

## Toward the indirect measurement of perceived auditory distance

DONALD H. MERSHON and WILLIAM E. HUTSON  
North Carolina State University, Raleigh, North Carolina

In previous research on perceived auditory distance, simple verbal reports have often been used, because alternative measurement procedures were limited. This research has explored the question of whether Gogel's "indirect" methods (for measuring visually perceived distance) could also measure auditory distance. Observers described the apparent horizontal alignment of two sound sources, presented either with the head stationary or before and after a lateral head movement. Room reverberation was manipulated in an effort to generate the sorts of errors in perceived distance that were necessary to test the predictions. Although consistent underperception occurred effectively when reverberation was low, the corresponding condition of overperception was not produced by the use of a highly reflective room. Nevertheless, taken as a whole, the present data were consistent with those predicted from Gogel's analysis of the consequences of errors in perceived distance. The development of an indirect measure of auditory distance appears feasible.

The perception of auditory distance has not been studied as extensively as the perception of visual distance. Nevertheless, quite a bit is now known about the types of cues that create (or modify) an impression of the distance between a listener and a source of sound (see Mershon, Ballenger, Little, McMurtry, & Buchanan, 1989). Among the factors or cues considered important are the reverberation patterns provided by the acoustical environment, the level and spectral characteristics of the sounds involved, binaural changes associated with directional locations away from the midline, and possible influences from familiarity with the sounds being presented. Additional factors are certainly possible, and even those indicated above may actually represent different effects that have been loosely grouped together under a single label.

Unfortunately, the study of auditory distance perception has been hampered until recently by the limited number of ways in which one might ask about the variable of interest. Most investigators ask for a direct verbal report (oral or written) describing the apparent distance (Mershon & Bowers, 1979). Other promising approaches include asking the listener to walk to the perceived position of the source (Loomis, Hebert, & Cicinelli, 1990). However, techniques such as walking may require a fairly large response space to implement and may, in addition, create problems for judgments of loudspeaker (as opposed to headphone) presentations, since it is difficult to remove the source from the listener's path without providing extraneous information about its position. There may also be limitations in the range of distances that can effectively be reported with such a technique. To describe a source

as seeming 200 m away would require that the listener be able (and willing) to walk 200 m.

In the present study, we evaluated a different approach. Gogel (1981) reviewed a technique in which the measure of one perceptual variable is used to determine the value of another. Such a procedure can provide an "indirect" measure of perceived distance, with the additional advantage that the listener is usually unaware of the relationship involved. The listener is thus less likely to introduce cognitive modifications of his or her responses than if a direct request were made for information about distance itself. (For example, a listener would be unlikely to restrict a choice of responses on the basis of knowledge that the laboratory was too small to allow a correct report that a target was 100 m away.)

One may consider the basis of Gogel's (1981) *head motion* task in relation to a simple diagram (see Figure 1). An observer makes a lateral head motion from Position 1 to Position 2 while watching a target at some physical distance  $D$ . If the target is correctly perceived in distance (i.e., perceived distance  $D' = D$ ), the target will appear to be stationary in space. If the target is misperceived in distance, however, the target will appear to move concomitantly with the observer's head—in the same direction as the head if the target is underperceived, and in the opposite direction to the head if the target is overperceived. The extent of the apparent target motion is directly related to the magnitude of the error in perceived distance, and the sign of the movement indicates the type of error. Thus, one can obtain measures of the apparent motion and use them to determine the perceived distance.

To simplify the observer's task, one may introduce physical motion in another direction (e.g., vertically) simultaneously with the lateral relocation of the observer's head. Errors in perceived distance will now create an ap-

---

William E. Hutson's current address is IBM Corporation, P.O. Box 12195, Research Triangle Park, NC 27709. Correspondence should be addressed to Donald H. Mershon, Department of Psychology (Box 7801), North Carolina State University, Raleigh, NC 27695-7801.

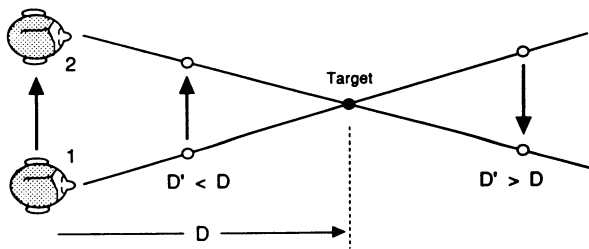


Figure 1. Simplified representation of the basis for Gogel's head motion technique for measuring perceived distance.

parent horizontal motion component. Since the real and apparent motions add vectorially, the target will appear to move along a diagonal path (see Gogel, 1979). The observer may then report the slope of this motion path itself as the indirect measure for perceived distance.

If, as Gogel has postulated, indirect techniques represent effective ways of measuring errors in perceived distance, then it should be irrelevant whether the perceptual distance variable is visual or auditory. The critical factors in predicting the outcome are based upon the theory of phenomenal geometry (Gogel, 1990). These factors include: (1) the perceived direction of the target, (2) the perceived distance to the target, and (3) the observer's own perception of self-movement. None of these factors is explicitly visual. Therefore, we set out to determine whether an indirect technique could potentially be applied to the measurement of perceived auditory distance. Rather than move a loudspeaker, however, we presented a single discrete burst of sound from each of two loudspeakers located at different heights in the auditory field. One burst occurred with the listener's head at a position to the right of center; the second occurred immediately after the listener had moved to a position left of center.

## METHOD

### Observers

Eighty students from introductory psychology classes served as the listeners for this experiment. All reported normal hearing in both ears, although no audiometry was done to confirm this. All were unfamiliar with the laboratory and the purpose of the experiment.

### Laboratory

The laboratory used for this research has been described previously (Mershon et al., 1989). Briefly, the experiment was conducted within a test room that was 7.3 m on a side and 3.6 m high. The plasterboard walls could be covered with fiberglass acoustical panels (each 0.6 m square). The floor could be covered by a low-pile carpet. Two conditions were employed. For one (the *live* room) condition, neither carpet nor wall panels were in the room and the average  $T_{60}$  for 500 Hz to 8 kHz was 1.66 sec.<sup>1</sup> For the other condition (the *dead* room), all four walls and the floor were covered with absorbant materials, resulting in  $T_{60} = 0.36$  sec. These reverberation values had previously been shown to produce systematic differences in the perception of auditory distance, at least over the first few presentations within each room. Specifically, the live room was expected to produce perceptions of distance that were greater than the 3-m distance to the physical source. The dead room was expected to produce perceptions of distance that were less than the physical distance to the source.

A special backdrop was used behind the listener. This consisted of a large section of leaded vinyl material covered by 4-in.-thick Sonex

foam in a standard acoustical wedge pattern. This backdrop minimized early reflections of sounds from the wall behind the listener.

A system of small overhead loudspeakers provided a low level of uniform masking noise to obscure any sounds intruding from the adjacent hallways or control room. Lighting was provided by a few relatively dim incandescent lights.

### Apparatus

Target sounds were generated, filtered, and gated using appropriate modules from Coulbourn Instruments. Resulting signals were sent to a Crown DL-2 preamplifier and PS-200 amplifier before being transduced by one of two 5-in. loudspeakers positioned on a stand in the test room. The final signals presented to the listener consisted of bursts of noise from 500 Hz to 8 kHz, lasting 2 sec each (with rise-decay ramps of 100 msec). They were presented at a level producing approximately 70 dB (A-weighted) sound at the position(s) of the listener's head.

The target loudspeakers were mounted on a support frame 3 m from the listener. One loudspeaker was always centered and fixed in the display 19 cm below the listener's ear level. The other loudspeaker was 19 cm above ear level and was mounted on a track that allowed it to be positioned either directly above the lower speaker or as much as 38 cm to the left or right of that position. A total of 11 positions for the upper speaker were used: the centered position and five equally spaced positions to either side of the fixed lower speaker. (See Figure 2.)

Responses were obtained by using two custom-built boxes containing synchronous motors. The shaft of each motor protruded from the top of its box. A lever was attached to each shaft. For the listener's box, the lever was a flat bar of metal about 20 cm long, which could be rotated around its center to represent different perceived orientations of the two speakers. (The entire assembly was tilted to increase the stimulus-response compatibility between the upper and lower ends of the lever and the upper and lower loudspeakers.) For the experimenter's box, the shaft was attached to a pointer, which moved across a protractor and thus provided a convenient way to read the setting made by the listener. Reorienting either shaft caused the other shaft to turn to a corresponding position.

### Procedure

Listeners were greeted at the door of the control room. Ear level was measured, and while the listener was reading a set of general instructions, the experimenter entered the test room to adjust the height of a platform on which the listener would stand. The listener was then blindfolded before being led into the test room, positioned on the platform (with feet slightly separated), and given additional instructions. For each trial, the listener heard two sounds—the first always from the lower speaker and the second from the upper speaker. There was a 1.5-sec pause between the sounds. Following completion of the second sound, the listener reported the apparent spatial relationship of the two speakers by using the preferred hand to rotate the movable bar on the response box to an appropriate orientation. When the experimenter had recorded the response, he moved the bar back to a centered position for the next trial.

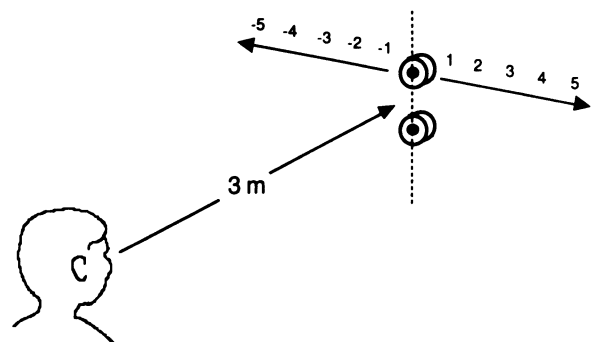


Figure 2. The arrangement of the two loudspeakers used in the present study. The upper speaker was moved randomly without replacement among eleven possible horizontal positions. The center or "0" position was directly above the lower speaker.

Each listener reported on the apparent spatial alignment of the upper loudspeaker relative to the lower one under each of two conditions: once with the head stationary in the midline of the display, and once with a lateral movement of the head. That is, for half of the observations, the listener heard the lower source with the head held approximately 30 cm to the right of the midline. The listener then moved during the 1.5-sec pause to a position in which the head was a similar distance to the left of the midline and heard the second (upper) source from that position. (Instructions and practice on moving the head in a specified, level, straight path were provided immediately prior to this part of the experiment.)

After completion of all the judgments of perceived alignment, the listener was presented with a single sound from the lower loudspeaker while standing in a centered position. The listener then provided a verbal report of the apparent distance to this final sound in "feet or inches or some combination of feet and inches."

### Design

Different groups of listeners were presented with the sounds in the two different rooms, because of the impossibility of changing the room conditions rapidly. Altogether, 40 listeners were presented with sounds in each room. Each listener, regardless of the room, heard the sounds under two conditions—the head stationary or the head moving—with the order of these conditions counterbalanced across listeners. For each such condition, the listeners reported on the alignment of the loudspeakers for each of the 11 possible horizontal positions of the upper speaker (with the order of positions randomized across different listeners). The resulting design for the main portion of the experiment was a  $2 \times 2 \times 11$  factorial, with room reverberation as a between-listeners variable and both head condition and physical alignment as within-listener variables.

## RESULTS

There were three primary expectations for this experiment. First, we expected that the different conditions of room reverberation would predictably modify the perceived distances to the sounds. Second, we expected that in the live room all of the alignment judgments of the upper speaker would shift to the right (consistent with conditions in which  $D' > D$ ), when the head moved between the presentations. We further expected that in the dead room all alignment judgments would shift to the left (consistent with conditions in which  $D' < D$ ), when the head moved between the presentations. (In neither room should there have been any shifts in alignment judgments when the head was stationary. Such conditions primarily provided a control situation to demonstrate the accuracy of such judgments in relation to the physically different alignments.) In the following sections, we will examine the data with respect to these expectations.

### Perceived Distance Reports

The physical distance to the plane of the loudspeakers was 3 m. Median reports of the apparent distance to the final sound were 3.04 m in the live room and 1.67 m in the dead room (semi-interquartile ranges were 1.07 and 0.46 m, respectively). Furthermore, in the live room, 15 reports were of values less than 3 m, while 11 were of values greater than 3 m. This difference is not significant by a sign test. In the dead room, on the other hand, 23 values were less than 3 m, while only 1 was greater. This difference is significant ( $p < .001$ ) by sign test.

Using these verbal judgments as the best available (albeit imperfect) indicator of apparent distance, it appears

clear that the expectation of manipulating the perceived distances by changing room reverberation was only partially successful. The dead room resulted in conditions in which the distance to the source was underperceived. There is no evidence, however, that the live room created any systematic error at all in the perceived distance. The alignment judgments in the live room, therefore, should show no shift, even for the condition with the moving head.

### Alignment Judgments

The reports of alignment were gathered as values from the protractor (i.e., as angular deflections of the pointer from  $90^\circ$ ). Values less than  $90$  indicate that the apparent location of the upper speaker was to the left of center; values greater than  $90$  indicate that the upper loudspeaker was to the right.

Before analyzing the results for the predicted shift in alignment, it is important to examine the ability of the listeners to perform the basic task. In every case, it is reasonable to expect that alignment reports should systematically change with the physical position of the upper loudspeaker. Indeed, the responses within a particular condition should provide at least a reasonable fit to a straight line. Before proceeding, therefore, the results for each listener were fitted by a least-squares technique. If a significant fit ( $\alpha = .05$ ) were not obtained for any condition, all the data from that listener were removed from the sample. This procedure resulted in a final pool of 33 listeners in the live room and 28 in the dead room. Inasmuch as each listener provided both head-stationary and head-moving reports, these sample sizes were considered sufficient. (Also, comparable full-data analyses yielded results essentially the same as those which follow.)

The data from listeners who performed the task adequately are shown in Figure 3. For the live room condition (Figure 3A), it is clear that no shift in judgments is present. The functions for head-stationary and head-moving overlap extensively. For the dead room condition (Figure 3B), however, the expected shift has occurred as predicted. All alignment reports shifted leftward when the head was moved between the two presentations. The results for the head-stationary condition, on the other hand, are virtually the same as for the functions from the live room. (Although there was also a general shift to the left for all functions, in that the zero position of the upper speaker usually yielded a response of less than  $90^\circ$ , this small constant error was most likely the result of unequal wrist flexion for right-handed listeners; inward flexion—i.e., leftward rotation—is a little easier than outward flexion, resulting in the observed bias.)

The difference in the dead room conditions was tested two ways. First, a two-way repeated measures analysis of variance was used. The factors of lateral position of the upper speaker and of head motion (stationary vs. moving) were both significant [ $F(10,270) = 68.29, p < .0001$ , and  $F(1,27) = 12.69, p < .002$ , respectively]. The interaction of these variables was not significant. Second, the set of fitted curves was used as the basis for another compari-

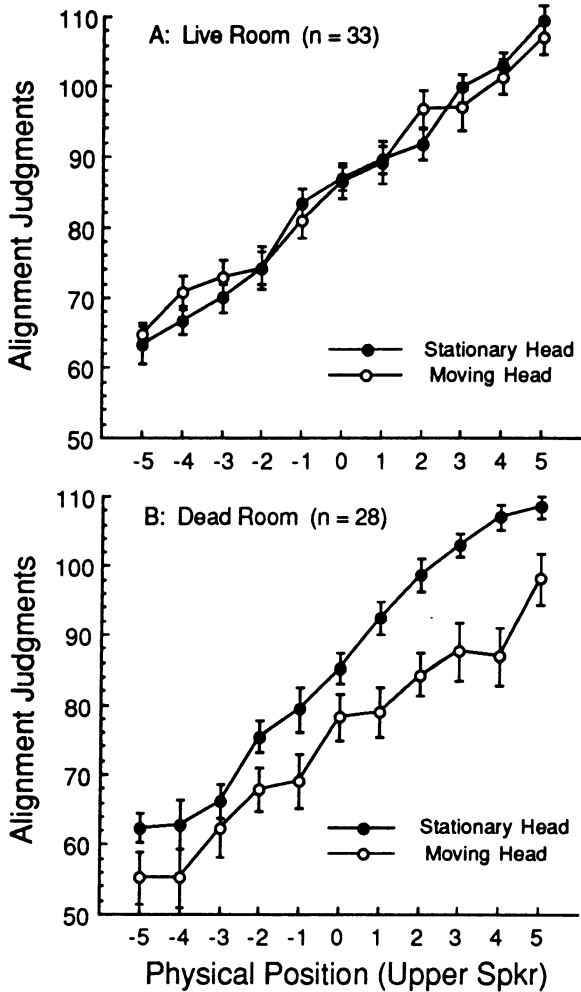


Figure 3. Mean judgments of the horizontal alignment of the upper speaker relative to the lower for the 11 positions of the upper speaker. Bars on the data points represent  $\pm 1$  SE.

son. A sign test of the intercept values for the head-stationary and head-moving conditions for each listener was also significant at  $p < .002$ .

## DISCUSSION

The results of this study lead to three major conclusions. First, the data are all in agreement with the predictions of Gogel's theory of phenomenal geometry. Although we were unable to produce a situation with overperception of distance, the outcome of the alignment judgments is at least consistent with that failure. Obviously, however, the argument will be stronger when a situation of overperception can be established for testing the predictions more explicitly. In the present experiment, it is likely that the listener was provided with too many presentations during the experimental session. As previous research has shown, repeated presentations in a reflective environment tend to move reports of perceived distance toward veridicality (Mershon et al., 1989). With the dead room conditions, however, there is no question that the expected shift in alignment did occur. It is difficult to suggest any alternate explanation for such a shift, since it only occurred for the head-moving condition. Unusual wall reflections are thus an unlikely basis for the shift, even if reflections had not already been minimized by the use of the absorbent materials.

Possible contributions from the (untreated) ceiling do not seem much more likely in that this surface was also present for the live room conditions in which no shift was observed.

The second conclusion is that an indirect technique based on the head motion effect shows promise for measuring perceived auditory distance. In the dead room, the calculation of perceived distance based on the shift in alignment judgments yields a  $D' = 253$  cm. Although this is not as close as the verbal reports for the same room, it is a significant shift in the correct direction. More importantly, without additional experimentation, there is no way to validate the precise values of the verbal reports. It is possible that the underlying perception of the source's distance was more similar to the results from the head motion task and that the verbal report was less reflective of the underlying perception. For the present, we must probably be content with demonstrating significant changes in the same direction.

The third and final conclusion concerns the status of an alternate explanation for the fundamental head motion effect itself. Leibowitz, Post, and Sheehy (1986), for example, have suggested that the apparent visual displacement that occurs for certain targets when one moves one's head is the consequence of peripheral eye movements. Specifically, they suggest that errors in convergence that may occur under reduced viewing conditions lead to inappropriate visual readjustments when the vestibulo-ocular reflex (VOR) is triggered by head movement. This error in reflexive compensation causes the image of the target to slip across the retina and initiates a pursuit eye movement to re-center the target's image. This corrective pursuit motion supposedly creates the appearance of motion and not any error in perceived distance.

It should be clear that the foregoing interpretation based on the VOR cannot readily be applied to the apparent motions of an auditory target. There is no visual target to be followed in the first place, and no corrective pursuit eye movement is necessary. Even if one wished to argue that the eyes assume some particular vergence value beneath the blindfold and that some eye movements may occur in the absence of any visual target, there is no reason for the shift in alignment judgments to occur for one room condition but not the other. Parsimony seems to call for an interpretation based on phenomenal geometry rather than one that is constrained to spatial judgments in the visual modality.

## REFERENCES

- GOGEL, W. C. (1979). The common occurrence of errors of perceived distance. *Perception & Psychophysics*, **25**, 2-11.
- GOGEL, W. C. (1981). Perceived depth is a necessary factor in apparent motion concomitant with head motion: A reply to Shebilske and Proffitt. *Perception & Psychophysics*, **29**, 173-177.
- GOGEL, W. C. (1990). A theory of phenomenal geometry and its applications. *Perception & Psychophysics*, **48**, 105-123.
- LEIBOWITZ, H. W., POST, R. B., & SHEEHY, J. B. (1986). Efference, perceived movement, and illusory displacement. *Acta Psychologica*, **63**, 23-34.
- LOOMIS, J. M., HEBERT, C., & CICINELLI, J. G. (1990). Active localization of virtual sounds. *Journal of the Acoustical Society of America*, **88**, 1757-1764.
- MERSHON, D. H., BALLENGER, W. L., LITTLE, A. D., McMURTRY, P. L., & BUCHANAN, J. L. (1989). Effects of room reflectance and background noise on perceived auditory distance. *Perception*, **18**, 403-416.
- MERSHON, D. H., & BOWERS, J. N. (1979). Absolute and relative cues for the auditory perception of egocentric distance. *Perception*, **8**, 311-322.

## NOTE

1. Room reverberation is often expressed by the  $T_{60}$  value. This represents the time required for a sound within the room to decay by 60 dB. It will differ for different frequencies and is usually expressed as an average across the range of interest.