Adapting to supernormal auditory localization cues. I. Bias and resolution

Barbara G. Shinn-Cunningham,[a] Nathaniel I. Durlach, and Richard M. Held
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 23 September 1996; revised 27 January 1998; accepted 2 March 1998)

Head-related transfer functions (HRTFs) were used to create spatialized stimuli for presentation through earphones. Subjects performed forced-choice, identification tests during which allowed response directions were indicated visually. In each experimental session, subjects were first presented with auditory stimuli in which the stimulus HRTFs corresponded to the allowed response directions. The correspondence between the HRTFs used to generate the stimuli and the directions was then changed so that response directions no longer corresponded to the HRTFs in the natural way. Feedback was used to train subjects as to which spatial cues corresponded to which of the allowed responses. Finally, the normal correspondence between direction and HRTFs was reinstated. This basic experimental paradigm was used to explore the effects of the type of feedback provided, the complexity of the simulated acoustic scene, the number of allowed response positions, and the magnitude of the HRTF transformation subjects had to learn. Data showed that (1) although subjects may not adapt completely to a new relationship between physical stimuli and direction, response bias decreases substantially with training, and (2) the ability to resolve different HRTFs depends both on the stimuli presented and on the state of adaptation of the subject. [S0001-4966(98)03306-2]

PACS numbers: 43.71.Hw, 43.72.Ew, 43.66.Qp [RHD]

INTRODUCTION

Many recent studies have investigated the utility of sophisticated auditory display techniques for presenting information to human operators. In particular, the maturation of auditory spatial display technologies over the last decade has enabled spatial auditory cues to be presented to operators involved in a variety of everyday tasks (e.g., see Durlach, 1991; Smith, 1991; McKinley and Ericson, 1992; Begault, 1993a, 1993b; Ericson, 1993; Begault and Pittman, 1994; Begault, 1995; Begault *et al.*, 1995; Shinn-Cunningham and Kulkarni, 1996). Most of these studies have examined the utility of presenting auditory spatial cues, rendered as realistically as is practical, compared to presenting the same information without such spatial cues.

A previous paper (Durlach *et al.*, 1993) pointed out that the normal human auditory system has relatively poor spatial resolution and that it should be possible to improve performance by synthesizing ''supernormal'' localization cues that are not constrained by the laws of physics which determine normal localization cues. More specifically, it should be possible to design localization cues that span a larger range of just-noticeable differences (jnd's) than do normal cues, and thereby allow listeners to improve their ability to resolve nearby spatial positions. A number of approaches for creating supernormal auditory localization cues were discussed, including simulating localization cues from a larger-than-normal head, remapping the relationship between normal localization cues (i.e., normal HRTFs) and the corresponding spatial directions to create regions of supernormal spatial acuity, increasing interaural differences [for instance, by raising the difference spectrum to a power greater than one, e.g., see Durlach and Pang (1986) and Van Veen and Jenison (1991)], and emphasizing cues that are thought to relate to the perception of source distance.

While such approaches should clearly improve subjects' ability to resolve sources in jnd-type experiments, the effects of such cue manipulation on resolution in other types of tasks (such as identification tasks in which a larger range of physical stimuli are presented) or on response bias are not well understood. In general, the apparent location of a source simulated with supernormal localization cues will be displaced from its desired location when a naive listener is first exposed to such cues. The degree to which such errors can be overcome with training is a measure of how well subjects ''adapt'' to the given sensorimotor spatial discrepancy.

Previous studies of sensorimotor adaptation have focused on how mean response affected by discrepancies between localization cues from one modality versus other modalities. However, no previous studies have examined the effect of adaptation on resolution, nor taken into account whether changes in mean response are significant relative to the variability in subject responses. In addition, most sensorimotor adaptation work is qualitative rather than quantitative, and thus is of limited use in trying to develop a quantitative description or model of adaptation. General background on adaptation can be found in Welch (1978) and Welch and Warren (1986).

The main goal of the current experiments was to determine the extent to which adaptation to supernormal localization cues is achievable. In addition, experimental conditions were designed to examine how a variety of experimental factors (which have previously been shown to affect different aspects of spatial perception) might affect how subjects adapt to such altered cues.

[a]Current address: Boston University, Department of Cognitive and Neural Systems, Room 311, 677 Beacon St., Boston, MA 02215. Electronic mail: shinn@cns.bu.edu

While previous studies provide few quantitative measures of adaptation, many such studies suggest that one of the most important factors affecting adaptation is the type of exposure to the rearranged spatial cues that the subjects receive. In particular, active motor tasks generally yield more complete adaptation than comparable experiments with passive exposure to the rearrangement (Freedman and Zacks, 1964; Pick and Hay, 1965). Two different types of feedback are investigated in the current study in order to determine how active motor involvement affects adaptation in our experimental paradigm.

Other studies of spatial perception have shown that the complexity of the visual or acoustic field can affect the perception of source motion. For example, if a single point light source is seen to move around a subject in an otherwise dark room, the subject perceives himself to be stationary and the source to be moving. However, if multiple lights move with a fixed angular velocity around a stationary subject, the subject perceives himself to be rotating within a fixed room (Lackner, personal communication). In our studies, results when subjects are presented with an acoustic field (two ongoing, stationary sources in addition to the target) are compared to results from an experiment in which only the target source is presented.

The degree to which different stimuli can be resolved is determined in part by the range of target stimuli presented in an experiment (Durlach and Braida, 1969). The effect of stimulus range on resolution is examined by comparing results in adaptation experiments using a stimulus range of 120 degrees compared to a range of 60 degrees.

Finally, the strength of the cue rearrangement is systematically varied in order to examine how the rate and degree of adaptation depend upon the quantitative strength of the change in acoustic localization cues and the overall range of cues presented.

In the current study, subjects were asked to adapt to supernormal remappings of auditory localization cues. In the experiments described, subjects are first tested with ''normal'' localization cues to yield baseline measures, then with the ''supernormal'' cues to examine how performance changes as subjects adapt. Finally, at the end of each experimental session, subjects are retested with the ''normal'' cues to look for any aftereffects in performance that may result from the training with the supernormal cues. Two quantitative measures were used to track how subject performance changed over the course of the experimental session. Bias (a measure of response error in units of standard deviation in subject response) was used to measure the degree to which subjects adapted to the supernormal cues. While bias is related to the error in mean response, the measures are not equivalent. In particular, since bias is measured in units of standard deviation, the absolute magnitude of response errors cannot be determined from bias results. As a metric, bias describes the magnitude of response errors relative to the variability in subject responses; thus, a decrease in bias could result either from a decrease in absolute error or an increase in variability. In the current study, bias is examined instead of mean response error in order to quantify the importance of errors relative to response variability. In particular, if response variability is large, the relative importance of a given magnitude error is much less than if response variability is small. The ability to resolve adjacent stimulus positions was measured in order to gain insight into whether better-than-normal resolution is achievable in a localization task using supernormal cues. Estimates of the standard psychophysical metric d' (again, a measure with units in standard deviation) were found for adjacent stimulus positions to summarize resolution.

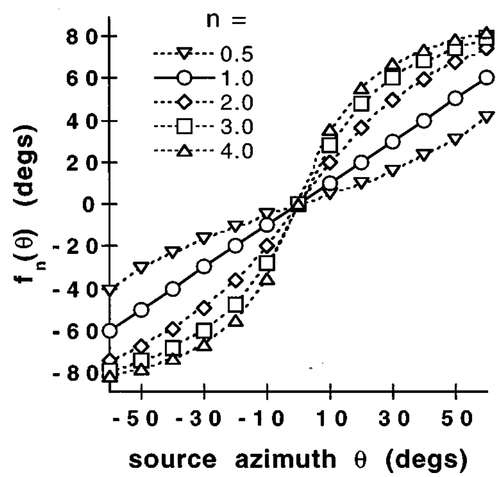

FIG. 1. Plot of the family of functions $f_n(\theta)$ used to transform auditory localization cues. With this transformation, a source from azimuth θ was synthesized using the HRTF that normally corresponded to the position $f_n(\theta)$.

I. SUPERNORMAL CUES

Supernormal localization cues were created in this study by remapping the relationship between source position and normal head-related transfer functions, or HRTFs. Normally, to simulate a source at azimuth θ and elevation ϕ, one simply uses the empirically measured HRTF for that direction, denoted in the frequency domain by $H(\omega,\theta,\phi)$, where ω denotes angular frequency. In the current study, the correspondence between HRTFs and azimuth values was remapped such that the HRTF used to simulate a source at position $[\theta,\phi]$ was given by

$$H'(\omega,\theta,\phi)=H(\omega,f_n(\theta),\phi). \tag{1}$$

In the current study, the family of mapping functions $f_n(\theta)$ used to transform azimuth cues is given by

$$f_n(\theta)=\frac{1}{2}\tan^{-1}\left[\frac{2n\sin(2\theta)}{1-n^2+(1+n^2)\cos(2\theta)}\right], \tag{2}$$

where the correct response azimuth is given by θ.[1] In other words, in the altered-cue situation, the HRTF associated with $[\theta,\phi]$ was equal to the HRTF normally associated with $[f_n(\theta),\phi]$. The parameter n in Eq. (2) corresponds to the slope of the transformation at $\theta=0$. With this function (shown in Fig. 1 for different values of n), the HRTFs used to generate stimuli are displaced laterally relative to the HRTFs normally used to present normal spatial sounds from those locations. The differences in localization cues for two positions in the frontal region are larger than normal with this remapping, while two positions off to the side would give rise to more similar cues than are normally heard.[2] As a result, subjects were expected to show better-than-normal resolution in the front and reduced resolution on the side, creating an enhanced ''acoustic fovea'' toward the front in which supernormal auditory localization could occur. In addition to affecting resolution, however, this transformation was expected to cause a bias whereby sources were perceived farther off-center than were their ''correct'' locations. If the same family of transformations is used with $n<1$, the opposite would be true: sources heard in front of the listener would have smaller-than-normal differences in localization cues and sources off to the sides of the listener would have larger-than-normal differences. The experiments discussed in the current paper used values of $n=1$ (normal cues, or no transformation), 2, 3, and 4.[3] The main questions of the study concern the extent to which (1) bias could be eliminated by subjects over time such that subjects interpreted the new relationship between HRTF and spatial location accurately, and (2) resolution was enhanced in the ''acoustic fovea'' with the transformed cues.[4]

Generating supernormal cues by remapping which HRTF corresponds to which direction has the advantage that subjects should hear a compact source image for every possible source position, because all localization cues are consistent with a normal source from some position [the position $f_n(\theta)$]. In other words, ignoring intersubject differences, all localization cues at all frequencies are consistent with a normal sound source at $f_n(\theta)$; subjects do not have to learn to interpret unusual combinations of interaural time, interaural level, and spectral cues.

Ideally, individualized HRTFs would be used to guarantee a ''realistic'' and compact sound image; however, in the current study, a single, standard set of nonindividualized HRTFs was used for all subjects. In general, the use of nonindividualized HRTFs may lead to mislocalizations in elevation, front/back confusions, and nonexternalized source images, especially when head tracking is not employed (e.g., see Wenzel *et al.*, 1993). In the current experiments, the response locations were always restricted to be in the horizontal front hemifield and front/back and up/down confusions cannot occur. In addition, the current experiments did not depend upon the subjective externalization of sound sources, the location dimension of interest (azimuth) does not rely heavily on the portions of the HRTF that show greatest intersubject variability, and the gross manipulations of HRTFs that were being studied should have a large effect on perception of source azimuth compared to any relatively minor effects caused by individual differences in HRTFs. Similarly, since the main effect, that of radically altering the azimuthal position of a source, should not show any bias due to any consistent overall spectral cues, the HRTFs used to generate stimuli were not equalized for the headphones used in the study. As a result, subjectively, sources were not always externalized; however, the ability to adapt to the required mapping of source cues to source location was not adversely affected by these subjective impressions. Finally, it should be noted that the focus of the current study is the ability of subjects to extract and use location information, not on subjective impressions of realism or source externalization.

The described approach to generating supernormal localization cues has the disadvantage that better-than-normal resolution is not possible for all locations around a listener; for positions where $df_n\theta/d\theta<1$ (sources off to the side with $n>1$), the transformation actually results in subnormal localization cues. One approach to generating supernormal localization cues for all positions around a listener would be to use HRTFs from a larger-than-normal head [a practical method for approximating such HRTFs is discussed in Rabinowitz *et al.* (1993)]. However, even this approach has its drawbacks. In particular, with such an approach, localization cues are not consistent with any normal-cue source position. As a result, source images may be broad, diffuse, and difficult to localize. This problem may be particularly true for sources off to the side since the interaural differences caused by such sources would be outside the range of interaural differences normally experienced. In contrast, although supernormal localization cues are only present for a limited range of directions in the current experiments, the sounds should be spatially compact and easy to localize.

II. GENERAL METHODS

A. Subjects

Subjects were recruited through the student employment agency at MIT. They were 18–28 years in age and either MIT students or related family members. All subjects reported normal hearing and were able to perform the localization tests employed in the study without difficulty. Different subjects were used in each of the seven experiments performed. The number of subjects completing a given experiment was between 3 and 8 (see Table III).

B. Stimuli

In all experiments, the target to be localized consisted of a periodic train of clicks (with a repetition rate of 10 clicks/s) generated by a Krohn-Hite model 5300A function generator. A 500-ms-long rectangular envelope gated the click train off and on so that roughly 5 clicks were heard per stimulus in each localization trial (described in Sec. II E below). In training runs (also described below), the click train was heard continuously until the subject completed his response. In some experiments, background sources were heard in addition to the click-train target. These background sources came from commercially recorded audio tapes, and consisted of a book on tape (Auel, 1980) and classical music (Handel, 1985) played from a Sony TCW490 tape deck.

C. Equipment

An auditory virtual environment (VE) was used to ''spatialize'' the acoustic stimuli in our experiments. The auditory VE consisted of a Convolvotron processor using HRTFs from subject SDO [measured and reported by Wightman and Kistler (1989)], a commercial, electromagnetic head tracker, a controlling PC, and headphones. In any given experiment, either the Isotrak (from Polhemus) or the Bird (from Ascension Technologies) head-tracking system was employed (for

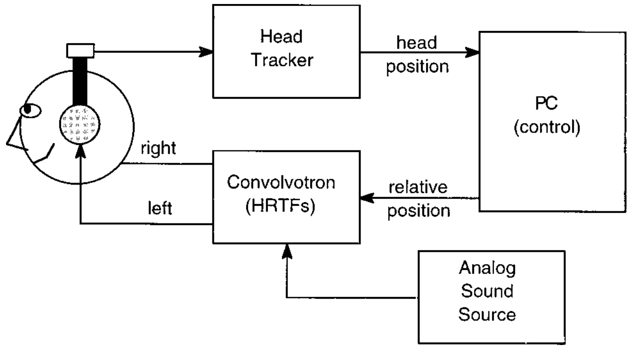

FIG. 2. Block diagram of the auditory virtual environment used to simulate acoustic sources.

all practical purposes, the two systems are interchangeable in their performance characteristics). A 486-based PC controlled the Convolvotron and the head tracker. The Convolvotron took monaural input stimuli (which were amplified, antialiased signals from the sources described above) and created appropriate binaural signals to simulate the stimuli from azimuthal locations specified by the controlling PC. Figure 2 shows a block diagram of the auditory virtual environment.

In most experiments, the binaural signals generated by the Convolvotron were played out through Etymotic Research 3A insert earphones with Bilsom earmuffs worn over the earphones. The combination of insert earphones and commercial hearing protectors helped to block background sounds from the subject during experiments. Taken together, the insert earphones and Bilsom earmuffs reduced sounds reaching the subjects via direct paths by roughly 40–50 dB across all audible frequencies. Experiment T_1 was performed before the insert earphones were incorporated into the experimental setup, using TDH-30 circumaural earphones. In this experiment, background sounds were not attenuated well. However, as this caused little difference in the observed pattern of results, this difference is thought to be of little consequence.

A visual display, consisting of 13 lights on a 5-ft-diam arc (every 10 degrees from -60 to $+60$ degrees azimuth), was located in front of the subjects at eye level throughout the experiments. The lights, labeled 1–13 from left to right, corresponded to the possible locations of the click-train target presented in the experiments (see Fig. 3). This visual display was used to present visual, spatial feedback about the simulated auditory sources used in the experiments.

D. Test procedure

Each subject performed eight identical test sessions over a period of between two to six weeks. Each session lasted roughly 2 h, and consisted of multiple runs separated by two 5-min breaks. Two types of experimental runs were used: localization test runs and training runs, described in detail below. In the training experiments (T_1 and T_3), no feedback was given during localization test runs, but training runs were interspersed with the localization test runs. In the feedback experiments (F_3, F_{3mid}, F_2, F_{4a}, and F_{4b}), correct-

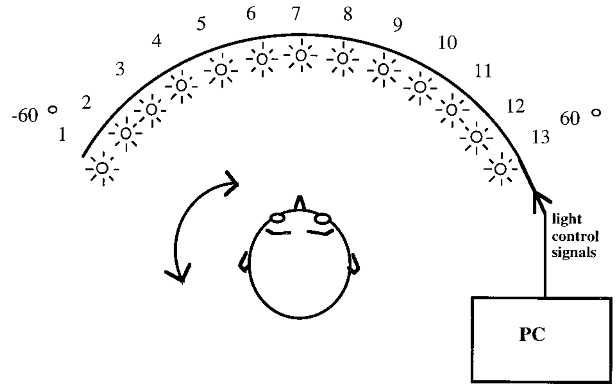

FIG. 3. Diagram of the visual display. Thirteen lights were positioned in front of the listener, spaced at 10-degree intervals ranging from -60 to $+60$ degree azimuth.

answer feedback was given during localization test runs and no training runs were performed (i.e., the training consisted of giving feedback during the test runs).

1. Test runs

In each localization test run, subjects were presented with a target stimulus simulated as coming from 1 of the 13 possible locations marked by the visual display, chosen at random. Each of the possible locations was presented exactly twice in each run. For most experiments, all 13 positions were employed for a total of 26 trials per run; however, in experiment F_{3mid}, only the middle 7 positions (positions 4–10) were used for a total of 14 trials per run. During the presentation of the target stimulus, subjects had to remain facing straight ahead (within 3 degrees of 0 degrees azimuth). If the head tracker reported that the subject turned his head off-center during a trial, that trial was thrown out and an additional trial was added to the run. In the test runs, the subject entered a number (1–13) corresponding to his/her best guess as to the location of the stimulus by typing on a laptop computer keyboard following presentation of the target. In experiments using correct-answer feedback (experiments F_3, F_{3mid}, F_2, F_{4a}, and F_{4b}), the light at the correct location was lit for 500 ms after the subject responded. In these experiments, the correct-answer feedback was the only information subjects received about the supernormal transformation employed (since no training runs were performed by the subjects). In the remaining experiments, no feedback was given during a test run (information about the transformation was obtained from the training runs, described below). A new trial began 500 ms after the subject entered his response to the previous trial.

2. Training runs

In experiments in which no feedback was given during test runs (T_1 and T_3), training runs were interspersed with test runs. During training runs, both synthetic auditory and real visual light sources were simultaneously turned on from 1 of the 13 possible locations, chosen at random. Subjects were instructed to turn their heads to face each audiovisual target. Once they faced the target (turned their head to within 1 degree of the target location), the light/sound source was

TABLE I. Order of runs in experiments T_1 and T_3. Each run consisted either of training or testing (shown in column one). The type of localization cues presented is given in columns two (for normal cues) and three (for supernormal or altered cues).

Run type	Norm	Super
Test 1	**x**	
Train	*x*	
Test 2	**x**	
	break	
Test 3		**x**
Train		*x*
Test 4		**x**
Train		*x*
Test 5		**x**
Train		*x*
Test 6		**x**
	break	
Test 7		**x**
Train		*x*
Test 8	**x**	
Train	*x*	
Test 9	**x**	
Train	*x*	
Test 10	**x**	

TABLE II. Order of runs in Experiments F_3, F_{3mid}, F_2, F_{4a}, and F_{4b}. Each test run used either normal cues or supernormal cues as shown in columns two and three.

Test runs	Norm	Super
1–2	*x*	
3–10		*x*
	break	
11–32		*x*
	break	
33–40	*x*	

turned off, a 500-ms pause occurred, and a new random location turned on. Training runs lasted 10 min each, with a variable number of trials (usually between 30 and 60 and determined by the speed with which subjects performed each trial) performed in each run. In training runs, exposure to the supernormal transformation entailed an active sensorimotor task (turning to face the audiovisual target). When this training method was employed, subjects never received feedback during the testing runs (and thus received no explicit feedback regarding any errors made during the testing portion of the experimental session).

3. Run order

In each session, auditory sources were first synthesized using ''normal'' HRTFs, then synthesized using the ''supernormal'' HRTF mapping, then synthesized with the normal HRTFs again. In the training experiments, a total of ten localization test runs were performed in each session: two normal-cue runs, five transformed-cue runs, and then three normal-cue runs. The training runs were performed in between test runs (see Table I). In the feedback experiments, 40 localization test runs were performed, one after another. In these experiments, 2 normal-cue tests were performed, followed by 30 supernormal tests, and then 8 normal-cue tests (see Table II). Two 5-min breaks were scheduled during each session for both types of experiments, as shown in the tables.

4. Subject instructions

Subjects were informed before the start of the experiment that they would be hearing both ''normal'' and ''transformed'' sound sources, and that the apparent location of transformed sources might not correspond to the ''correct'' location. They were instructed to always try to localize the sound sources correctly. Prior to the experiment, subjects were given a list of the experimental runs, including information about when they would hear normal sounds and when they would hear transformed sounds. In addition, prior to any change of cues (from normal to supernormal or from supernormal back to normal), subjects were reminded that the sources were about to change, and that they should answer as accurately as they could for the current sources. Beyond being told that sources were ''transformed,'' subjects were given no information about how the apparent source location might differ from the correct answer.

E. Experimental conditions

As mentioned above, the various experiments undertaken in this study were designed to probe some parameters that might affect how quickly and completely subjects adapted to remapped localization cues. Table III summarizes the important differences between the experiments. The effect of the complexity of the simulated sound field is shown

TABLE III. Summary of experiments performed. The altered-cue transformation ''strength'' [defined in Eq. (2)] is given in the second column. The number of subjects who completed eight sessions is shown in column 3. The ''Exp type'' describes whether subjects were exposed to training runs or given correct-answer feedback in order to cause adaptation. The number of source positions used in the experiment is given in column 5, and the number of acoustic sources simulated in the experiment (target plus additional background sources) is given in column 6. The head tracker used in the experiments is shown in the final column.

Exp	*n*	Subs	Exp type	Pos	Sources	Tracker
T_1	3	4	training	13	1	Isotrak
T_3	3	8	training	13	3	Isotrak
F_3	3	5	feedback	13	3	Bird
F_{3mid}	3	4	feedback	7	3	Bird
F_2	2	4	feedback	13	1	Bird
F_{4a}	4	3	feedback	13	1	Bird
F_{4b}	4, 0.5	3	feedback	13	1	Bird

by comparing results from experiments T_1 and T_3. If a complex sound field allows subjects to extract more information about the cue transformation than does a sound field consisting of a single source, more complete adaptation might be found in experiment T_3 than in experiment T_1. Comparisons between results from experiments T_3 and F_3 contrast the effects of active sensorimotor training (experiment T_3) versus correct-answer, cognitive feedback (experiment F_3). Experiments F_3 and F_{3mid} address the question of how the number of stimuli presented affects adaptation and resolution. Different strength transformations were employed in order to gather data that could lead to the development of a quantitative model of the adaptation process (comparison of results from experiments F_3, F_2, F_{4a}, and F_{4b}). Finally, experiment F_{4b} was identical to experiment F_{4a} except that subjects were exposed to a transformation of strength $n=0.5$ after exposure to the supernormal transformation. This final experiment investigated whether exposure to an inverse transformation might allow subjects to readapt to normal localization cues more quickly than without explicit inverse training.

III. RESULTS AND DISCUSSION

Although many experiments on adaptation to transformed sensorimotor cues have shown that exposure on one day can affect performance on a subsequent day (e.g., see Welch and Warren, 1980), no such effects are seen in the current experiments. However, this may be due to the fact that there are too little data to show any significant effects. In any case, any differences from session to session were small relative to the differences within a session. Thus, all the data reported here were combined across the eight identical sessions performed by each subject. This resulted in 16 trials (2 trials from each of 8 sessions) for each position and run, for each subject.

A. Analysis

We were interested in estimating how large subject response errors were relative to the variability in subject responses and how well subjects could distinguish stimuli from each other. Two metrics were used to summarize these quantities: bias and resolution.

A maximum likelihood method was used to estimate bias and resolution from the confusion matrices (pattern of responses observed for every possible physical stimulus) for each subject and run (combining data across the eight experimental sessions). This approach assumed that each presentation of a physical stimulus gives rise to a random variable whose value falls along a unidimensional, internal decision axis (see Fig. 4). The mean of the random variable depends upon which physical stimulus is presented. On each trial, the subject decides how to respond based upon the value of the unidimensional decision variable; in particular, it is assumed that the decision axis is divided into N contiguous segments by $N-1$ thresholds (or criteria) in order to decide which of N possible stimuli was presented on that trial (for an N-alternative, forced-choice experiment). It was further assumed that the decision variable has a Gaussian distribution for all physical stimuli, and that the standard deviation of the

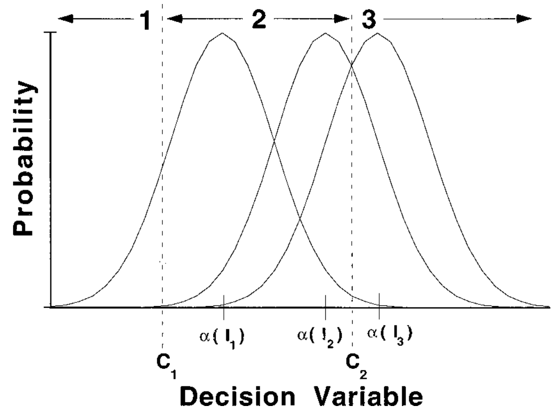

FIG. 4. Diagram of the assumed underlying decision space for a sample three-alternative, forced-choice experiment. The abscissa corresponds to the internal, unidimensional decision variable (which is assumed to relate monotonically to the azimuth of the physical location cues presented), and the ordinate shows the probability of hearing a given value of the decision variable. Shown are three Gaussian distributions with means $\alpha(I_1)$, $\alpha(I_2)$, and $\alpha(I_3)$, which result from the presentation of the three corresponding physical stimuli, I_1, I_2, and I_3. Note that it is assumed that the three distributions have equal standard deviation. The internal decision axis is broken into three contiguous regions by the placement of two criteria (labeled C_1 and C_2). On a given trial, the subject is assumed to respond that he heard stimulus i if the value of the internal decision variable falls into the ith contiguous region (see arrows at top of figure). As shown, C_2 is placed optimally, halfway between the means of the distributions for stimuli I_2 and I_3, resulting in a zero bias for C_2. Here C_1 is displaced from its optimal location, resulting in a nonzero bias for C_1. This placement of C_1 will cause the mean response to stimulus I_1 to be larger than 1, the ''correct'' answer. Resolution between stimuli I_1 and I_2 will be better than between stimuli I_2 and I_3, since the distance between $\alpha(I_1)$ and $\alpha(I_2)$ is larger than the distance between (I_2), and $\alpha(I_3)$.

Gaussian decision variable was independent of the physical stimulus presented in a run.[5] Finally, the mean of the Gaussian distribution was assumed to be monotonically related to the ''correct'' response of the physical stimulus presented on a given trial (specifically, the mean was assumed to vary monotonically with the azimuth of the HRTF used in a given trial). These assumptions are consistent with standard decision-theory models of psychophysical tasks. Since the hypothesized decision axis has arbitrary units, it can further be assumed without loss of generality that the standard deviation of the distribution is equal to one. Bias and resolution are given in units of standard deviation.

With these assumptions, the bias for each response depends on the placement of the $N-1$ criteria that divide the decision axis into N regions. If all stimuli are equally likely to be presented, subjects will maximize the probability of answering correctly by placing the nth criterion exactly halfway between the means of the distributions for the Nth and the $(N+1)$th stimuli (see Fig. 4).[6] If the means of the distributions corresponding to the different stimuli are equally spaced along the internal decision axis, then this optimal placement of the criteria will lead to mean responses which are roughly equal to the correct response.[7] Bias can then be defined as the difference between the optimal criteria placement and the actual criteria placement, measured in units of standard deviation. Errors in mean response arise when criteria are displaced from their optimal locations (i.e., when there is nonzero bias).

The ability to resolve adjacent positions depends only on the distance between the means of the corresponding distributions, measured in units of standard deviation in the distributions (the standard d' measure often used in psychophysical experiments). When the means of two physical stimuli are relatively close, subjects will be less able to resolve the stimuli from each other, independent of where criteria are placed. Similarly, when the means of two distributions are relatively far apart, subjects will rarely confuse the two stimuli (see Fig. 4).

In order to determine bias and resolution using this underlying model, the relative locations of the $N-1$ criteria and the means of the N decision-variable distributions (corresponding to the N physical stimuli) were estimated. These values were found using a gradient-descent algorithm which maximized the likelihood of obtaining the actual confusion matrix observed in a given run. Estimates were made for each subject and run (averaging all data across the eight experimental sessions). Bias was estimated by subtracting the estimated location of the Nth criterion from the average of the estimated means of the distributions corresponding to the Nth and $(N+1)$th stimuli (with the standard deviation in the underlying distributions assumed equal to one). Resolution was estimated by subtracting the estimated mean of the $(N+1)$th distribution from the estimated mean of the Nth distribution. This approach has previously been used to estimate resolution in intensity experiments (Lippmann *et al.*, 1976).

All subjects showed similar patterns of results for both bias and resolution. Thus, the results plotted below were found by averaging bias and resolution across all subjects in an experiment. While the basic patterns of results were identical across subjects, there were differences in the magnitudes of the effects; in general, these differences were due mainly to differences in the magnitude of the standard deviation in responses (i.e., some subjects showed greater response variability and consequently smaller overall bias and resolution than did other subjects).

In addition to the above-described maximum-likelihood estimates, simple estimates of bias and resolution were computed. With this approach, bias was determined simply by subtracting the mean response from the correct response for each position and normalizing the result by the experimental standard deviation in the response for that position.[8] Resolution was determined by finding the difference between means for adjacent stimuli, and normalizing by the geometric mean of the experimental standard deviations for those two positions. In general, this simple method gave results which were consistent with the more complex, maximum-likelihood method. However, the maximum-likelihood approach partially compensates for inaccuracies inherent with this simpler method. For instance, in both cases, there will be an edge effect in estimating bias. In particular, estimated bias tends to be positive for the leftmost position and negative for the rightmost position because subjects could not respond that a source was left of position 1 or right of position 13, even if a stimulus sounded far off to one side. With the simple method for estimating bias, bias is *always* less than or equal to zero for position 1, since the mean response must be greater than or equal to one. With the maximum-likelihood

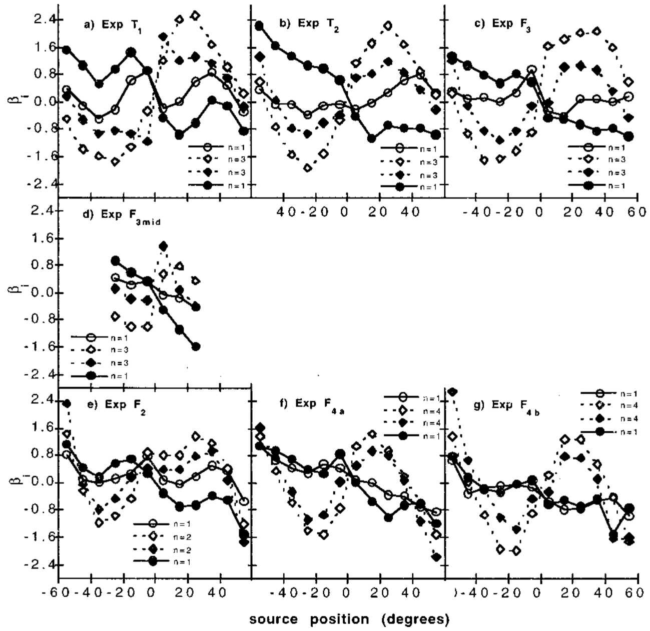

FIG. 5. Bias results for the seven experiments. In each panel, the average estimated bias is plotted as a function of correct source position (in degrees). For each panel, bias was estimated for each subject and run in the experiment and then averaged across all subjects in the experiment. Four runs are plotted in each panel: the first run in the experiment (open circles), the first run with the transformed cues (open diamonds), the final run with altered cues (filled diamonds), and the first normal-cue run following altered-cue exposure (filled circles). In each panel, the legend details the strength of the cue transformation employed in the normal- (n always equals 1) and altered-cue runs ($n=2$, 3, or 4).

estimation, bias will *tend* to be less than zero for position 1 for the same reason, but the approach also takes into account the variability in subject responses in order to estimate bias. For this reason, the maximum-likelihood estimates are better at dealing with edge effects and consequently are more accurate. Only the maximum-likelihood estimates are presented here.

B. Bias

The first test using normal cues was expected to show little systematic bias since the cues presented were consistent with everyday experience (at least within the limits of the simulation method). As such, they provided a baseline for performance. The first test with the transformed localization cues was expected to show a strong bias whereby source locations were heard farther off center than were their ''correct'' locations since the ''correct'' location was suddenly and arbitrarily changed. After training with the altered cues, bias was expected to decrease in magnitude for all positions. Finally, results from the first, posttraining, normal-cue test were expected to show either (1) a bias in the direction opposite that shown when the transformed cues were first presented (if changes in subject performance were unconscious, and therefore could not be immediately ''turned off'' by the listener following training) or (2) little bias (if subjects were capable of consciously interpreting cues as either normal or altered, as appropriate).

Figure 5 shows bias results as a function of source position for the seven experiments. Normal-cue runs are plotted with circles and solid lines; altered-cue runs with diamonds and dashed lines. The open symbols represent runs prior to altered-cue training while filled symbols correspond to results from the tests following exposure to the altered cues. In all experiments, four runs are plotted: the initial run (which uses normal cues); run 3, the first run with altered cues; the final run with altered cues (run 7 in the training experiments, run 32 in the feedback experiments); and the first run with normal cues following altered-cue exposure (run 8 in the training experiments, run 33 in experiments F_3, F_{3mid}, F_2, and F_{4a}, and run 37 in experiment F_{4b}).

Independent of the exact training method employed, the range of source positions presented, the number of sources simulated in the acoustic field, and the strength of the cue transformation employed, bias results were similar. First of all, in all experiments, bias results are roughly left–right symmetrical, as expected. Since there is no reason to expect asymmetrical results, the degree of left–right symmetry in the results is one measure of variability in estimates of bias. Results from the first normal-cue run (open circles) showed some bias, although these errors were much smaller than those found in other runs. A strong bias occurred in the first test with transformed cues (open diamonds) in the direction predicted by the transformation and the aforementioned edge effect (subjects heard sources farther off-center than they were). Without these edge effects, one would expect bias to increase monotonically with the magnitude of the source azimuth. Results for the test using transformed cues after training (filled diamonds) showed a clear reduction in bias over the whole range of positions tested; however, this adaptation was not complete. In all experiments, the final bias when testing with altered cues is roughly one-third to half of the initial bias (compare filled to open diamonds). Finally, a negative aftereffect is seen in the results from the final normal-cue test following exposure (filled circles), indicating that performance was not controlled solely by conscious correction which could be easily ''turned off'' at will.

In general, the four runs plotted in Fig. 5 summarize the important aspects of performance across all runs. In particular, these runs show the starting and ending points for performance as subjects adapt to the transformation in localization cues. Intermediate runs, which are not shown in Fig. 5, show intermediate levels of performance changing from the initial to the final results shown in Fig. 5. It is important to note that the changes in performance (evidenced by differences between the open and filled symbols in Fig. 5) actually occurred rapidly during the course of the experiments and that performance had clearly stabilized by the end of the altered-cue test period in all seven experiments (performance was close to the final values within approximately two altered-cue runs in the training experiments and four altered-cue runs in the feedback experiments). The exact time course of these changes will be fully explored in a subsequent paper. However, because performance was stable by the last altered-cue run in all experiments, direct comparisons of the final results from the various experiments can be made, even across the training and feedback experiments, for which completely different training paradigms were investigated.

Differences between experiments appear to depend mainly on differences in the transformation employed. First, examine results from experiments T_1, T_3, and F_3 [panels (a), (b), and (c) in Fig. 5]. Results for these three experiments are roughly the same, despite the differences in the exposure method (experiment T_1 had only a single source on at all times, while experiment T_3 presented two continuous background sources in addition to the target source in order to give a more complete auditory scene; experiments T_1 and T_3 used active sensorimotor training, while correct-answer feedback was used in experiment F_3.) In these three experiments, the largest biases occur for positions at -30 and $+30$ degrees. Similar results are seen in experiment F_{3mid} [Fig. 5(d)], except that the largest biases occur for sources near -20 and $+20$ degrees, slightly closer to the center position than occurred with experiments T_1, T_3, and F_3. This result is consistent with the fact that a smaller range of positions was used in experiment F_{3mid}, so that the edge effects were significant for sources closer to the 0 degrees azimuth. In experiment F_2 [Fig. 5(e)], a less extreme transformation was employed. In this experiment, the largest biases occurred for source positions nearer to -40 and $+40$ degrees. Once again, this result can be explained on the basis of the transformation employed. In this experiment, smaller biases are expected for a given source position than occurred in experiments T_1, T_3, and F_3 because the transformation was less extreme. In addition, the edge effects described above should affect fewer positions, since fewer sources will be heard outside of the allowed response range in this experiment. Finally, results from experiments F_{4a} and F_{4b} [Fig. 5(f) and (g)] show a larger bias than the other experiments for the centermost source positions. The largest biases occur for sources near -20 and $+20$ degrees. For these experiments, the more extreme cue transformation resulted in more sources falling outside the normal range of source positions, so that a larger number of sources were affected by the edge effects. Any differences between the results for experiments F_{4a} (in which normal cues were presented in runs 33–40) and F_{4b} (in which a transformation of 0.5 was used in runs 33–36 and normal cues were presented in runs 37–40) should be evident by comparing the final, normal-cue run results (filled circles). However, any such differences are small relative to the variability seen in the results.

C. Resolution

In general, resolution was expected to be somewhat better in the center region than at the edges of the range of source positions in the first run using normal cues, since resolution is best straight ahead of the listener. Independent of the exact pattern of results, the first normal-cue test provides a baseline measure against which results from later tests could be compared. Given the remapping function shown in Fig. 1, results from the first test with altered cues were expected to show improved resolution (relative to the first, normal-cue test) in the center of the range and decreased resolution at the sides. Following training with the altered cues, resolution was expected to either (1) remain as it had been for the first altered-cue test (if resolution depended solely on the difference between the magnitudes of

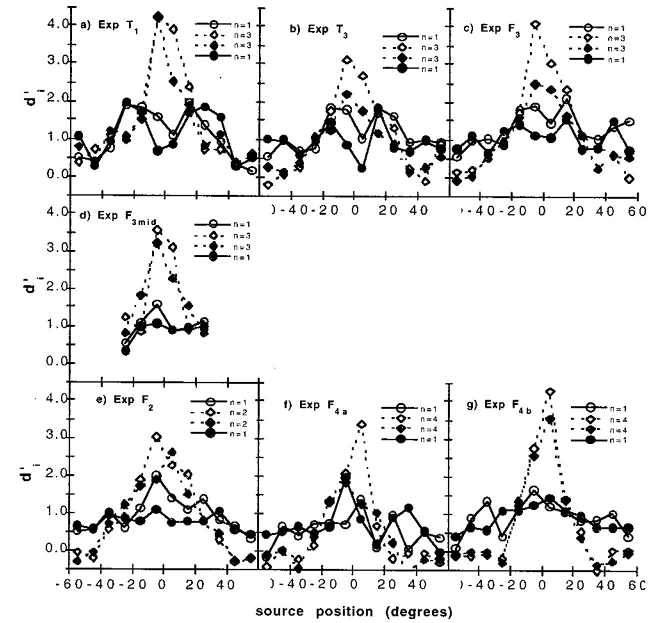

FIG. 6. Resolution results for the seven experiments. In each panel, the average estimated d' is plotted as a function of correct source position (in degrees). For each panel, resolution was estimated for each subject and run in the experiment and then averaged across all subjects in the experiment. Four runs are plotted in each panel: the first run in the experiment (open circles), the first run with the transformed cues (open diamonds), the final run with altered cues (filled diamonds), and the first normal-cue run following altered-cue exposure (filled circles). In each panel, the legend details the strength of the cue transformation employed in the normal- (n always equals 1) and altered-cue runs ($n=2$, 3, or 4).

the cues at adjacent positions) or (2) change with time (if performance depended upon high-level cognitive factors which were affected by training).

Resolution results for all seven experiments are shown in Fig. 6. The same four runs are plotted in this figure as were plotted in the bias results in Fig. 5. The initial test with normal cues is plotted with open circles and solid lines, the initial test with altered cues is plotted with open diamonds and dashed lines, the final test with altered cues is plotted with filled diamonds and dashed lines, and the first normal-cue test following altered-cue exposure is plotted with filled circles and solid lines.

As with bias, the basic patterns of results are the same across all experiments, and results are roughly left–right symmetric. Overall, resolution for normal cue runs showed a consistent pattern in which resolution appeared to be slightly *worse* for the center positions compared to positions just off center, rather than slightly better as was expected. This result may be due to positional dependencies on the accuracy of the simulation as well as true differences in resolution arising from perceptual issues. In any case, the results from the initial run using normal cues provides a baseline against which results from the other runs can be measured. As expected for the transformation employed, resolution on the first run using the transformed cues was enhanced for positions in the central region and degraded at the edges of the range compared to results from the initial run. Of particular interest are the results for the final, altered-cue test (filled diamonds): although resolution remains enhanced over that achieved with normal cues, there tends to be a decrease in resolution compared to results for the first test using the transformed cues, demonstrating that resolution does not depend solely on physical cues. A similar decrease can be seen when comparing the final normal-cue test (following training) with the initial, preexposure normal-cue test. There is a tendency for resolution to decrease with training, both for the normal-cue and for the altered-cue results.

Comparing results for experiments T_1, T_3, and F_3 [Fig. 6(a)–(c)], there is substantially more variability in the resolution results across these three experiments than was seen in the bias results. In particular, the increase in resolution with the first altered-cue run seen in experiment T_1 is larger than that seen in experiments T_3 and F_3. However, in all cases, the estimated value of d' is quite large. Generally speaking, when d' for two stimuli is larger than about 3.0, there are large changes in the amount of overlap of the distributions for relatively small changes in distance between the means of the distributions. As a result, estimates of d' are very sensitive to small changes in the pattern of responses when d' is relatively large. Thus, although the apparent differences between results for experiments T_1, T_3, and F_3 are pronounced, they arise in part from the numerical instability of estimating large values of d'. This numerical instability can also be seen in left–right asymmetries in many experiments for some of the large values of d' [e.g., examine the estimates of d' for the middle two positions in experiment T_1, particularly for the final test with altered cues (filled diamonds)]. Results from experiment F_{3mid} [Fig. 6(d)] are roughly consistent with results for experiments T_1, T_3, and F_3. The increase in resolution seen in experiment F_2 for the first test with altered cues is slightly smaller than was seen in the first four experiments, consistent with the fact that the transformation in experiment F_2 is less extreme than in the first four experiments. Finally, the increase in resolution in experiments F_{4a} and F_{4b} tends to be greater than in the other five experiments, all of which used a less extreme transformation. As with bias, there are no obvious differences in resolution results for experiments F_{4a} and F_{4b}, despite the fact that a transformation of 0.5 was used in runs 33–36 in experiment F_{4b}.

IV. CONCLUSIONS AND FUTURE WORK

The results demonstrate that subjects are able to learn remappings between acoustic cues and physical locations in the sense that they are able to reduce bias with training. However, subjects never completely overcome their systematic errors when responding to altered localization cues. Instead, over time, their errors grow smaller in magnitude, but retain the same pattern of results as is seen in their initial errors with altered cues (i.e., larger errors at the center of the range, smaller errors at the edges of the range). Since performance was stable by the final test with altered cues, it appears that subjects cannot adapt completely to the transformation employed in these experiments (shown in Fig. 1). This result is consistent with previous results investigating sensorimotor adaptation (Welch, 1986) which show that adaptation usually occurs, but is seldom complete; instead, systematic biases remain even after performance is stable (and additional exposure causes no further change in localization performance). A negative after-effect was found for all experiments, implying that changes in performance were not based solely on conscious correction; instead, changes occurred gradually, with training. These gradual changes occurred both when subjects were first exposed to the altered cues (adaptation), and when subjects were returned to normal cues at the end of the experiments (recovery).

Unlike previous experiments investigating sensorimotor adaptation, our experiments imply that there is no qualitative difference in the final level of adaptation achieved when using training paradigms that involve subjects in active sensorimotor tasks (experiments T_1 and T_3) compared to the adaptation achieved in experiments in which simple correct-answer feedback is provided (experiments F_3, F_{3mid}, F_2, F_{4a}, and F_{4b}). In fact, the relative insensitivity of bias and resolution results to the various experimental conditions is somewhat surprising. Bias and resolution appeared to be insensitive to the complexity of the auditory scene, since results from experiments T_1 and T_3 are comparable. Even when subjects are explicitly trained to an inverse transformation in an attempt to allow their normal-cue test results to return to preexposure patterns more rapidly, no clear effect is seen (compare results for experiments F_{4a} and F_{4b}).

While changing exposure conditions causes little difference in results, changing the strength of the transformation (compare results for experiments F_3, F_2, F_{4a}, and F_{4b}) and/or the range of stimuli used (compare results for experiments F_3 and F_{3mid}) did cause differences in bias and resolution. In particular, the stronger the transformation, the larger the initial errors in performance (bias) and the larger the initial increase in resolution for center positions. Similarly, the bias in the final run with the altered cues (filled diamonds in Fig. 5) varies with transformation strength, with larger final errors tending to occur for experiments using more extreme transformations. While the absolute size of initial bias errors depended on the transformation employed, it appears that the decrease in bias with training is roughly proportional to the initial error. For all experiments, final bias was roughly one-third to one-half of the bias initially measured in the first altered-cue run. It should be noted that a reduction in bias may arise either from decreases in mean error or increases in response variability. Subsequent analysis (described in Shinn-Cunningham *et al.*, 1998) that examines mean response in detail shows that localization errors decrease with exposure to supernormal localization cues, and that this is the main reason that bias decreases over time. However, there is a small increase in the estimated internal decision noise as subjects adapt which contributes to the decrease in response bias.

Increasing the physical differences in the stimuli that correspond to the 13 possible responses allows subjects to achieve better-than-normal resolution. However, as subjects adapt to these changes, their ability to resolve adjacent stimuli appears to decrease. Previous investigations of the dependence of resolution on experimental conditions have shown that resolution depends on both the physical cues used in an experiment, and on the range of physical cues presented during an experiment. For instance, earlier models of resolution explain that resolution of two stimuli in a jnd task is better than resolution of the same two stimuli during a many-alternative, forced-choice task by positing that internal decision noise grows as the range of the physical stimuli increases (Durlach and Braida, 1969; Braida and Durlach, 1972, 1985). However, the current results show that resolution also depends upon the past history of exposure of the subject. If resolution depended only on the range and values of the physical stimuli, then resolution would be identical in the first and last runs using altered cues. In these two runs, an identical set of 13 physical stimuli are presented; however, resolution decreases over time.

One possibility is that the decrease in resolution arises because data are averaged over time periods during which the decision criteria shift significantly. If this were the case, estimates of the variability in the underlying probability densities would be too large and estimates of d' too small simply due to estimation errors. Thus, an apparent decrease in resolution might be due to changes in criteria placement during the final test with altered cues. However, this explanation is inconsistent with other aspects of the data. In particular, analysis in a companion paper (Shinn-Cunningham *et al.*, 1998) shows that the greatest changes in mean response occur at the beginning of the altered-cue exposure period, and that performance is stable well before the final test run with altered cues. In other words, shifts in response criteria may cause estimates of d' for the initial test with altered cues to be too small but will have little effect on the estimates of d' for the final test with altered cues.

An alternative explanation assumes that resolution depends upon the range of stimuli being attended to by the subject (not the range of the physical stimuli used in a run). Prior to adaptation, subjects expect positions to span 120 degrees (from -60 to $+60$ degrees), but during training, they hear acoustic cues covering a much larger range (see Fig. 1) and learn to attend to this broader range. Perhaps, over time, as subjects adapt to the change in cues, they begin to attend to a larger and larger range of physical stimuli. Earlier models of resolution performance may be extended to include specifications for how performance changes over time in an adaptation experiment by specifying how the effective mean response and the effective range depend upon training. A preliminary model of adaptation based on the model of intensity perception by Durlach and Braida (Durlach and Braida, 1969; Braida and Durlach, 1972, 1985) is currently under development.

ACKNOWLEDGMENTS

This work was mainly supported by AFOSR Grant Nos. 90-200, 93NL387, and 96NL145. Supplementary support was received from Navy Grant No. N61339-93-C-0047. Portions of this work were presented at the 124th (November 1992) and 127th (June 1994) meetings of the Acoustical Society of America. We are grateful to Lorraine Delhorne and John Crouch for their help with data collection.

[1]This family of transformations was chosen because it changes smoothly as a function of angular position, and maps the positions -90, 0, and 90 degrees to themselves for all values of n.

[2]For the family of transformations $f_n(\theta)$, the point at which the slope

$df_n(\theta)/d\theta$ equals one is given by $|\theta| = \frac{1}{2}\cos^{-1}((n^2-2n+1)/(n^2-1))$. Thus, the range of positions that have larger-than-normal changes in physical cues for a given azimuthal increment are $(-35,+35)$ degrees when $n=2$; $(-30,+30)$ degrees when $n=3$; and $(-27,+27)$ degrees when $n=4$.

[3]In Experiment F_{4b}, subjects were exposed to a transformation of strength $n=0.5$ after the normal ''supernormal'' exposure period in order to test whether such retraining might increase the speed with which their performance returned to preexposure patterns.

[4]In the current experiments, only azimuthal position was transformed to achieve better-than-normal resolution. In the future, we hope to perform similar experiments involving elevation and distance. Since the main cue for azimuthal position is interaural time delay (ITD; see, for example Wightman and Kistler, 1992), it is likely that similar adaptation results would obtain for experiments involving only ITD transformations.

[5]Since the standard deviation of the internal decision variable is assumed independent of the physical stimulus, nonuniform sensitivity to the physical variable of interest (i.e., azimuth) implies a nonlinear dependence of the decision-variable distribution mean on the physical stimulus. In particular, the means of the decision variable distributions for sources at 0 and 10 degrees (sources that are relatively easy to distinguish from one another) will be farther apart than will the means of the decision variable distributions for sources at 50 and 60 degrees (sources that are harder to distinguish from one another). This is consistent with other decision-theory models, for instance, see Durlach and Braida (1969).

[6]It is possible that subjects change the placement of criteria as a run progresses to reflect the fact that there is less uncertainty in which stimuli will be presented at the end of the run than at the start of the run. For instance, if the position directly in front of the listener is not presented until the last two trials of the run, in principal it is possible that the listener is aware that the last two trials are more likely to come from the center than anywhere else, and may shift his criteria to reflect this fact. In our analysis, we ignore any such effects. In practice, there were enough trials in each run that subjects did not keep track of how often each of the stimuli had been presented during the run.

[7]While this is true in general, mean response for sources at the edges of the response range will not have this property. For instance, in Fig. 4, the mean response to stimulus 3 will always be less than 3, since the left tail of the distribution for stimulus 3 falls into the ''2'' response range and there is no ''4'' response allowed to counterbalance the effect of these ''2'' responses on the mean response to stimulus 4.

[8]Note that, using this simple estimation method, bias values are calculated for each of the N stimulus values presented, not for the $N-1$ criteria. Assuming the same underlying stimulus distributions and Thurstonian decision model, the ith bias estimated in this manner reflects ''error'' in the placement of all the criteria in the vicinity of the mean of the distribution of the ith stimulus value, while the bias estimated by the maximum likelihood estimate reflects only the ''error'' in the placement of the ith criteria.

Auel, J. (**1980**). *Clan of the Cave Bear* (Crown, New York).

Begault, D. (**1993a**). ''Call sign intelligibility improvement using a spatial auditory display,'' Technical Report 104014, NASA Ames Research Center.

Begault, D. (**1993b**). ''Head-up auditory displays for traffic collision avoidance system advisories: A preliminary investigation,'' Hum. Factors **35**, 707–717.

Begault, D. R. (**1995**). ''Virtual acoustic displays for teleconferencing: Intelligibility advantage for telephone grade audio,'' 95th Convention of the Audio Engineering Society, 25–28 February 1995, Paris, France.

Begault, D., and Pittman, M. T. (**1994**). ''3-D audio versus head down TCAS displays,'' Technical Report 177636, NASA Ames Research Center.

Begault, D. R., Wenzel, E. M., Miller, J., and Shrum, R. (**1995**). ''Preliminary investigation of spatial audio cues for use during aircraft taxi under low visibility conditions,'' NASA Ames Research Center.

Braida, L. D., and Durlach, N. I. (**1972**). ''Intensity perception. II. Resolution in one-interval paradigms,'' J. Acoust. Soc. Am. **51**, 483–502.

Braida, L. D., and Durlach, N. I. (**1985**). ''Peripheral and central factors in intensity perception,'' in *Auditory Function: Neurobiological Bases of Hearing*, edited by G. M. Edelman, W. E. Gall, and W. M. Cowan (Wiley, New York), pp. 559–583.

Durlach, N. I. (**1991**). ''Auditory localization in teleoperator and virtual environment systems: ideas, issues, and problems,'' Perception **20**, 543–554.

Durlach, N. I., and Braida, L. D. (**1969**). ''Intensity perception. I. Preliminary theory of intensity resolution,'' J. Acoust. Soc. Am. **46**, 372–383.

Durlach, N. I., and Pang, X. D. (**1986**). ''Interaural magnification,'' J. Acoust. Soc. Am. **80**, 1849–1850.

Durlach, N. I., Shinn-Cunningham, B. G., and Held, R. (**1993**). ''Supernormal auditory localization. I. General background,'' Presence **2**, 89–103.

Ericson, M. A. (**1993**). ''A comparison of maskers on spatially separated competing messages,'' J. Acoust. Soc. Am. **93**, 2317.

Freedman, S. J., and Zacks, J. L. (**1964**). ''Effects of active and passive movement upon auditory function during prolonged atypical stimulation,'' Perceptual and Motor Skills **18**, 361–366.

Handel, G. F. (**1985**). ''Concerti a due Corni,'' recorded by the English Chamber Orchestra, Philips.

Lippmann, R. P., Braida, L. D., and Durlach, N. I. (**1976**). ''Intensity perception. V. Effect of payoff matrix on absolute identification,'' J. Acoust. Soc. Am. **59**, 129–134.

McKinley, R. L., and Ericson, M. A. (**1992**). ''Minimum audible angles for synthesized localization cues presented over headphones,'' J. Acoust. Soc. Am. **92**, 2297.

Pick, H. L., and Hay, J. C. (**1965**). ''A passive test of the Held reafference hypothesis,'' Perceptual and Motor Skills **20**, 1070–1072.

Rabinowitz, W. R., Maxwell, J., Shao, Y., and Wei, M. (**1993**). ''Sound localization cues for a magnified head: Implications from sound diffraction about a rigid sphere,'' Presence **2**, 125–129.

Shinn-Cunningham, B. G., and Kulkarni, A. (**1996**). ''Recent Developments in Virtual Auditory Space,'' in *The Generation and Applications of Virtual Auditory Space*, edited by S. Carlile (Landes, Austin, TX), pp. 185–243.

Shinn-Cunningham, B. G., Durlach, N. I., and Held, R. M. (**1998**). ''Adapting to supernormal auditory localization cues. II. Constraints on adaptation of mean response,'' J. Acoust. Soc. Am. **103**, 3667–3676.

Smith, S. (**1991**). ''An auditory display for exploratory visualization of multidimensional data,'' in *Workstations for Experiment*, edited by G. Grinstein and J. Encarnacao (Springer-Verlag, Berlin).

Van Veen, B. D., and Jenison, R. L. (**1991**). ''Auditory space expansion via linear filtering,'' J. Acoust. Soc. Am. **90**, 231–240.

Welch, R. (**1978**). *Perceptual Modification: Adapting to Altered Sensory Environments* (Academic, New York).

Welch, R. (**1986**). ''Adaptation of Space Perception,'' in *Handbook of Perception and Human Performance*, edited by K. R. Boff, L. Kaufman, and J. P. Thomas (Wiley, New York), Vol. I, pp. 24.1–24.45.

Welch, R., and Warren, D. H. (**1986**). ''Intersensory interactions,'' in *Handbook of Perception and Human Performance*, edited by K. R. Boff, L. Kaufman, and J. P. Thomas (Wiley, New York), Vol. I, pp. 25.1–25.36.

Welch, R. B., and Warren, D. H. (**1980**). ''Immediate perceptual response to intersensory discrepancy,'' Psychol. Bull. **88**, 638–667.

Wenzel, E. M., Arruda, M., Kistler, D. J., and Wightman, F. L. (**1993**). ''Localization using nonindividualized head-related transfer functions,'' J. Acoust. Soc. Am. **94**, 111–123.

Wightman, F. L., and Kistler, D. J. (**1989**). ''Headphone simulation of free-field listening. I. Stimulus synthesis,'' J. Acoust. Soc. Am. **85**, 858–867.

Wightman, F. L., and Kistler, D. J. (**1992**). ''The dominant role of low-frequency interaural time differences in sound localization,'' J. Acoust. Soc. Am. **91**, 1648–1661.