Trading Directional Accuracy for Realism in a Virtual Auditory Display

Barbara G. Shinn-Cunningham, I-Fan Lin, and Tim Streeter

Hearing Research Center, Boston University
677 Beacon St., Boston, MA 02215 USA
{shinn, ifan1976, timstr}@cns.bu.edu

Abstract

In a virtual auditory display, reverberant energy is important for cuing source distance and creating sounds that are externalized (perceived as coming from outside the listener's head). The benefits of adding reverberation to a virtual auditory display are seen even for small levels of reverberant energy, like that in an ordinary, small room. While previous studies have shown that large amounts of reverberant energy can disrupt speech intelligibility and degrade localization accuracy, systematic deficits have not been observed in most studies conducted in or simulating typical small rooms, like classrooms. Moreover, while past studies have measured either distance or directional perception in virtual auditory displays, little is known about how these two dimensions interact when both vary. The dual goals of the current study were to see whether 1) increasing the task demands by having subjects simultaneously judge both distance and direction (compared to judging either distance or direction) interfered with their ability to judge each dimension, and 2) the value of one stimulus dimension affected how well listeners could judge the other dimension. Results show that mean judgments of distance and direction are independent of task demands. However, localization judgments are more variable if subjects must simultaneously judge distance and direction compared to when they are only asked to judge only one dimension. This "dual-task cost" is larger for distance judgments than for directional judgments. Mean perceived distance is independent of source direction; however, directional accuracy for lateral sources depends on the source distance: judgments of source laterality become increasingly more biased towards midline as source distance increases. These results show that directional abilities degrade with the inclusion of even modest amounts of reverberation (which are needed to encode source distance) and highlight the importance of understanding the goals of a particular application when designing a spatial auditory display. When creating a virtual auditory display, directional accuracy trades off directly with distance accuracy and realism.

1 Introduction

Many past studies of virtual auditory displays have considered how to represent either source direction or distance. The current study explored both source distance and direction to see whether the two dimensions interact in a virtual auditory display. In particular, we are interested in whether direction and distance perception are "independent," that is, whether perception of one dimension is unaffected when the other dimension is varied. This question is of great practical importance in designing auditory displays, because if the dimensions are independent, than each can be used to encode different information in the display without any interference. Distance and direction judgments could fail to be perceptually independent for at least two reasons: 1) task demands of simultaneously judging the two dimensions may cause competition for central resources such as attention and / or memory, degrading performance in the dual task, and 2) the stimulus attributes that encode source direction and source distance may interfere with one another at a perceptual level.

If central resources are not a limiting factor, then subjects should be able to report both distance and direction of a sound source with an accuracy that is predicted directly by performance when reporting either distance and direction alone. However, if central resources limit performance on the dual task, then performance in the dual task should be worse than predicted when only one stimulus dimension is judged at a time (e.g., see Durlach *et al.*, 1989).

If distance and direction cues in a stimulus do not interfere at a perceptual level, then the ability to judge one dimension should not depend on the value presented in the other dimension; for instance, directional abilities should be independent of the source distance that is simulated. Given that the main cues for source direction (interaural time differences or ITDs, interaural level differences or ILDs, and spectral cues) and the primary cue for source distance (the ratio of the direct sound energy compared to reverberant energy or D/R) are distinct, one might expect directional and distance judgments to be independent at a perceptual level. However, when sources are very close to the listener, ILD cues vary not only with source laterality, but also with source distance (Brungart & Rabinowitz, 1999; Shinn-Cunningham *et al.*, 2000). Although ILD cues are not the most important acoustic cue for either of the two dimensions measured here (source laterality, where ITD is generally dominant, and source distance, where D/R is generally dominant; Macpherson & Middlebrooks, 2002; Mershon & King, 1975; Wightman & Kistler, 1997; Zahorik, 2002b), the fact that ILD cues co-vary with direction and distance for nearby sources may make it harder to jointly estimate distance and direction, compared to estimating each dimension separately.

In addition to the fact that ILD cues co-vary with distance and direction, there is reason to expect simulated source distance to directly affect directional accuracy in a virtual auditory display. Reverberant energy causes ITDs and ILDs to vary over time and reduces the magnitude of ILDs (Shinn-Cunningham *et al.*, 2005). Reverberant energy also reduces the depth of spectral notches, making these cues (important for conveying sound source elevation) less reliable (Shinn-Cunningham et al., 2005). These effects increase with increasing sound source distance because direct sound energy (which contains the ITDs, ILDs, and spectral cues signalling source direction) decreases as distance increases. Thus, one might expect directional accuracy to decrease with increasing source distance (and decreasing D/R). However, past experiments investigating sound localization in rooms have found little effect of moderate amounts of reverberant energy on localization accuracy (Shinn-Cunningham, 2000).

In the current study, binaural room impulse responses from a classroom were used to create a virtual auditory display in which both distance and direction could be controlled. Localization performance was measured in 1) a dual task in which subjects had to simultaneously judge both distance and direction and task demands are great, and 2) simpler, single tasks in which subjects judged only distance with direction held fixed, or vice-versa.

2 Methods

2.1 Subjects

Four subjects (three female, one male) were paid for their participation in the experiments (one additional subject was excluded because his performance was comparable to chance performance, i.e., similar to what would happen if he had guessed his responses, independent of the stimulus presented). All subjects had normal hearing as confirmed by an audiometric screening.

2.2 Stimuli

Binaural room impulse responses (BRIRs) were measured using microphones positioned in the ear canals of a KEMAR acoustic research manikin (for details of the measurement approach and analysis of the impulse responses, see Shinn-Cunningham et al., 2005). KEMAR was located in the center of a classroom (T_{60} of approximately 650 ms) and BRIRs were recorded for a loudspeaker located at all combinations of 10 different directions (azimuths every 10˚ from 0˚ and 90˚ to the right of the listener, in the horizontal plane containing KEMAR's ears) and seven distances (from 20 to 228 cm, spaced on a logarithmic scale). BRIRs were convolved with a train of pink noise bursts (five 160-ms-long noise bursts separated by 20 ms silent gaps) to simulate noise sources at different locations relative to KEMAR. Signals were presented through Tucker-Davis Technologies hardware to EAR 3a insert earphones. All experiments were performed in a sound-treated booth that contained a computer terminal. The corresponding computer, located outside the booth, controlled stimulus presentation and recorded subject responses via a graphical user interface (GUI). The GUI showed a top-down diagram of the subject with the 70 different possible locations (at all valid locations relative to the listener) indicated by dots. Following each trial, subjects clicked on one of the dots on the screen using the computer mouse and the computer recorded their response. After each response, the correct response location on the GUI was illuminated to provide feedback.

2.3 Procedures

Runs were either fixed-azimuth, fixed-distance, or full-field. In the fixed-azimuth runs, the source azimuth was fixed at one of the 10 possible source directions throughout the run and each of the seven distances was presented twice, in random order (14 trials in each of 10 distinct runs, one per azimuth). In this single-task condition, subjects were asked only to judge source distance. In the single-task, fixed-distance runs, distance was held fixed throughout the run and the azimuth randomly selected on each trial (producing 7 different fixed-distance runs of 20 trials per run); subjects indicated the perceived direction. In a full-field run, all possible source locations (all combinations of seven distances and 10 directions) were presented in random order, exactly once per run, for 70 trials per run. In this dual task, subjects judged both distance and direction on each trial. Subjects performed multiple experimental sessions (consisting of multiple runs), with at most one session per day.

Each subject performed 10 training sessions to familiarize them with the task. Each training session consisted of a full set of 10 fixed-azimuth runs (one for each direction) and seven fixed-distance runs (one for each distance) for a total of 280 trials per training session. Inspection of these results showed some learning across the first two or three sessions, but for all subjects, learning had asymptoted well before the final training session.

Following training, each subject performed ten test sessions (with at most one session per day). Each test session consisted of 1) either a block of runs containing a full set of fixed-azimuth runs (10 runs of 14 trials) or a block of runs consisting of a full set of fixed-distance runs (7 runs of 20 trials), and 2) a 70-trial full-field run (which were only presented during formal testing, not during training). We refer to the former case (a session consisting of a set of fixed-azimuth runs and a full-field run) as a "fixed-azimuth session" and the latter as a "fixed-distance session." Thus, each formal test session consisted of exactly 210 trials, 140 trials in a "fixed block" and 70 trials in a "full-field block." The order of blocks within each test session was randomly set for each subject. All subjects alternated between performing fixed-azimuth sessions and fixed-distance sessions on subsequent days. Whether a subject performed a fixed-azimuth session or fixed-distance session on the first test day was determined randomly.

Each subject performed exactly five fixed-azimuth sessions and five fixed-distance sessions over the course of the ten test sessions. Thus, each subject performed 700 fixed-azimuth trials, 700 fixed-distance trials, and 700 full-field trials in the test sessions. Moreover, in all three cases, the 700 trials consisted of exactly the same distribution of stimuli (10 repetitions of all possible combinations of 10 azimuths and 7 distances); however, the stimuli were blocked differently in each case and task demands differed in the different conditions, as noted above.

3 Analysis

In each condition (fixed-azimuth, fixed-distance, and full-field), the 700 trials from each subject were used to generate appropriate confusion matrices describing the frequency with which subjects gave each possible response for each of the stimuli presented. In the fixed-azimuth condition, there were 10 different 7x7 confusion matrices for each subject, corresponding to the 10 fixed azimuths tested. Similarly, there were seven 10x10 confusion matrices for each subject in the fixed-distance condition. Finally, a 70x70 confusion matrix summarized how subjects responded to each distance-direction pair for the dual-task, full-field task.

3.1 Perceived Location

Results from the confusion matrices for each subject were analyzed to determine the mean and standard deviation in individual subject responses. For the fixed-azimuth condition, the mean and standard deviation in perceived distance were computed as a function of distance for each direction presented; in the fixed-distance condition, the mean and standard deviation in perceived azimuth was computed as a function of direction for each of the fixed distances; in the full-field condition, the mean and standard deviation were computed for both stimulus dimensions. Results were inspected for each subject, then combined across subjects to find across-subject means and standard deviations.

3.2 Information Transfer

In addition to considering response mean and standard deviation, information theory was applied to estimate the number of bits of information that the listener was able to extract in the various conditions. In order to examine the extent to which the two display dimensions interfered with one another, we wished to directly compare the information transfer in the fixed and full-field conditions. First, to compare the total amount of information that could be displayed using one or both dimensions, the information transfer was calculated for each of the raw confusion matrices. For instance, in the fixed-distance condition, we computed the number of bits a listener could extract from source azimuth from each of the seven 10x10 confusion matrices, one value for each of the seven fixed distances. In these computations, the total number of bits possible varied with condition (2.81 in the fixed-azimuth condition, 3.32 in the fixed-distance condition, and 6.13 for the full-field condition). Second, to compare the average loss of information associated with judging a source dimension in the dual task compared to the fixed tasks, we needed to collapse the results so that we had the same number of trials and same number of possible bits for both dual and single tasks. We combined all of the fixed-azimuth confusion matrices to produce a single 7x7 matrix (with 100 trials for each distance; the maximum number of bits transferred in this case was 2.81) and all of the fixed-distance confusion matrices to produce a single 10x10 matrix (with 70 trials for each azimuth; maximum number of bits 3.32). For comparison with the fixed-azimuth conditions, the full-field confusion matrix was collapsed across all azimuths to produce a single 7x7 confusion matrix; for comparison with fixed-distance results, the full-field confusion matrix was collapsed across distance to generate a single 10x10 confusion matrix.

4 Results

4.1 Perceived Location

4.1.1 Mean Responses

Figure 1 plots the across-subject average response distance as a function of simulated source distance (panels A and B) and average response direction as a function of simulated direction (panels C and D). For clarity, results are shown for subsets (the two extreme values and one intermediate value) of the other parameter (direction in panels A and B and distance in panels C and D). Error bars show the across-subject standard deviation of the individual-subject means. Panels A and C show results for the fixed conditions, while panels B and D show results for the corresponding full-field analysis.

In general, mean perceived location varied strongly with the simulated location, both in distance and direction (results generally fall near the diagonal). In addition, mean response in both distance and direction were similar in the fixed and the full-field conditions (i.e., results in panels A and B are similar, as are the results in panels C and D). There are a few cases where there appear to be systematic differences between the fixed and full-field results (e.g., the overestimation of source azimuth for sources with azimuths less than 50˚ in the full-field condition, shown as circles in panel D, which is not present in the fixed-distance results of panel C; the tendency to overestimate the distance of nearby sources in the full-field tests more than in the fixed-azimuth tests, comparing the diamonds and circles in panels B and A). However, closer inspection shows that these differences are small compared to the inter-subject variability. Moreover, and consistent with the large inter-subject variability at these points, inspection of the individual results shows that these trends are due to idiosyncratic responses of one subject and do not reflect any consistent difference between the mean responses in the fixed and full-field conditions.

Mean distance judgments did not vary with simulated source direction; within both panels A and B, results generally fall on top of one another, independent of azimuth. In contrast, there was a consistent effect of source distance on the mean azimuth responses for lateral sources (panels C and D) that was also present in the data for each individual subject (not shown). In both the fixed-distance (panel C) and full-field (panel D) results, mean perceived azimuth was consistently underestimated at the largest azimuth values, an effect that increased systematically with simulated source distance. In particular, the results for the nearest sources were most accurate, with an underestimation of less than 10˚ for a source at 90˚ (see rightmost circles in panels C and D). However, when the simulated source was at 228 cm, the mean perceived azimuth of a 90˚ source was underestimated by almost 30˚ (see rightmost triangles in panels C and D).

4.1.2 Variability in Responses

Figure 2 shows the across-subject averages of the standard deviation in responses as a function of distance (panel A) and direction (panel C) for subsets of both the fixed conditions (filled symbols and solid lines) and the full-field condition (open symbols and dashed lines). In order to more directly compare results in the dual and single tasks, the ratio of the standard deviation in the full-field condition over the standard deviation in the fixed conditions was first computed for each subject and then averaged across subject for each stimulus location. These results are shown in panels B and D (for the fixed-direction and fixed-distance conditions, respectively).

Results in panel A show that the standard deviation in the subject responses increases with increasing stimulus distance, growing roughly linearly with source distance when plotted as a function of log-distance. This result is consistent with past experiments investigating distance perception, which suggest that distance sensitivity is roughly constant on a log scale (Mershon & Bowers, 1979; Zahorik, 2002a). Comparing the full-field to fixed-azimuth results in panel A shows that distance judgments were generally more variable in the full-field condition than in the corresponding fixed-azimuth condition (open symbols fall above filled symbols). This result is confirmed by results in panel B, which show that the standard deviation in the full-field condition divided by the standard deviation in the fixed-azimuth conditions consistently is greater than one. This ratio decreases with increasing source distance (compare the ratio plotted in panel B for sources at distances of 20-30 cm, where values range from 0.8-4.0, to the ratio for sources at a distance of a meter and beyond, where values generally fall between 0.7 and 1.5), suggesting that the cost associated with judging both distance and direction compared to judging only distance decreases as the source distance increases.

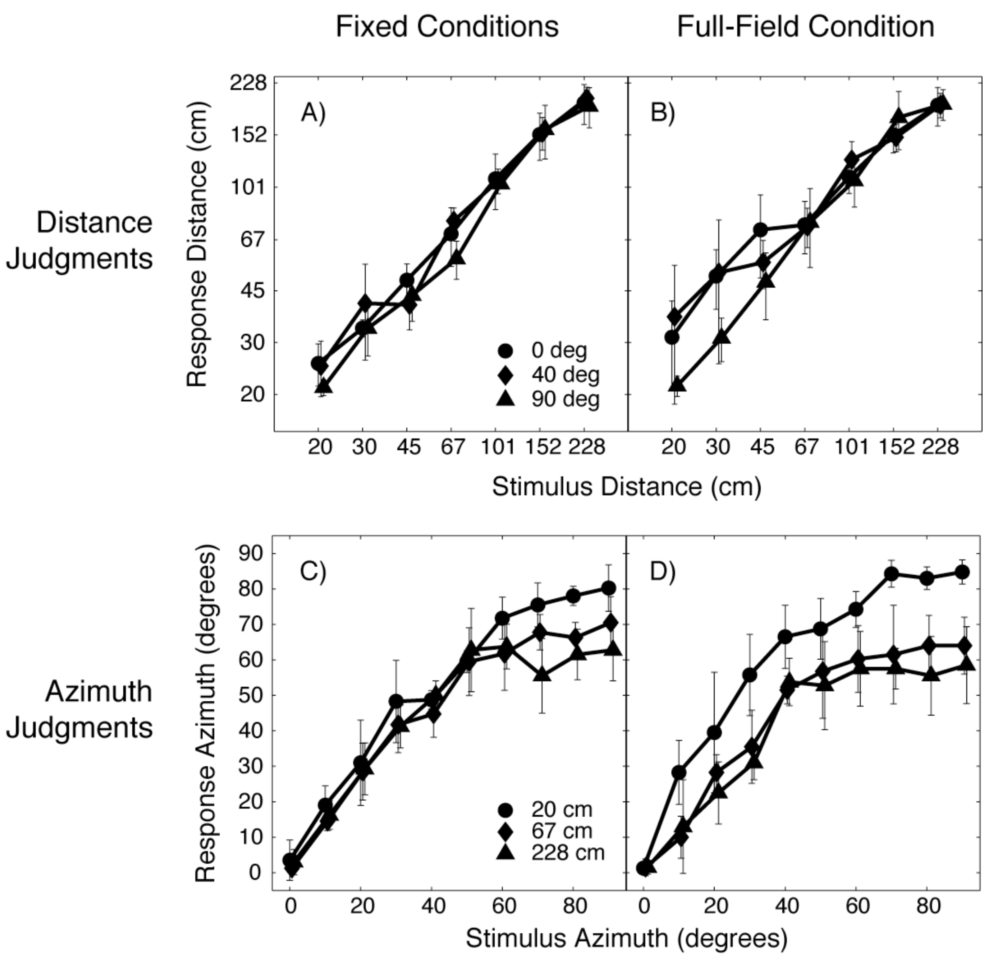

Figure 1. Mean perceived location in fixed and full-field conditions is relatively accurate and similar in fixed and full-field conditions; however, perceived azimuth varies with source distance. A) Mean perceived distance in the fixed-azimuth condition. B) Mean perceived distance in the full-field condition. C) Mean perceived azimuth in the fixed-distance condition. D) Mean perceived azimuth in the full-field condition.

Results in panel C of Figure 2 compare the standard deviation in judgments of source azimuth in the fixed-distance and full-field conditions. For both the fixed-distance and full-field conditions, response variability is small (on the order of 5˚) for sources near the midline, increases with source laterality out to source azimuths of about 30˚, and then is roughly constant for sources between 30˚ and 90˚ (with a standard deviation on the order of about 15˚). This finding is consistent with results of past localization experiments that show that sensitivity to changes in source laterality is greatest for sources near the median plane and worse as sources near the interaural axis (Mills, 1958). As with distance judgments, variability is generally larger in the dual-task, full-field condition (open symbols) than in the fixed-distance condition (filled symbols); however, in contrast with the distance results, this increase in variability is relatively modest. Results in panel D emphasize this point: the ratio of the standard deviation in the full-field task over the standard deviation in the fixed-distance task generally falls between 0.5 and 2.0, independent of source distance and source azimuth.

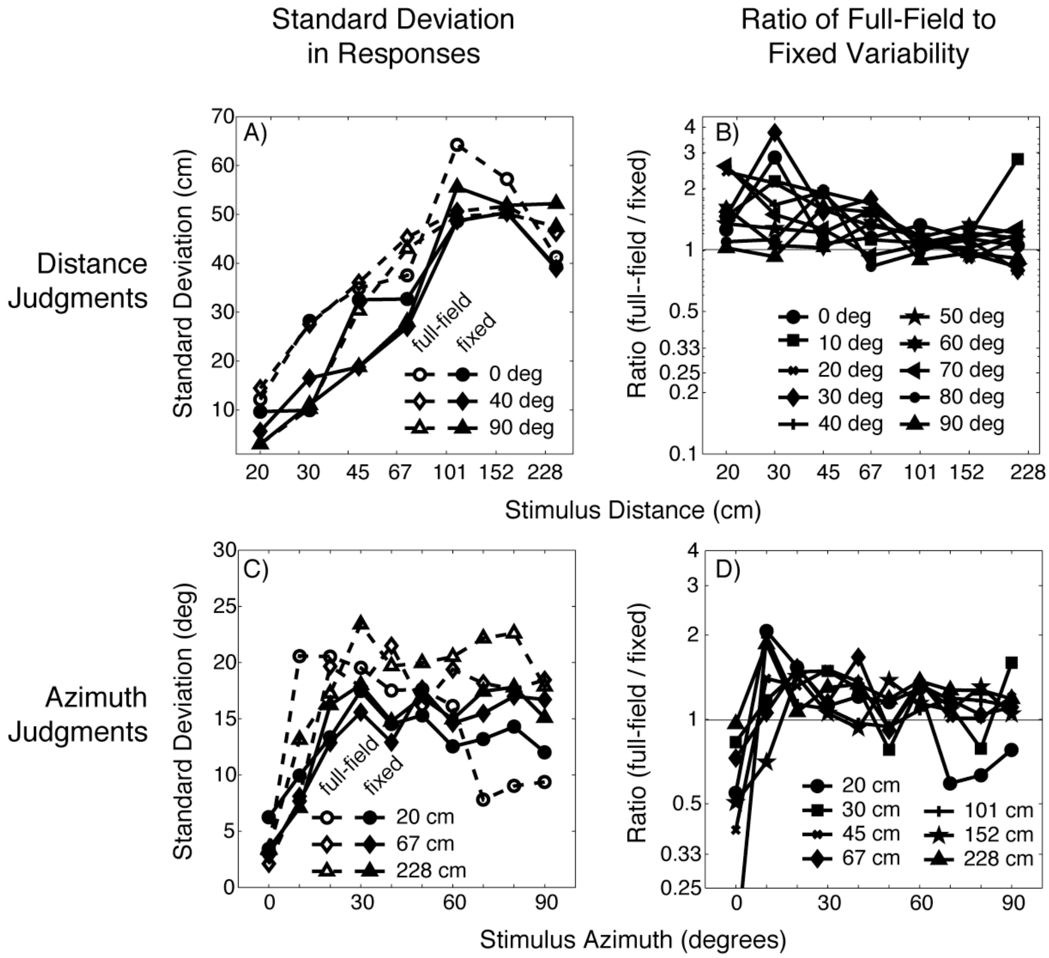

Figure 2. Response variability in both distance and direction is larger in the full-field condition than the corresponding fixed conditions. A) Standard deviation in perceived distance in full-field (open symbols) and fixed-azimuth (filled symbols) conditions. B) Ratio of standard deviation in full-field to standard deviation in fixed-azimuth conditions. C) Standard deviation in perceived direction in full-field (open symbols) and fixed-distance (filled symbols) conditions. D) Ratio of standard deviation in full-field over standard deviation in fixed-distance conditions.

4.1.3 Discussion

Mean distance and mean direction judgments are similar in single-task and dual-task conditions, suggesting that mean perceived location is unaffected by the task demands placed on a listener. While mean perceived distance is also independent of the simulated source azimuth, the mean perceived direction varies systematically with source distance. In particular, as source distance increases, judgments of source laterality for sources to the side of the listener become increasingly biased toward midline. This result may reflect the fact that reverberant energy, whose influence increases with increasing source distance, decreases the magnitude of ILD cues (Shinn-Cunningham et al., 2005). To the extent that ILD cues contribute to the perception of source laterality, the bias in azimuth judgments for lateral sources may be due to the influence of reverberant energy on ILDs in the stimuli.

Response variability is larger in both spatial dimensions (distance and direction) in the dual, full-field task than in the corresponding fixed conditions. The increase in response variability was especially pronounced for judgments of distance (where the full-field response variability was sometimes a factor of four greater than in the fixed-azimuth conditions). The cost (in increased response variability) associated with simultaneously judging source direction is also greater for sources close to the listener than for sources more distant from the listener. This effect may be due to the fact that, for nearby sources, ILD cues co-vary with distance and direction (Brungart & Rabinowitz, 1999; Shinn-Cunningham et al., 2000). For fixed-direction conditions, ILDs can unambiguously contribute to judgments of source distance; however, when distance and direction both vary from trial to trial, the listener must determine source direction in order to correctly intrepret ILDs to judge source distance.

Overall, these results show that perceived distance and direction are not independent in a virtual auditory display. When asked to simultaneously judge two source dimensions instead of a single dimension, response variability in each dimension is greater than in corresponding single tasks. This degradation in performance is larger for distance than direction. In addition to a central limit on performance, there is evidence that distance and direction interfere perceptually. In particular, mean perceived direction depends on source distance, with direction of lateral sources more biased when sources are distant from the listener than when they are close.

4.2 Information Transfer

4.2.1 Overall Information Transfer

Figure 3 plots the total information transfer in the fixed-azimuth, fixed-distance, and full-field conditions.

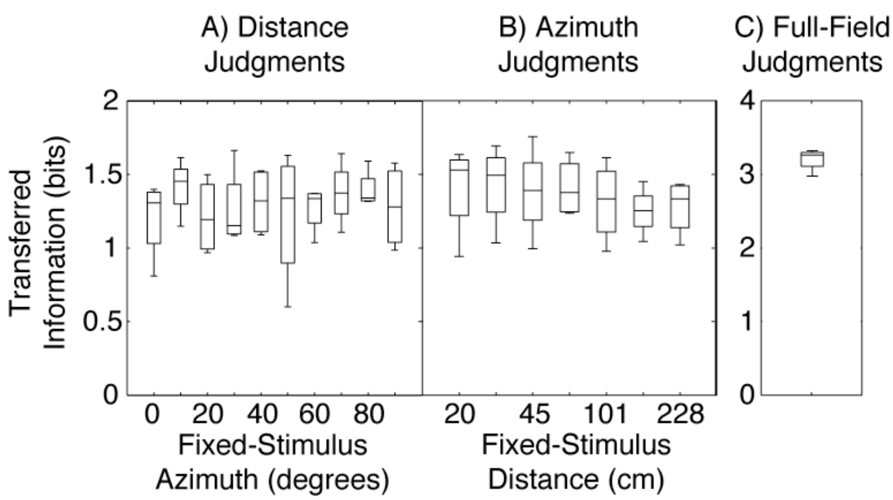

Figure 3. Total information transfer in the full-field conditions is roughly equal to the sum of information transferred in the two fixed conditions. A) Number of bits transferred by source distance in the fixed-azimuth conditions plotted as a function of azimuth. B) Number of bits transferred by source azimuth in the fixed-distance condition, plotted as a function of distance. C) Number of bits transferred by combinations of distance and direction in the full-field condition.

In the fixed-azimuth condition (panel A), information transfer does not vary systematically with azimuth, varying between slightly more than one bit to slightly less than 1.5 bits. In the fixed-distance condition (panel B), there is a weak trend for information transfer to decrease with increasing source distance, although the inter-subject variability is roughly of the same magnitude as these small changes with source distance. The total number of bits of information transferred by varying direction is slightly larger than that transferred by distance, averaging about 1.5 bits for nearby sources and 1.25 bits for the more distant sources. Overall information transfer is much larger in the full-field task than in either of the single tasks, averaging about 3.25 bits, which is slightly larger than the sum of the average number of bits transmitted by source distance and transmitted by source direction.

At first glance, these results are somewhat surprising. If the two source dimensions are independent, then the total number of bits transferred in the full-field condition should be, at most, the sum of the bits transferred by distance and the bits transferred by direction. If there is interference between the dimensions, then the sum of the bits in the two fixed conditions should be greater than in the full-field condition. Results in Section 4.1 show that there is increased response variability in each of the response dimensions in the dual task compared to the single tasks, which should decrease the information transferred in the full-field condition. The analysis shown in Figure 3 estimates that the total number of bits transferred in the full-field condition is greater than the sum of the bits transferred in the two dimensions in the fixed conditions. However, these direct comparisons across condition are not completely "fair," as the total number of responses in the cells of the confusion matrix differ across conditions. In the full-field condition, there are only 10 responses for each of the 70 possible stimuli presented. In the fixed-azimuth condition, there are 70 responses for each stimulus; in the fixed-distance condition, there are 100 responses per stimulus. Estimates of information transfer become increasingly less reliable and more positively biased as the ratio of the number of samples in each cell over the number of response categories (70 in the full-field condition) decreases (Castelloe & Woodworth, 1996; Miller & Nicely, 1955). Thus, it is likely that the estimate of information transfer in the full-field condition is biased towards a larger value than would be found with a larger sample. Nonetheless, these results suggest that the loss in information transfer caused by the dual-task demands in the full-field condition is not too great.

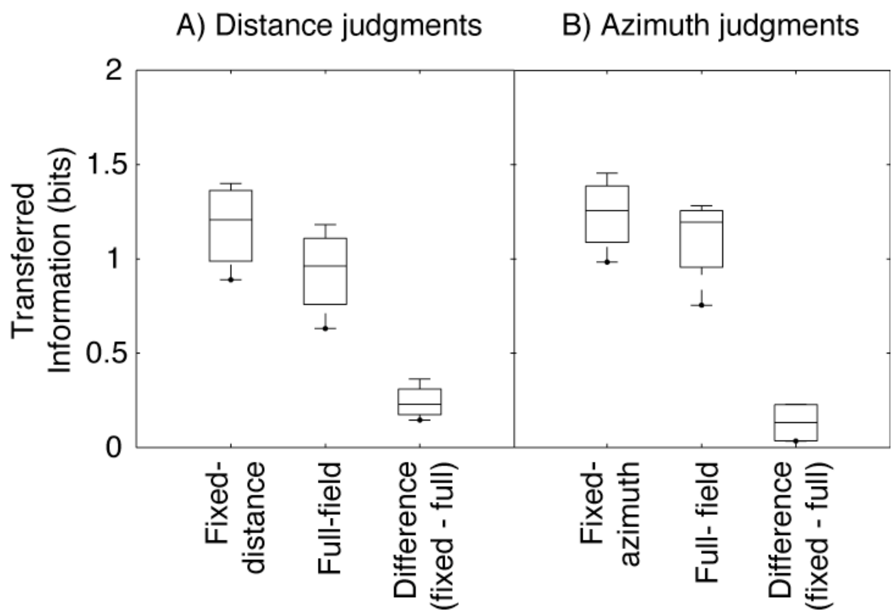

Figure 4. Information transfer in both the distance and direction dimensions is smaller in the full-field condition (collapsed across the other dimension) than in the corresponding fixed conditions (combined across the other dimension). A) Number of bits transferred by source distance in the fixed-azimuth conditions after combining all fixed azimuths compared to in the full-field condition after collapsing across azimuth. The difference in the number of bits is shown on the right. B) Number of bits transferred by source direction in the fixed-distance conditions after combining all fixed distances compared to in the full-field condition after collapsing across distance. The difference in the number of bits is shown on the right.

4.2.2 Information Transfer in Single Stimulus Dimensions for Single and Dual Tasks

Figure 4 plots the number of bits of information transferred in the data when the fixed and full-field conditions' confusion matrices are combined and collapsed to have the same number of cells and the same number of trials per cell, as described in Section 3.2. Panel A shows results for fixed-direction conditions (combined across direction) compared to results in the full-field condition when they are collapsed across direction. Panel B compares the results of the fixed-distance conditions (combined across distance) and full-field results (collapsed across distance). In both panels, the difference in the number of bits transferred in the single task and the dual task are plotted at the right of the panel.

In both panels, the number of bits transferred in the single task (left of panels A and B) exceeds the number of bits transferred in that stimulus dimension in the dual task condition (middle of panels). The difference on the right of each panel is also positive, consistent with this observation. The loss of information transferred in the full-field condition compared to in the fixed condition is larger when considering information transfer in the stimulus dimension of source distance (panel A) compared to source distance (panel B); the difference between fixed and full-field conditions is more positive in panel A than in panel B.

The absolute number of bits transferred by source distance and the number of bits transferred by source direction are remarkably similar when data are collapsed as described in Section 3.2. However, this similarity is somewhat misleading. In particular, mean perceived azimuth changes with simulated source distance. Thus, when estimating information transfer from azimuth judgments after collapsing across distance, the information transfer is reduced because data with different means (from near and far distances) are combined into a single entry in the confusion matrix. This factor should affect both the fixed-distance condition and the full-field condition approximately equally, since mean perceived direction varies similarly with source distance in both conditions (and thus is unlikely to affect the difference in information transfer plotted in the far right of panel B). However, this factor will reduce the estimated amount of information transferred by azimuth.

4.2.3 Discussion

Direct comparison of the number of bits transferred in the fixed-azimuth, fixed-distance, and full-field conditions shows that simultaneously varying both direction and distance enables a listener to extract more information than holding one stimulus dimension constant and varying only the other stimulus dimension. However, the amount of information extracted in one stimulus dimension is reduced when listeners must judge two stimulus dimensions on each trial, rather than concentrated all of their efforts on one dimension alone.

Overall, information transfer in the single dimensions of distance and direction used here are on the order of 1.5 bits each. When both distance and direction vary jointly, information transfer is approximately 3 bits, or roughly the sum of the number of bits that can be conveyed by distance and by direction (although, as mentioned above, the information transfer in the full-field condition is overestimated due to estimation bias). These results suggest that, despite the fact that distance and direction are not independent dimensions in an auditory display (e.g., see results in Section 4.2.1), the total information transferred through distance and direction jointly is of the same magnitude as the amount of information that would be transferred if the two dimensions were independent.

5 Summary and Conclusions

In these experiments, task demands have no significant effect on perceived source location; mean judgments of source distance and direction are similar when only one stimulus dimension varies at a time and when both distance and direction vary. However, distance and direction are not perceptually independent. For lateral sources, perceived source direction becomes increasingly biased towards the median plane as source distance increases, presumably because reverberant energy decreases the magnitude of ILD cues signifying source direction. While the physical effects of reverberant energy on binaural cues may explain the interaction between simulated source distance and perceived source direction, no previous studies have demonstrated the kind of systematic localization bias of directional judgments with increasing source distance reported here.

Response variability in both distance and direction increases when subjects must judge both stimulus dimensions compared to when only one spatial dimension varies. This effect is relatively modest for source direction, but can be quite large for source distance. This increase in response variability with increased task demands suggests that there is a central processing limit influencing performance, such that the ability to judge source distance is reduced when listeners must simultaneously judge source direction (and vice versa).

Despite the fact that distance and direction are not independent, the total information conveyed by distance and direction is roughly the same as the theoretical maximum that could be obtained if the two dimensions were independent. For instance, even though mean perceived direction depends on source distance, information transferred by varying source direction is nearly the same when distance is held constant and when distance and direction vary jointly.

Practically speaking, the most important result of these experiments is that judgments of source laterality become increasingly more biased and more variable with (1) increasing angular displacement from the median plane, and (2) with increasing distance. This result has important implications for designing a virtual auditory display, showing that even low levels of reverberant energy cause significant degradations in directional localization performance. In particular, directional localization accuracy is likely to be degraded whenever reverberation is included in a virtual auditory display, at least for lateral sources. However, reverberant energy is necessary to allow accurate distance perception and to improve the subjective realism of a virtual auditory display. Given that localization of sources near the median plane shows little degradation with the levels of reverberation used here, the benefits of adding reverberant energy to a virtual auditory display may often be worth any loss in localization accuracy for sources off to the sides of the listener. Nonetheless, these results directly demonstrate the kinds of tradeoffs in performance that must be considered when designing a virtual auditory display.

6 Acknowledgements

This work was supported by grants from the National Institutes of Health (R01 DC005778-02) and the National Science Foundation.

References

Brungart, D. S., & Rabinowitz, W. M. (1999). Auditory localization of nearby sources I: Head-related transfer functions. Journal of the Acoustical Society of America, 106(3), 1465-1479.

Castelloe, J., & Woodworth, G. (1996). Jacknife and bootstrap bias correction for single-subject information transfer in audiological testing (available at www.Stat.Uiowa.Edu/~gwoodwor/jackbootit.Pdf): University of Iowa.

Durlach, N. I., Tan, H. Z., Macmillan, N. A., Rabinowitz, W. R., & Braida, L. D. (1989). Resolution in one dimension with random variations in background dimensions. Perception and Psychophysics, 46, 293- 296.

Macpherson, E. A., & Middlebrooks, J. C. (2002). Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited. J Acoust Soc Am, 111(5 Pt 1), 2219-2236.

Mershon, D. H., & Bowers, J. N. (1979). Absolute and relative cues for the auditory perception of egocentric distance. Perception, 8, 311-322.

Mershon, D. H., & King, L. E. (1975). Intensity and reverberation as factors in auditory perception of egocentric distance. Perception and Psychophysics, 18, 409-415.

Miller, G. A., & Nicely, P. E. (1955). An analysis of perceptual confusions among some english consonants. J Acoust Soc Am, 27, 338-352.

Mills, A. W. (1958). On the minimum audible angle. Journal of the Acoustical Society of America, 30, 237-246.

Shinn-Cunningham, B. G. (2000, 2-5 April 2000). Learning reverberation: Implications for spatial auditory displays. Paper presented at the International Conference on Auditory Displays, Atlanta, GA.

Shinn-Cunningham, B. G., Kopco, N., & Martin, T. J. (2005). Effects of listener location and source location on reverberant binaural room impulse responses from a classroom. Journal of the Acoustical Society of America, in press.

Shinn-Cunningham, B. G., Santarelli, S., & Kopco, N. (2000). Tori of confusion: Binaural localization cues for sources within reach of a listener. J Acoust Soc Am, 107(3), 1627-1636.

Wightman, F. L., & Kistler, D. J. (1997). Factors affecting the relative salience of sound localization cues. In R. Gilkey & T. Anderson (Eds.), Binaural and spatial hearing in real and virtual environments (pp. 1-24). New York: Erlbaum.

Zahorik, P. (2002a). Assessing auditory distance perception using virtual acoustics. Journal of the Acoustical Society of America, 111(4), 1832-1846.

Zahorik, P. (2002b). Auditory display of sound source distance. Paper presented at the International Conference on Auditory Displays, Kyoto, Japan.