruction was to have the subjects make saccades to the 20° LED location. Next the subject instructions on how to respond in the single stimulus trials were addressed.

In my computations, the dual stimuli reaction times were compared to the LED and speaker reaction times. Therefore the manner in which these reaction times were obtained was of vital importance. Since the LEDs were positioned 20° left and right throughout the dual stimuli trials it followed that the LED-only trials should also be completed at the 20° locations. Determining the best subject response in the speaker-only trials was less straight forward.

The ideal speaker location for the speaker-only trials would meet three criteria:

1. Desired saccade size should equal the desired saccade size for the dual stimulus trials.
2. The subjects should make saccades to the location of the speakers.
3. The speakers should be in the same location as in the dual stimulus trials.

Since the dual stimuli speaker locations moved towards and then across the fixation point during the study it was impossible to meet all three criteria. I had three options:

1. The speakers could be placed in the same position as on the dual stimulus trials and subjects could make saccades to the position of the speakers. This would be contrary to criterion 1.
2. The speakers could be placed in the same position as on the dual stimulus trials and subjects could make saccades to the 20° position. This would be contrary to criterion 2.
3. The speakers could be placed at the 20° location and the subjects could make a saccade to the speakers. This would break criterion three.

Arguments could be made for using any of these options. However, I chose option three, to keep the speaker-only position at 20°, because it would provided the most conservative estimate of the violation. There is an inverse relationship between auditory saccade reaction time and target eccentricity for the eccentricities used in my study (Zambarbieri et al, 1982)(Yao and Peck, 1996). Therefore, keeping the speakers at the more distant 20° position, a violation of criterion three, decreased saccadic reaction times in the speaker-only condition and thus created a more stringent boundary for statistical facilitation.

The least conservative option would have been to require constant amplitude (20°) saccades while keeping the speakers in the same position as on the dual stimulus trials. Under these conditions, on speaker-only trials subjects would have been required to make saccades to a location where no target was presented. In addition, for disparities that crossed the midline the subjects would have been required to make saccades to the side opposite the speaker. Saccades made to positions opposite the stimulus are known as antisaccades and their latencies are significantly longer than target directed saccades (Guitton et al, 1985)(Forbes and Klein, 1996). Requiring antisaccades in some of the speaker-only trials would have increased the estimate of the extent of the violation for disparities crossing midline.

Interpretation

The objective of my study, to describe the effect of spatial disparity on the neural summation of auditory and visual information dictated that I use an intrasubject experimental design which was repeated for a total of three subjects. Successful interpretation of the data collected was dependent on the ability to compare and constrast the findings for the three subjects. The three subjects were similar in that their mean reaction times to the auditory targets were faster than to the visual targets at the respective stimulus intensities. All three subjects had large violations of the statistical facilitation boundary at no disparity. They all maintained this violation at disparities that crossed the midline but then lost it before a level of 40° disparity was reached. None of the subjects produced any recognizable express saccades.

The subjects also had some differences. Subject three maintained evidence of neural summation past a disparity of 35° while subject one made it to 25° and subject two made it to 30°. Subject two demonstrated a large and prolonged practice effect that was not as prevalent for subjects one and three. Subject two was also unique in that the peak violation shifted to shorter reaction times as the disparity increased. Subject one had the greatest dispersion in reaction times for all types of stimulus presentation.

Intersubject variability can complicate the interpretation of the results. For instance if the interpretation was based solely on the results of subject one (Figure 7) a possible conclusion would be that there is a sudden decrease in the violation between 4° and 10° of disparity and another sudden loss between 25° and 30° of disparity. However, when looking at the results from all three subjects (and taking in consideration the variability of subject one's reaction times) a much stronger argument can be made for the gradual loss of the violation with increasing disparity.

There are many possible explanations for the violations of the statistical facilitation boundary found in this experiment. These explanations include: practice effects, general alerting, facilitation of the directional decision process, an interaction between stimulus condition and the modality or shape of the reaction time distribution, and neural summation at multisensory neurons.

Practice effects are a perpetual problem with reaction time experiments. It is not unusual for reaction times to continue to improve over several hundred trials (Woodworth and Schlosberg, 1954). I attempted to minimize practice effects by running 300 practice trials with each subject before beginning these experiments. However, my data clearly show that reaction times continued to improve throughout my experiment (see Figure 6). This improvement in performance with practice should not impact the results of this study because the computations of latency facilitation were derived from comparisons of reaction times <u>within</u> a given session. In addition, intrasession order effects were counterbalanced.

While practice effects were straight forward to address in my experimental design, an explanation based on general alerting cannot be discarded so easily. In studies of alerting, a temporally relevant but spatially irrelevant non-target is presented with the target stimulus and results in a reduced latency. In a study on the effects of stimulus characteristics on saccadic reaction time, Engelken and Stevens (1989) demonstrated that reaction times to visual targets could be reduced by a synchronous overhead auditory signal. This auditory signal was considered to be spatially irrelevant because it provided no directional information about the saccade target, and because multisensory convergence should be minimal with the auditory target outside of the visual field. I am unable to estimate the extent to which general alerting contributed to my findings. The experimental design did not control for alerting and my results do not preclude an alerting component. However, it is clear that general alerting effects are insufficient to explain the pattern of reduction in reaction times across disparities. Alerting should reduce reaction times uniformly over all disparities, while my results showed a pattern of decreasing violations with increasing spatial disparities.

Another possible explanation of the reduced reaction time is that the dual stimuli facilitated the directional (left or right) decision process. The reaction times in this study are disjunctive or choice reactions not simple reactions. Disjunctive reactions have been shown to have latencies 20 to 200 ms longer than simple reactions with the amount of increase being related to the difficulty of the decision. For example the more alike the stimuli the longer the disjunctive reaction time (Woodworth and Schlosberg 1954). In my study the varying disparities may have impacted the similarity of the targets. However, if having dual stimuli did aid in the decision process this advantage would be lost as the speakers moved closer to center.

Other authors have found evidence that the dual stimuli may affect the directional decision process. Hughes et al (1994) looked at the effect of response modality on the magnitude of race inequality violations in reaction times to visual and auditory targets. In comparing saccadic, manual directed (joy stick), and undirected simple reaction times, they found that directed responses, responses that required a directional decision, showed greater violations than simple reaction times. This finding suggests that the single and dual stimulus conditions affect the simple and directed responses differently. This stimulus condition by response modality interaction implies that the dual stimuli condition facilitates the decision process required in directed responses. It is important to note that in the study of Hughes et al (1994) violations also occurred in the undirected response trials, so that the decision facilitation was insufficient to explain the entire reduction in reaction time.

Some of my results are difficult to explain on the basis of decision facilitation. When the speaker was placed at center fixation violations were moderate to large. I also failed to find an abrupt loss of violation at the midline as would be expected if decision facilitation played a major role in my results. The argument could be made that disparities that crossed the midline still provided directional data with the subjects learning to make saccades away from the speakers. This is unlikely in that the reaction times for anti-saccades (those made in the direction opposite to the target) are significantly slower than target directed saccades (Forbes and Klein, 1996). The fact that moderate to large violations were found with a central (non directional) speaker argues that a directional decision facilitation did not contribute strongly to my results, but I cannot rule out the possibility of a subtle effect.

Another process which might partially explain my data is an interaction between stimulus condition and the modality or shape of the reaction time distribution. Some authors have reported a bimodal or even a trimodal distribution for saccades. Saccades that occur around the shortest latency mode, approximately 100 ms, are often referred to as express saccades. The remaining saccades, belonging to the longer latency peaks, are referred to as regular saccades if the distribution is bimodal, or fast regular and slow regular if the distribution is trimodal. The multimodal nature of saccade latencies in some subjects has led to speculation that there are multiple neural pathways that can generate saccades and that the varying amount of processing required for each pathway determines the resulting reaction times (Fischer and Weber, 1993).

Express saccades are typically demonstrated with a gap (dark interval) between the offset of the fixation target and the onset of the saccade target. With no gap or with overlap between the fixation target and the saccade target the reaction time distributions exhibit fewer short latency saccades. Because I did not use the gap paradigm in my study, I did not expect many express saccades nor, in fact, did I record many express saccades. However the relationship between express saccades and the gap paradigm prompts the question: If a small change in the fixation target can influence the shape of the saccade distribution, could small changes in the saccade stimulus influence it as well? Assuming that there are multiple neural pathways for generating saccades, it is possible that the selection of the pathway in part depends on the nature of the saccade stimulus. A potential explanation for my findings is that the change from the single to the dual stimulus conditions resulted in a shift in saccades from a slower to a faster neural pathway. Histograms of data from individual recording sessions did not contain a sufficient number of trials to analyze the form of the distributions (unimodal vs. bimodal or trimodal). Combining reaction times from different sessions was not possible due to practice effects.

I started this experiment with the expectation of finding a behavioral correlate to the overlapping auditory and visual receptive fields found in multisensory neurons of the superior colliculus. We know that neural summation (multisensory convergence onto a common neurological pathway) of auditory and visual information occurs at these neurons. We also know the intermediate and deep layers of the superior colliculus plays an important role in the generation of saccades. Single cell recordings from these multisensory neurons have shown that the auditory and visual receptive fields are large and overlapping (Meredith and Stein, 1996). If spatial disparity is introduced so that one of the stimuli falls outside of its receptive field the neurons response rate often drops below what it would normally be for the remaining stimulus. The presence of a stimulus outside of the neuron's receptive field can interfere with or inhibit the neuron's response to a second stimulus inside the receptive field. These known characteristics of superior colliculus seem to be sufficient to explain the gradual loss of neural summation I found with increasing disparity. For example, when two multisensory stimuli are aligned there are many superior colliculus neurons whose receptive fields contained both targets. These neurons, which are known to demonstrate an enhanced response in the dual stimuli condition, then contribute to the generation of the saccade. However as increasing disparity is introduced fewer of these neurons have receptive fields that contain both targets. The enhanced response is gradually lost and saccadic reaction times increase.

Of course there is a significant gap in this hypothesis. How could a large number of multisensory neurons influence the saccadic reaction time? Hanes and Schall (1996) may have provided some insight on this issue. They compared mathematical models of decision and response preparation to physiological data from saccadic reaction times and neuron recordings in the frontal eye fields. Their results suggested that the neurons that control saccades have a set threshold firing rate for saccade initiation and that the speed at which that threshold is met determines the latency of the saccade. Therefore, if the saccade control neuron's dendritic tree synapsed with all the multisensory neurons whose receptive fields include the location of the desired saccade, the inputs from the multisensory neurons could be integrated at the control neuron. With this integration increasing the number of responding multisensory neurons would increase the firing rate of the control neuron. The control neuron's threshold would be reached sooner resulting in an expedited response.

The finding that neural summation is gradually lost with increasing disparity is supported by a recent study by Frens et al (1995). Frens and colleagues used a different experimental design and different stimulus configuration to look at the effects of spatial disparity on dual stimuli reaction times. They found that when their visual saccade targets and their auditory non-targets were aligned reaction times where facilitated. This facilitation was gradually lost as spatial disparity was introduced between the visual target and the auditory non-target.

The finding that neural summation occurs over a wide range of disparities could be considered to conflict with the findings of Stein et al (1989). They found that a auditory non-target decreased the likelihood that a cat would approach a low intensity visual target for a food reward, when the auditory target was presented at 15^{o} of disparity. This level of disparity was well within the range of neural summation found in my study. However differences in species and required responses makes it difficult to compare these studies directly.

Each of the possible interpretations of the reduced reaction times described so far is compatible with, or may result from, neural summation. Other unidentified explanations for the reduced reaction times maybe accounted for in the race model. If the proposition that violations of the statistical facilitation boundary can only be explained by neural summation is accepted I have demonstrated that neural summation occurs over a wide range of disparities. My hypothesis that this neural summation occurs at the superior colliculus is supported by previous anatomical and physiological findings, is consistent with results of this behavioral study, and is worthy of additional exploration.

REFERENCES

Duncan J (1980) The Locus of Interference in the Perception of Simultaneous Stimuli. Psychological Review 87: 272-300

Engelken E, Stevens K(1989) Saccadic Eye Movements in Response to Visual, Auditory, and Bisensory Stimuli. Aviation, Space, and Environmental Medicine. August 762-768

Fischer B, Weber H(1993) Express Saccades in Visual Attention. Behavioral and Brain Sciences 16:553-610

Frens M, Van Opstal J, Van der Willigen R(1995) Spatial and Temporal Factors Determine Auditory-Visual Interactions in Human Saccadic Eye Movements. Perception and Psychophysics 57:802-816

Forbes K, Klein R(1996) The Magnitude of the Fixation Offset Effect with Endogenously and Exogenously Controlled Saccades. Journal of Cognitive Neuroscience 8: 344-352

Guitton D, Buchtel H, Douglas R(1985) Frontal Lobe Lesions in Man Cause Difficulties In Suppressing Reflexive Glances and in Generating Goal-Directed Saccades. Experimental Brain Research 58:455-472.

Hanes P, Schall J(1996) Neural Control of Voluntary Movement Initiation. Science 274:427-429

Hershenson M(1962) Reaction Time as a Measure of Intersensory Facilitation. Journal of Experimental Psychology. 63:289-293

Hughes H, Reuter-Lorenz P, Nozawa G, Fendrich R(1994)Visual-Auditory Interactions in Sensorimotor Processing: Saccades Versus Manual Responses. Journal of Experimental Psychology: Human Perception and Performance 20:131-153

Knudsen E, Brainard M(1995) Creating a Unified Representation of Visual and Auditory Space in the Brain. Annu. Rev. Neurosci. 18:19-43

Kornblum S(1973) Simple Reaction Time as a Race Between Signal Detection and Time Estimation: a Paradigm and Model. Perception and Psychophysics. 13:108-112

Meijers L, Eijkman E(1977) Distributions of Simple RT With Single and Double Stimuli. Perception and Psychophysics 22,41-48

Merideth M, Stein B(1996) Spatial Determinants of Multisensory Integration in Cat Superior Colliculus Neurons. Journal of Neurophysiology 75: 1843-1857

Miller J(1986) Time Course of Coactivation in Bimodal Divided Attention. Perception and Psychophysics 40:331-343

Miller J (1982) Divided Attention: Evidence for Coactivation with Redundant Signals. Cognitive Psychology 14, 247-279

Raab D(1962)Statistical Facilitation of Simple Reaction Times. Transactions of the New York Academy of Sciences 24:574-590

Stein B, Meredith M, Huneycutt W, McDade L(1989) Behavioral Indices of Multisensory Integration: Orientation to Visual Cues is Affected by Auditory Stimuli. Journal of Cognitive Neuroscience 1:12-24

Woodworth S, Schlosberg H (1954) Experimental Psychology. Holt, Rinehart and Winston. New York

Yao L, Peck C(1996) Saccadic Eye Movements to Visual and Auditory Targets. Experimental Brain Research: In Press

Zambarbieri D, Schmid R, Magenes G, Prablanc C(1982) Saccadic Responses Evoked by Presentation of Visual and Auditory Targets. Experimental Brain Research 47:417-427