Image Cover Sheet

CLASSIFICATION	SYSTEM NUMBER 513090
UNCLASSIFIED	

TITLE

Relation Between Performance and Confidence Ratings for Sound Localization in Free-Field and Virtual Acoustic Space

System Number:

Patron Number:

Requester:

Notes:

DSIS Use only:

Deliver to:

SL 1999-093

Relation between Performance and Confidence Ratings for Sound Localization in Free-Field and Virtual Acoustic Space

G.R. Arrabito, J.R. Mendelson, S.L. Van Blyderveen, R.B. Crabtree and M.J. Boultbee
DCIEM, 1133 Sheppard Ave. W., P.O. Box 2000, Toronto, ON M3M 3B9

INTRODUCTION

A relatively new technology, the three-dimensional (3-D) audio display, is being explored for improving aircrew performance. The presentation of virtual auditory cues over headphones to serve as warnings (e.g., weapons, other aircraft) to the aircrew is one application. Accurate localization of the virtual cues necessitates that the aircrew exhibit a high degree of confidence when making localization judgements.

In general, individuals are overconfident in their performance on knowledge-based tasks [1]. This overconfidence is exacerbated the more difficult the knowledge-based task becomes [2]. Underconfidence is more common in perceptual-based tasks, particularly when the task is not difficult [1]. If the perceptual task is comparable in difficulty to a knowledge-based task, then overconfidence is exhibited [1].

It could be detrimental if a pilot, who is in error of his/her localization judgements, is nonetheless extremely confident in making that judgement. To the best of our knowledge, there is no reported study on confidence ratings for localization judgements. The use of confidence ratings on localization judgements in free-field and virtual acoustic space was investigated in this study. Subjects made confidence ratings on a seven-element Likert scale after each localization judgement.

METHOD

Subjects. Five females and five males participated (mean age was 28.1 years). All subjects had normal hearing as verified by audiometric screening.

Stimuli and Apparatus. The stimuli were low-pass 4 kHz and 14 kHz, and high-pass 4 kHz white noise with duration of 300 ms with a 50 ms onset and decay. These were chosen to allow an assessment of the effectiveness of binaural and spectral cues.

Testing was performed in an Industrial Acoustics Company sound-attenuating listening booth. The booth contained a chair that was positioned in the centre of eight Axiom Millennia loudspeakers (model AX1.2) arranged in a circle with a radius of one metre centered at the seated listener's head. Each loudspeaker was mounted on a Yorkville adjustable microphone stand so that the vertical midpoint of each speaker was at the same height as the listener's ear level. The loudspeakers were placed at 45° intervals ranging from 0° to 315° azimuth increasing clockwise on the horizontal plane with 0° positioned directly in front of the listener. The transmitter of a Polhemus 3Space Fastrak magnetic head tracker was suspended from the ceiling of the sound booth at approximately 12cm directly above the listener's head to monitor the subject's head position. The tracker's receiver was placed on the top of the listener's head and was held in position by a headband worn by the listener. Localization judgements were made on a response box whose buttons were arranged in the same configuration as the loudspeaker array. A separate seven button response box was used for collecting confidence ratings. These buttons were labeled -3 (very unconfident), -2 (moderately unconfident), -1 (somewhat unconfident), 0 (neutral), +1 (somewhat confident), +2 (moderately confident), and +3 (very confident).

A Stax electrostatic headphone (model SR-l Signature) was used to present the stimulus in virtual acoustic space (VAS). The head-related transfer functions (HRTFs), digital filters for synthesizing the location of a sound in VAS, used in this study were measured from the first author. The Tucker-Davis Technologies system in conjunction with custom software running on a personal computer was used to present the stimulus in free-field and VAS.

Experimental Design. A 2 (auditory presentation mode) X 3 (acoustic stimulus) X 8 (azimuth) X 4 (session) within-subject repeated measures design was employed to measure localization performance and confidence ratings. Each acoustic stimulus condition was presented in both free-field and VAS at the eight azimuth positions on the horizontal plane described above. A Latin Square was employed to counterbalance auditory presentation modes, acoustic stimuli, and azimuth positions across subjects and sessions.

Procedure. On each day of testing the subject completed the three acoustic stimuli conditions in free-field and VAS. After all three acoustic stimuli conditions were completed in the same auditory presentation mode (e.g., free-field), the subject proceeded to the other auditory presentation mode. The ordering of auditory presentation modes was randomly determined. Each acoustic stimulus condition consisted of 8 practice trials followed by 104 experimental trials, with each azimuth position used once in the practice trials and 13 times in the experimental trials. Each trial began by flashing a 0.5 second light on the wall in front of the subject followed by a 0.5 second delay prior to the presentation of the stimulus. The subject's task was to identify the perceived location of the stimulus followed by a confidence rating on the seven-element Likert scale for that localization judgement.

Localization judgements and confidence ratings were made on the response boxes described above. If, during the presentation of the acoustic stimulus, the subjects moved their heads more than 2° in any direction of yaw, pitch or roll as monitored by the head tracker, then the subjects were notified by flashing LEDs to reposition their heads to the "straight-ahead" position. The trial was discarded and was presented again at the end of the current block of trials. No feedback was given to the subjects. Subjects completed two practice and four experimental sessions, each on separate days.

RESULTS AND DISCUSSION

A preliminary analysis of the pooled data on the four experimental sessions revealed that localization performance as measured by percent correct localization judgements was better in free-field (95.9%) than VAS (76.1%). Subjects made more reversals (i.e., perceiving the mirror image of the presented sound source) in VAS (9.2%) compared to free-field (1.6%). Performance was nearly identical regardless of the acoustic stimulus. These findings appear to be in partial agreement with those reported in [3].

Figure 1 shows the confidence ratings corresponding to correct

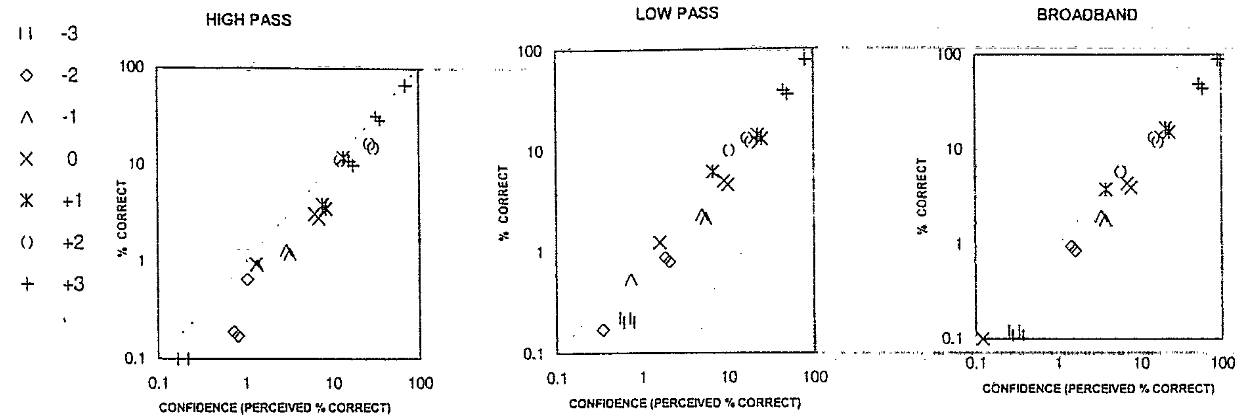

Figure 1. Confidence ratings corresponding to correct versus perceived localization performance. Data are shown for each acoustic stimulus condition in free-field (non-shadowed) and VAS (shadowed). The dotted line represents perfect calibration.

localization judgements as a function of confidence ratings corresponding to perceived localization performance. Subjects had higher confidence in their free-field responses compared to those made in VAS. In the free-field, subjects were close to "perfect calibration". An individual is said to be well calibrated if, in the long run, for all propositions assigned a given probability the proportion correct is equal to the probability assigned [2]. However, subjects were overconfident in VAS, which suggests that the choice of the HRTFs made the localization task too difficult.

The VAS localization performance of the first author, who was one of the subjects, was nearly identical to his free-field performance. He had closer to perfect calibration than all the other subjects. This was most likely due to the fact that he was the only subject who used his own HRTF. Indeed it has been argued that listeners localize a virtual sound better when using HRTFs measured from their own heads ("personal"), as compared to HRTFs measured from a different head ("generic") [3].

The present results suggest that if virtual auditory cues are to serve as warnings to the aircrew then the use of generic HRTFs may convey an inaccurate sense of confidence for making correct localization judgements. They further suggest that the use of personal HRTFs may provide closer to perfect calibration compared to generic HRTFs. However, acquiring personal HRTFs requires a substantial investment in infrastructure and equipment and is currently impractical in many applications. Methods need to be developed to quickly and accurately select a generic HRTF for conveying accurate localization.

REFERENCES

1. Keren G. On the ability of monitoring non-verdical perceptions and uncertain knowledge: Some calibration studies. Acta Psychologica 1988; 67:95-119.

2. Lichtenstein S, Fischoff B & Phillips D. Calibration of probabilities: The state of the art to 1980. Chapter 22. In: D. Kahneman, P. Slovic and A. Tversky, eds. Judgement under uncertainty: Heuristics and biases. Cambridge University Press, Cambridge, 1982.

3. Wightman FL & Kistler DJ. Headphone simulation of free-field listening part II: Psychophysical validation. J Acoust Soc Am 1989; 85(2): 868-878.

#513090