Auditory Navigation Performance is Affected by Waypoint Capture Radius

Bruce N. Walker and Jeff Lindsay

Sonification Lab, School of Psychology
Georgia Institute of Technology
654 Cherry Street NW
Atlanta, GA, USA 30332-0170
bruce.walker@psych.gatech.edu , gte457e@prism.gatech.edu

ABSTRACT

Non-speech audio navigation systems can be very effective mobility aids for persons with either temporary or permanent vision loss. Sound design has certainly been shown to be important in such devices [e.g. 1]. In this study we consider the added factor of capture radius. The capture radius of an auditory beacon is defined as the range at which the system considers a user to have reached the waypoint where the beacon is located. 108 participants successfully navigated paths through a virtual world using only non-speech beacon cues. Performance differed across the capture radius conditions. Further, there was a speed-accuracy tradeoff, which complicates the design decision process. Implications of these results for the design of auditory navigation aids are discussed, as are other ongoing and future studies.

1. INTRODUCTION

When visual information is not available, either due to permanent vision loss or to temporary conditions such as smoke or darkness, the fundamental task of navigating through our environment can become a major challenge. Technological solutions involving auditory displays may be able to assist a wide range of users in the non-visual navigation task, but there are many unanswered questions regarding the best ways to present assistive information through an auditory interface, as well as how to design the human-system interaction. Walker and Lindsay [1] described the results of an initial study aimed at determining the best auditory beacons to be used in audio navigation aids. In the present paper we provide an update to that initial report, including expanded data and additional conclusions, and we discuss new and ongoing investigations in the development of an effective auditory navigation system.

There are approximately 11.4 million people with vision loss in the United States. Of this population, ten percent have no usable vision; and by 2010 these numbers will nearly double [2]. The prevalence of blindness rises steadily with age to the extent that nearly two-thirds of people with vision loss are 65 years of age or older [3, 4]. This rise in the average age of people with severe visual impairment is the result of the increase in average age of the general population, and the increased prevalence of diabetic retinopathy, macular degeneration and glaucoma in this country. Just as everyone else, visually impaired individuals wish to lead an active, independent life which includes freely moving about the community, running errands, shopping, working, and taking advantage of the service and entertainment opportunities in their environment.

In addition to persons with vision loss, there are whole classes of persons who have normal vision but for whom temporary smoke, fog, darkness, fire, or other environmental conditions prevent them from seeing their immediate surroundings, and can lead to disorientation and an inability to navigate from place to place. Firefighters in a smoke-filled building may not be able to locate the stairwell; military personnel in darkness or under water may not be able to reach a particular rendezvous location; police in the midst of a protest may lose orientation due to thick tear gas. Also, even when people can see, during some tasks they may be unable to use vision for navigation since it is required for another concurrent task. Broadly, then, the groups of individuals who can benefit from auditory navigation interfaces can be summarized as those who cannot see and those who cannot look.

1.1. Prior Investigation

There are several systems that have been developed that use sound to help persons with vision loss through an environment. They have typically been designed for persons with visual impairments One representative system is the Personal Guidance System (PGS) [5, 6]. The PGS interface consists of a virtual 3D auditory environment where a computer creates spatialized speech beacons such that the perceived location of the beacon is at the place that the semantic content in the beacon refers (e.g. “Doorway here” as an auditory beacon).

However, *non-speech* audio cues are important to consider because there are several drawbacks to using exclusively speech sounds. First, speech beacons are harder to localize in a virtual audio environment than non-speech beacons [7]. Users also give speech beacons low ratings for quality and acceptance [7]. Second, the speech-based interface cannot display a large amount of information, as two or more speech beacons presented simultaneously are difficult to attend to, given the limited human speech processing capacity [e.g., 8]. It would also most certainly not be possible to use the speech-based interface and carry on a conversation at the same time [e.g., 9]. Third, spoken messages in such a system are each generally more than a second long, so the system is often talking. This inefficiency of speech can result in a cluttered and annoying listening environment.

As a result of these findings, our own auditory navigation projects have focused on non-speech audio. Walker and Lindsay [1] evaluated different beacon sounds for use in a navigation aid, based on the findings of Tran et al. [7]. Specifically, Tran et al. studied a number of sounds (both speech and non-speech) to be used in navigation aids and concluded that a wide-band, non-speech sound was most effective. The sonar-like sound they used was indicated as the best out of their stimulus set. Walker and Lindsay [1] determined that a broadband noise beacon was, indeed, the most effective for their navigation task. In that study, however, the sonar-like sound was actually the worst performer. These results parallel the findings of Lokki, Grohn, Svioja, and Takala [10], in which a (pink) noise source was found to be more effective than musical instrument sounds in moving through a virtual world. Thus it seems that beacon sound does have an effect on navigation effectiveness, but it is likely the general nature of the sound (e.g., broadband), and not the specific sound that should be considered as paramount.

In the course of the work reported by Walker and Lindsay [1] there were naturally questions that remained to be explored. In particular, issues were raised about how the user would actually move from waypoint to waypoint along a path. In the world of computerized navigation aids each waypoint is specified by exact x, y, and z coordinates. However, in the course of moving towards and past one waypoint, and heading on to the next one, the precise location of the user might never actually coincide exactly with the waypoint, despite having passed pretty much right over it. A computer system might say that the user failed to traverse the path correctly, since she never technically arrived at the penny-sized point. A human observer would, on the other hand, say she was definitely “close enough” to each of the points. This is the concept of *capture radius*. That is, there is a radius around the waypoint that is considered close enough, so that the next beacon sound can appear, and the user can carry on down the next path segment. If the capture radius is too small, the user may *overshoot* the waypoint, and may walk past the corner, off the sidewalk, and into the street. If the capture radius is too large, the person may be told she has reached the turning point too soon, and as a result either cut across the grass or run into a building on the corner. Thus, to keep the person on the intended path—neither missing the marks nor turning too soon—an optimal capture radius needs to be determined. In the work described here we present an update and extension of the previous study, focusing exclusively on capture radius as it affects navigation success with an auditory display.

2. EXPERIMENTAL INVESTIGATION

We set out to study just how precisely the listeners could maneuver along a path using our audio navigation system, and how it was affected by varying the capture radius of the waypoints that defined the path.

2.1. Participants

Undergraduates from the Georgia Institute of Technology participated for partial course credit. All reported normal or corrected-to-normal vision and hearing. There were a total of 108 participants (71 male, 37 female; mean age 20.2, range 18 to 30). Note that a small subset of these data were discussed in [1]. The present data includes data from 69 new participants, and includes an additional level of the capture radius factor (the middle capture radius is entirely new, here).

2.2. Apparatus

As before, our audio navigation studies were conducted using a virtual reality-based testing environment, constructed using the Simple Virtual Environments (SVE) software package developed by the College of Computing at the Georgia Institute of Technology [11]. The beacon sounds were played through closed ear headphones. To change direction participants rotated on the spot where they were sitting. They used two buttons on a joystick to control forward and backward movement in the VR environments (they did not actually walk forward, they were only required to rotate in place). Their orientation within the environment was tracked by an Intersense InertiaCube 2 head-mounted tracking cube attached to the headphones.

Each participant was asked to navigate three different paths in the VR world. The VR environment in which these paths (or *maps*) were located was essentially a large empty (virtual) room with four plain walls and a simple tiled floor (recall, there is no need for any visual fidelity, since the participant only has an auditory interface). In addition to the starting point, Map 1 had five waypoints and Maps 2 and 3 each had 10 waypoints. The three maps differed simply in the layout of the waypoints. The visual rendering of the space, including cubes to mark the waypoints, was available to the experimenter to monitor a participant’s progress through the map. The SVE software logged the participant’s location in the environment (in terms of X, Y, and Z coordinates), head orientation (angular pitch, yaw, and roll), and the waypoint she was currently moving towards, every 2 ms.

Even though it is not a focus of the present discussion (and analyses are collapsed across this variable), it is important to point out that there were different beacon sounds involved in the full study design. The participants were divided into three groups, with each group being guided through the maps by a different navigation beacon sound. The beacon sounds for all three groups were 1 s long, with a center frequency of 1000 Hz and equal loudness. The sounds differed greatly in timbre, however. The first sound beacon was a burst of broadband noise centered on 1 kHz. The second beacon was a pure sine wave with a frequency of 1 kHz. The third beacon sound was a sonar pulse, similar to the sound that Tran el al. [7] found to be one of the best sounds for use as a navigation beacon. Thus, each participant navigated using the same sound throughout their three maps. At the start of a map the beacon sound played in an on-off pattern, where the sound was on for 1 s and off for 1 s of silence. As the listener moved closer to the next waypoint the silence was shortened to effectively make the beacon tempo faster. Hence, increasing proximity to the waypoint was mapped to increasing tempo, which is consistent with our findings for population stereotypes or preferred mappings between proximity and tempo [12].

We were primarily interested in examining how precisely a listener can navigate such an auditory environment with different capture radii for the waypoints. Thus, within eachbeacon-sound group, one third of the participants had a small (50 cm) capture radius, one third had a medium (1.5 m) capture radius, and the final third had a large (15 m) capture radius. Thus, the beacon sound and capture radius were constant throughout the three maps for a given participant. These different beacon radii were chosen based on the results of earlier pilot testing. In practice, as a participant reached a waypoint (or, rather, got close enough to be within the capture radius), a “success” chime sounded and the beacon sound moved in space so as to lead the listener towards the next waypoint in the map. It did not matter from which direction the user approached a waypoint; simply getting within the capture radius was sufficient. So if a participant missed a waypoint, then turned around and came back to the waypoint, the waypoint could be reached successfully from the “wrong” side. Of course, this would mean that the person had moved some extra distance and wasted time, which is reflected in the efficiency and rate calculations for that map. For the large capture radius the participant might actually not come very close to the real waypoint, given the broad region that was “close enough”.

2.3. Procedure

Participants were welcomed and the task was explained in some detail. In particular, the techniques for moving through the map and the concept of a front-back confusion were discussed. A brief questionnaire was administered to gather demographic information and informed consent was obtained. Once the study began, participants moved through the three maps one after the other, with a brief rest between maps. The map order was the same for all participants. Following completion of the third map, the experimenter explained the purpose of the study, answered any questions, and thanked the participant.

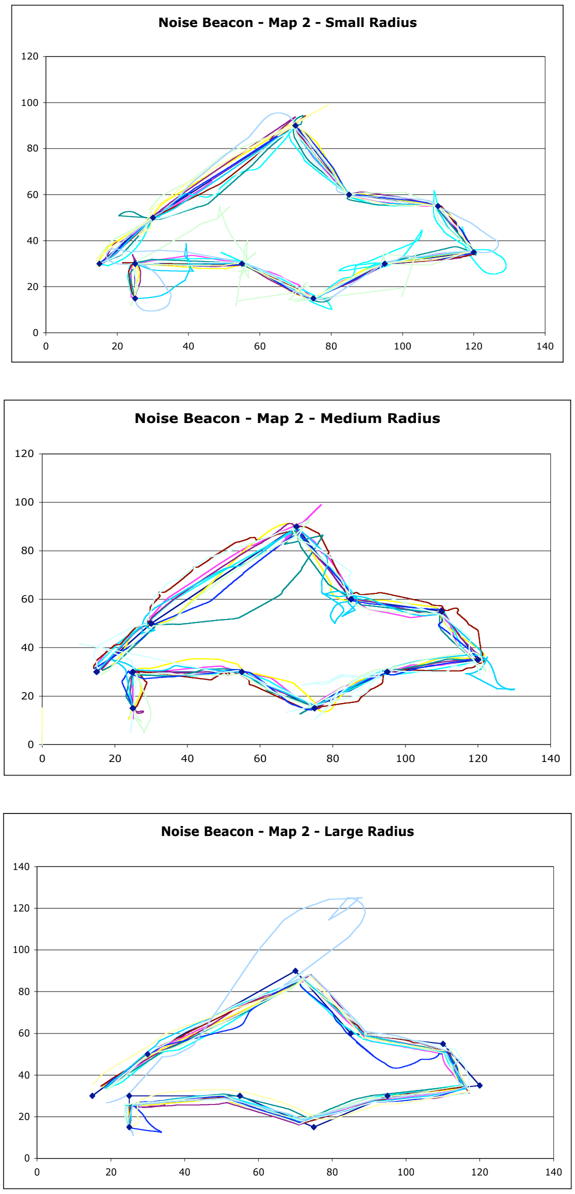

Figure 1. *Movement traces for participants in Map 2, with the noise beacon, and all three capture radius conditions. Note the overall navigation success, and the variability between conditions.*

3. RESULTS

Clearly there are a number of ways to consider the results of such a large study with several variables to consider. For the present paper we will focus only on the findings relevant to our question of capture radius. Full discussion of the comprehensive analysis can be found elsewhere [13]. We first considered the global question of whether participants would be able to complete the navigation tasks using only non-speech auditory cues. Figure 1 presents the movement traces of participants for all three capture radii, when navigating through the second map, and with the noise beacon. These particular results are shown because they are representative and instructive; results for the first and third maps, and for the other beacon sounds are very similar. The straight dark solid line between the waypoints represents the scheduled path, and the other lines in each panel represent the different paths traveled by the participants. The first result to note is the relatively successful navigation through the map by nearly all participants. Not only could participants complete the maps, they picked up the task very quickly and with little instruction. Nevertheless, it is important to note that in some cases there are significant departures from the scheduled path. Most often these result from a participant walking just past a waypoint and not realizing it for some time because the beacon sound is mis-localized as coming from the front instead of from the rear (i.e., an overshoot exacerbated by front-back confusion). This navigation error occurs most often with the smallest capture radius. In the top panel of Figure 1 (the smallest capture radius) several of the waypoints have a “star-like” pattern of movement traces around them. This is the result of a participant overshooting the waypoint, turning around and heading back towards it, then overshooting again. This *hunting* behavior does not appear nearly as often for the medium capture radius, and is very rare for the largest capture radius.

The second result to note is the difference between the general movement patterns in the different capture radius conditions. In the smallest capture radius condition the participants stick very close to the scheduled path, and pass precisely over the waypoints enroute. There is a sort of “pinch point” at each waypoint that is very small for the small radius (naturally). In the larger two capture radius conditions the pinch point is more relaxed, and if a person strays off the scheduled path, he or she need not come exactly back to the path in order to carry on—the capture radius allows some flexibility (or “slop”, depending on one’s perspective) in the path. For the medium radius the participants seem to move off the path in some cases, but still come back to the waypoint.

Finally, in the largest capture radius condition, the participants often never even reach the actual waypoint. They come close enough for the capture radius to be satisfied, but their overall path is actually quite different, geometrically, from the scheduled path. The turning angles are often considerably more or less acute than the angles in the path they were supposed to travel. Certainly the severity of this depends on the context and the reasons for which the person is traversing the path. For practical purposes, the medium capture radius has a compromise between relatively

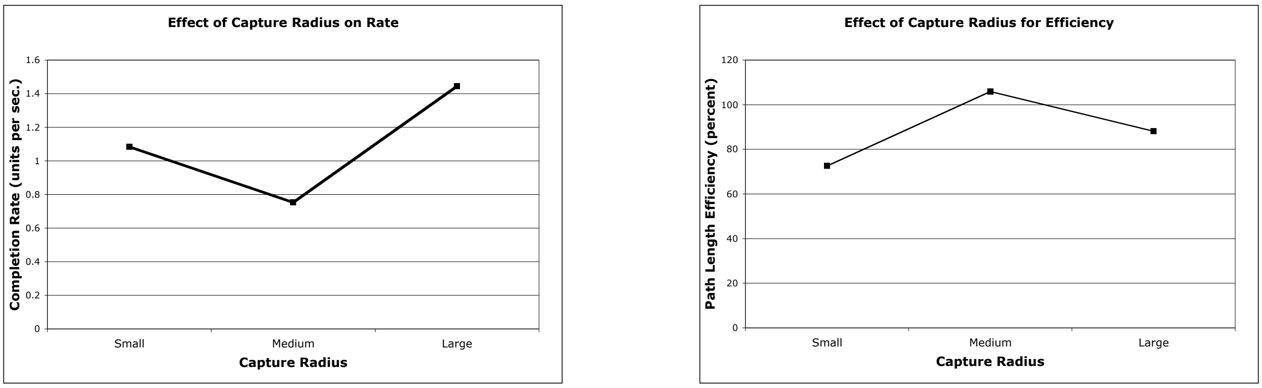

Figure 2. *Effect of capture radius on completion rate and efficiency. The main effect indicates the equivalent of a speed (rate) – accuracy (efficiency) tradeoff in performance.*

little overshooting and hunting, and relatively close passage by the waypoints.

At this point we turned to a more quantitative analysis of the rate of completion and the path length efficiency in the various conditions. We analyzed the data using a mixed factors multivariate analysis of variance (MANOVA). The between-subjects factors were the beacon sound used and the capture radius. The within-subjects factor was map number. The dependent measures being recorded for each map were the participant's overall completion rate (time through the map divided by map length) and their navigation efficiency (total distance traveled divided by scheduled map length). We define the *scheduled map length* as the sum of the lengths of the shortest-distance path segments. That is, if a person moved directly from waypoint to waypoint to waypoint in a map, she would travel a distance equal to the scheduled length. Since participants typically veered off the shortest path somewhat and in some cases overshot the waypoints, the actual distance the participants traveled was usually (but not always) longer than the scheduled map length. The "extra" distance they traveled can be considered as wasted time and effort, and we predicted that comparing the distance the person was supposed to travel to the distance they actually followed would be a useful metric of the movement efficiency afforded by the different beacon sounds and capture radii. This efficiency metric can also be viewed as an indicator of how effective the map might be in guiding a visually impaired person along a specific path (e.g., along a sidewalk). With this in mind, deviation from the path could potentially be very dangerous if there were environmental hazards near the path (e.g. roads, a ditch, etc.). Thus, a priori we assumed an optimal efficiency score would be 100 percent, indicating that for the most part the participant stayed very close to the scheduled map route. In the multivariate analyses we used Wilks' Lambda to determine F values, and throughout all analyses we set an alpha level of .05.

In terms of capture radius, the MANOVA on the combined dependent variables revealed a significant effect of the capture radius being employed, $F(4, 196) = 63.67$, $p < .001$, Wilks' Lambda = .19. This significant multivariate effect led us to seek further clarification in the results for the two dependent variables considered separately.

There was a main effect of capture radius on both rate, $F(2, 99) = 29.07$, $p < .001$, and efficiency, $F(2, 99) = 24.16$, $p < .001$. These results are presented in Figure 2. In the case of rate (Figure 2, left panel), overall the largest capture radius (15 m) yielded the fastest completion rate (1.44), the medium capture radius (1.5 m) led to the slowest rate (0.75), and the smallest capture radius (0.5 m) led to an intermediate rate (1.08). In the case of efficiency, however, the results are quite different (see Figure 2, right panel). The largest capture radius led to a moderate efficiency (88.1%), while the medium capture radius led to the greatest efficiency (105.8%), and the smallest radius led to the lowest efficiency (72.5%). Note that efficiency can be greater than 100% since the implementation of a capture radius makes it possible to traverse a path that is actually shorter than the scheduled map length. Taken together, these two results for rate and efficiency are analogous to a speed-accuracy tradeoff. For example, in the case of the medium capture radius the participants were slow but very efficient. Participants using the large capture radius were fast but inefficient. That is, they spent less time orienting themselves to the beacon sounds, and subsequently traversed a longer path than necessary. However, the large capture radius was very "forgiving", and as a result they were still able to complete the maps quickly. The importance of these various strategies will be discussed shortly.

4. DISCUSSION

There are several important ideas to be drawn from the results presented here. The first and most important is that the non-speech auditory interface can definitely be used for successful navigation. Participants were able to follow the paths in the virtual environment using only the spatialized beacon sounds. Their ability to do so is well illustrated by the traces in Figure 1. Even in the least effective cases, participants strayed relatively little from the path designated by the beacons. This successful performance amongst almost all individuals is a good indicator that the interface can be successful. This is important since the likelihood of simultaneous conversation, use of radio or mobile telephone, or other speech communication points to the need for a non-speech navigation system. In the few cases where a participant's path did deviate significantly from the beacon path, it was most often due to overshooting a beacon by passing just outside its capture radius. Once that happened the participant might have experienced front back confusion and did not turn around to find the beacon because it still sounded as if it was ahead of them. This can lead to a dramatic departure from the planned route, so it must not be dismissed. In debriefing participants it seems clear that some listeners just do not seem to "get" the interface, and never really navigate very well. It may be important to isolate what leads to such confusion with the navigation cues. However, we should be clear that these instances are quite rare, in our experience. Most people pick up the task immediately, show good performance from the start, and improvements with practice. We have considered that the overshoot likelihood is exacerbated by the smaller capture radius. Thus, given that some participants will miss the target waypoint sometimes, we have considered a number of ways to make passing by the waypoint more salient. Studies of a variety of waypoint passing and front-back disambiguation methods are currently underway and will be reported separately.

Next, the effect of capture radius on performance found here appears to be more practically significant than that of beacon sound. Tran et al. [7] investigated the effectiveness of various types of beacon sounds, but did not consider other potentially important factors. Capture radius was also not a main focus of Walker and Lindsay [1]. The results of the present study provide evidence that while sound design should certainly be considered and evaluated carefully, there are additional critical aspects.

Finally, this study also highlights the difference between theoretical and real world considerations. Since for our application we are primarily designing for visually impaired individuals, safety (or remaining on the path) is a paramount concern. There is an obvious speed-accuracy tradeoff occurring between rate and efficiency for different capture radii. Given that fact, and given our primary concern for accuracy, we would first look at the capture radius that led to the best efficiency, and then consider other factors that may affect rate as the situation permits. In a real world application it does not matter if a person using this type of system to navigate down the sidewalk does not move quite as quickly, so long as he or she can manage to remain on the sidewalk throughout the path. Thus, as is often the case, a true human-centered approach must be taken in order to avoid "optimizing" the system at the expense of the user.

Of course, the development of a non-speech auditory interface of this type remains a work in progress. Since such an interface is novel for all users, in addition to the general effects of interface design elements we are beginning to study the effectiveness of different training methods on performance. This includes an evaluation of the basic learnability of different interface sounds and the most effective types of training. Further, it is not clear whether there are individual differences in the perception, understanding, and learning of auditory displays (speech or non-speech), nor how one might predict performance with such a system. Also, to our knowledge, none of the speech-based navigation systems to date involves context- or task-dependent adjustments to the information that is presented. The needs of the listener, within her present acoustical and functional environment, must be factored in so the interface can adapt appropriately. For example, if a user is on target to a waypoint 30 meters down a straight hall, with no obstacles in the way, then the system should stay relatively quiet and let the person use the mobility skills she already has. Approaching the target, the system can gracefully chime in again. A related issue is communicating to the listener the degree of certainty about location, orientation, and items in the surroundings. Knowing that there is some uncertainty in the location (perhaps due to relying solely on GPS) is important for the user in order to adjust attention and other movement techniques. In the present virtual reality-based tests of the interface, the exact location is known, and the listener need not rely on any other sensory input for guidance. Certainly, it will be interesting to see the similarities and differences in effectiveness of the interface when used in an actual movement situation (i.e., not in the VR environment). As the result of pilot studies and our own experience with the outdoor version of our system (which uses a wearable computer), we are confident that the localization of the beacons and the interaction with the system remains similar, and the overall navigation remains robust. We do, however, see some differences in the actual movement style that the users employ. For example, we have noticed that outdoors users tend to walk a little bit more slowly than without the system. However, this effect diminishes with continued usage and increased confidence. Similarly in the virtual environment, participants tend to engage in a type of stop and go pattern of motion that is not necessarily typical of normal movement. Once again this lessens as familiarity with the system increases. Also, it remains to be studied how effectively a user can navigate with an auditory wayfinding system, while at the same time completing other cognitive tasks such as decision making and planning. This multitask proficiency will be important for success of any system aimed at assisting in navigation. Finally, and perhaps most importantly, we are presently conducting several of these listening and navigation studies with low vision and blind participants. Clearly this is a crucial group for comparison with the results reported thus far. As Walker and Lane [14] observed, there are many similarities, but often notable differences in the way participants with visual impairments react and respond to what they hear. This is especially true when it comes to navigation without vision. Preliminary results of these studies seem to indicate that severely visually impaired and blind users, including elderly individuals, are able to use the system at or near the same proficiency level as the sighted undergraduates if they are given sufficient training and practice.

In summary, we have shown the effectiveness of non-speech auditory beacons in guiding a listener along a path, and have highlighted the importance of considering both sound design and other interaction aspects such as capture radius that affect the performance of users in an auditory-only navigation task. Current and future additional studies involving both sighted and visually impaired participants, are also considered.

5. REFERENCES

[1] B. N. Walker and J. Lindsay, "Effect of beacon sounds on navigation performance in a virtual reality environment," Proceedings of Ninth International Conference on Auditory Display ICAD2003, Boston, MA, 2003.

[2] W. De l'Aune, *Legal Blindness and Visual Impairment in the Veteran Population 1990-2025*. Decatur, GA: VA Rehabilitation R&D Center, 2002.

[3] National Center for Veteran Analysis and Statistics, *National Survey of Veterans*: Assistant Secretary for Policy and Planning, Department of Veterans Affairs, Gov't Printing Office, 1994.

[4] G. L. Goodrich, *Growth in a Shrinking Population: 1995-2010*. Palo Alto, CA: Palo Alto Health Care System, 1997.

[5] J. Loomis, R. Golledge, R. Klatzky, J. M. Speigle, and J. Tietz, "Personal guidance system for the visually imapired," Proceedings of First Annual International ACM/SIGCAPH Conference on Assistive Technologies (ASSETS94), Marina del Rey, CA, 1994.

[6] J. Loomis, C. Hebert, and J. G. Cicinelli, "Active localization of virtual sounds," *Journal of the*

Acoustical Society of America, vol. 88, pp. 1757-1764, 1990.

[7] T. V. Tran, T. Letowski, and K. S. Abouchacra, "Evaluation of acoustic beacon characteristics for navigation tasks," *Ergonomics*, vol. 43, pp. 807-827, 2000.

[8] G. H. Mowbray and J. W. Gebhard, "Man's senses as informational channels," in *Human factors in the design and use of control systems*, H. W. Sinaiko, Ed. New York: Dover, 1961, pp. 115-149.

[9] C. D. Wickens, *Engineering psychology and human performance*, 2nd ed. New York, NY: HarperCollins Publishers, 1992.

[10] T. Lokki, M. Grohn, L. Savioja, and T. Takala, "A case study of auditory navigation in virtual acoustic environments," Proceedings of international Conference on Auditory Display (ICAD2000), Atlanta, GA, 2000.

[11] GVU Virtual Environments Group, "SVE Toolkit," vol. 2000, 1997.

[12] B. N. Walker, "Stability of magnitude estimation for auditory data representations," in preparation.

[13] B. N. Walker and J. Lindsay, "Navigation performance with a virtual auditory display: Effects of beacon sound, capture radius, and practice," under review.

[14] B. N. Walker and D. M. Lane, "Psychophysical scaling of sonification mappings: A comparision of visually impaired and sighted listeners," Proceedings of 7th International Conference on Auditory Display, Espoo, Finland, 2001.

***Acknowledgments**. This research was partly funded by a Training Grant from the VA Atlanta Rehabilitation R&D Center.*