Navigation Performance With a Virtual Auditory Display: Effects of Beacon Sound, Capture Radius, and Practice

Bruce N. Walker and **Jeffrey Lindsay,** Georgia Institute of Technology, Atlanta, Georgia

Objective: We examined whether spatialized nonspeech beacons could guide navigation and how sound timbre, waypoint capture radius, and practice affect performance. **Background:** Auditory displays may assist mobility and wayfinding for those with temporary or permanent visual impairment, but they remain understudied. Previous systems have used speech-based interfaces. **Method:** Participants (108 undergraduates) navigated three maps, guided by one of three beacons (pink noise, sonar ping, or 1000-Hz pure tone) spatialized by a virtual reality engine. Dependent measures were efficiency of time and path length. **Results:** Overall navigation was very successful, with significant effects of practice and capture radius, and interactions with beacon sound. Overshooting and subsequent hunting for waypoints was exacerbated for small radius conditions. A human-scale capture radius (1.5 m) and sonar-like beacon yielded the optimal combination for safety and efficiency. **Conclusion:** The selection of beacon sound and capture radius depend on the specific application, including whether speed of travel or adherence to path are of primary concern. Extended use affects sound preferences and quickly leads to improvements in both speed and accuracy. **Application:** These findings should lead to improved wayfinding systems for the visually impaired as well as for first responders (e.g., firefighters) and soldiers.

INTRODUCTION

For a person with vision loss, the two fundamental tasks of navigating through a space and knowing what is around can be a great challenge. At home, difficulties in navigating and learning about the environment can mean diminished mobility and increased danger. At school, visually impaired students may have difficulties in simply getting to class or locating a teacher's office. In the workplace, such difficulties can be an outright impediment to full participation in the corporate or urban culture. Just getting to work can require navigating through a mixture of spaces such as public transit stations, underground malls, city streets, and office buildings. What is a navigational challenge for sighted commuters can be nearly impossible for those with visual impairments.

There are approximately 11.4 million people with vision loss in the United States, 10% of whom have no usable vision, and by 2010 these numbers will nearly double (De l'Aune, 2002; Goodrich,

1997; Leonard, 2002). Because the prevalence of blindness rises steadily with age, as the population of the United States ages there will continue to be more workers with age-related visual impairments resulting from, for example, glaucoma, macular degeneration, and diabetic retinopathy. A great many of these employees can remain very productive even with diminished eyesight, so long as they are able to get to and from work and to move about the office building safely and effectively. Spatial orientation is the major mobility problem encountered by all individuals with profound vision loss (LaGrow & Weessies, 1994; Welsh & Blasch, 1997) but is especially difficult for people whose onset of vision loss occurs later in life (Levy & Gordon, 1988; Welsh & Blasch, 1997). Wayfinding, the ability to find one's way to a destination, is dependent on the ability to remain oriented in the environment in terms of the current location and heading and the direction of a destination. Even highly experienced blind pedestrians exhibit random movement error large enough to

occasionally veer into a wall or into a parallel street when crossing an intersection (Guth & LaDuke, 1995). Although there has been considerable research in the area of electronic travel aids for obstacle avoidance, there has not been comparable research in the development of orientation devices that keep one apprised of both location and heading (Blasch, Wiener, & Welsh, 1997).

It is important to point out that wayfinding can be a challenge for many reasons, not just visual disability. For example, firefighters in a smoke-filled building may not be able to locate the stairwell; military personnel in darkness may not be able to reach a particular rendezvous; police in the midst of a protest may lose orientation because of thick tear gas. Also, even when people can see, during some tasks they may be unable to use vision for navigation because it is required for another concurrent task. It is therefore highly important to develop a system that communicates a range of information about the environment in a nonvisual manner, in order to allow a person more effective knowledge of, connection to, and navigation through the space. Of the candidate alternative display modalities, audition is an obvious choice.

There is strong evidence of the benefits of auditory aids for the blind, in general (e.g., Golledge & Stimpson, 1997; see also Massof, 2003). Most such systems present directions and information via synthesized speech. MoBIC (Petrie et al., 1997) focused on navigation but did not make use of orientation at all. Atlas Speaks® and GPS Talk® from Sendero, LLC, provide a GPS mapping system with information about the user's location and heading, the direction of a particular destination, and limited information about the surroundings (Busboom & May, 1999). For many years, Loomis and his colleagues (e.g., Loomis, Klatzky, & Golledge, 2001) have been developing a Personal Guidance System (PGS) that uses differential GPS and compass data to guide a user along a route. The Drishti system (Helal, Moore, & Ramachandran, 2001) has all of the features of PGS, plus a more sophisticated mapping system. In the more recent systems, the spoken labels are played in spatialized audio, such that the word seems to emanate from the actual location of an object. For example, "doorway here" would sound as if it came from the real doorway.

However, it is also important to consider non-speech audio cues because there are several drawbacks to using exclusively speech sounds. Speech beacons are harder to localize in a virtual environment than are nonspeech beacons (Tran, Letowski, & Abouchacra, 2000). Users also give speech beacons low ratings for quality and acceptance (Tran et al., 2000). 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., Mowbray, 1953; Mowbray & Gebhard, 1961). It is also difficult to use a speech-based interface for navigation and carry on a conversation at the same time (see, e.g., Wickens, 1992). Further, spoken messages in such a system are each generally more than 1 s long. Simpson and Marchioni-Frost (1984, as cited in Stokes, Wickens, & Kite, 1988) pointed out that speech messages are often not understood until the whole phrase is spoken, which may delay perception of urgent messages if speech is used. Also, the length of spoken segments means that the system is often talking. For occasional spoken directions (e.g., "turn left"), this is not a major issue. However, if the system is describing a complex or curving route or is simultaneously presenting other sounds that represent the upcoming curb cut, a low-hanging branch, and the location of a bus stop, the inherent inefficiency of speech can result in a cluttered listening environment (see Stokes et al., 1988, for more on this issue). One final concern is that it simply takes many words to describe nonrectilinear movement: Simply walking toward a sound is easier than translating "57 degrees" into a movement action.

Thus, although speech-based navigation sounds have been useful in some cases, there is a need to understand how to utilize nonspeech sounds as well. Only recently have researchers begun to study nonspeech sounds that can be used in this sort of system. Tran et al. (2000) investigated the effect of sound characteristics on localization and subsequent navigation, the effect of real environments compared with virtual acoustic environments, and the qualitative aspects of various types of acoustic beacons. They summarized their findings by suggesting that the sounds should be wideband and nonspeech, with a proper balance between low- and high-frequency energy to make it pleasant and easy to localize. They also found that a user's rating of the quality of a sound was

highly correlated with localization performance, suggesting that subjective ratings could be a useful metric for initially selecting sounds. Walker and colleagues have also studied the use of nonspeech sounds to convey complex information and have looked at the attributes of the listener (e.g., Walker & Lane, 2001), the design of the sounds (Walker, 2002), and the training given to the listeners (Smith & Walker, 2002).

As part of a larger project to develop a System for Wearable Audio Navigation (SWAN), we have begun to incorporate as much of the existing literature as possible into an auditory display for navigation and environmental awareness. Clearly there remain many questions about the best display design and interaction methods. We report here on the results of the initial studies, in which the factors under investigation included (a) different classes of sounds used as navigation beacons, (b) the effect of varying the capture radius of the navigation system (described shortly), and (c) the effects of some practice with an auditory interface.

BEACONS AND CAPTURE RADIUS IN THE SWAN

The SWAN interface utilizes a repertoire of auditory icons and earcons within a specific framework to allow users to navigate successfully. The nonspeech sounds in SWAN include navigation beacon sounds, object sounds, and surface transitions. They are presented in a 3-D audio environment, with each sound being spatialized to seem as if it were located at the corresponding real-world location. SWAN is able to spatialize these sounds by tracking the user's location and momentary orientation and then applying a dynamic head-related transfer function (HRTF; see also Begault & Wenzel, 1992; Shilling & Shinn-Cunningham, 2002; Wenzel, Arruda, Kistler, & Wightman, 1993).

A complete path that a user might wish to travel is broken down into shorter, straight, unobstructed path segments, joined by waypoints. In order to travel along a path, the user listens for the beacon that seems to emanate from the next waypoint and simply walks toward its apparent location. Once the user reaches the waypoint indicated by the beacon, the sound shifts to represent the location of the next waypoint, the user reorients,

and then the user sets off on the next path segment. As pointed out, Tran et al. (2000) reported that specific sounds can lead to better localization and thus better performance in this sort of task.

Each waypoint is specified by exact coordinates, but the precise location of the user might never exactly "reach" the waypoint's location. That is, a person might never actually step right on top of the exact (x, y) locations that specify a waypoint, despite passing pretty much right over it. This indicates the need for a *capture 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 path. If the capture radius is too small, a person might overshoot the waypoint, walk off the sidewalk, and go into the street. If the radius is too large, the user may think he or she has reached the turning point too soon and may either cut across the grass or run into the side of the building. An optimal capture radius will keep the person close to the intended path while still allowing some flexibility.

EXPERIMENTAL VALIDATION

In studying the effectiveness of the proposed interface, we hypothesized first that users would be able to navigate quite readily with the system and that performance would increase with practice, as is often the case with the use of a new interface. More specifically, improvements should follow a decelerating improvement function (e.g., Lane, 1987; Newell & Rosenbloom, 1981). In terms of what attributes of the display affect navigation performance, we hypothesized that certain types of sounds would lead to more effective navigation, largely because listeners could more easily localize them. This would favor broadband sounds such as noise bursts and complex tones, as compared with, for example, pure tones (Begault & Wenzel, 1992; Middlebrooks & Green, 1991; Shilling & Shinn-Cunningham, 2002). Our final hypothesis was that there would be an optimal capture radius at or about the length of the human stride. If smaller than this, even with excellent sound localization and movement tracking, it would be likely that the user would not step exactly on the waypoint location; if larger than this, it would mean that the sound would shift to the next waypoint while the user was still at least a step away from the turning point.

Method

Participants. A total of 108 undergraduates from the Georgia Institute of Technology participated for credit (71 men, 37 women; mean age 20.2 years, range 18–30 years). All reported normal or corrected-to-normal vision and hearing and completed demographic surveys.

Apparatus. To study a variety of aspects of the audio navigation (SWAN) interface, we have developed a virtual-reality-based prototyping environment. This allows us to implement and rapidly evaluate our sounds, menus, and interaction devices in a safe and controlled lab environment before testing with the full SWAN system. Our virtual environment (VE) was constructed using the Simple Virtual Environments (SVE Version 2.0) software package developed at Georgia Tech (Kessler, Kooper, & Hodges, 1998). SVE, run on a personal computer, managed the tracking of the user's location in the VE and integrated the momentary head orientation as reported by an InterSense InertiaCube2 head tracker (http://www.intersense.com). SVE logged the participant's *x, y,* and *z* coordinates and head pitch, yaw, and roll every 2 ms. The auditory display was rendered in virtual 3-D audio by a Creative SoundBlaster Extigy external sound card, via the OpenAL audio language (http://www.openal.org) and the Direct-Sound3D Application Programming Interfaces (APIs) in Microsoft Windows 2000. The nonindividualized HRTF included in the Extigy hardware was used for spatializing the audio. (Technical specification of the HRTF employed by the Extigy is not available. However, experience with this system indicates that although it may not represent the best HRTF, it is certainly useful in producing spatialized audio; see also Wenzel et al., 1993.) The beacon sounds were played through Sony MDR-7506 closed-ear headphones.

Participants stood in the middle of a quiet testing room wearing the headphones and tracker, holding a modified joystick. To move forward in the virtual space, listeners pulled the joystick trigger (they did not actually walk forward). While the trigger was pressed, the rate of movement through the VE was constant at 1.2 VE distance units/s, which approximates a walking rate of about 4.3 km/hr (2.7 miles/hr; 1 VE unit corresponds to 1 m in the real world). To change direction, the participant rotated on the spot where he or she was standing. Note that changing direction and moving could be done simultaneously. With only brief practice, the users' movement became fluid: They were able to move and adjust orientation in a way that was very similar to moving (walking) in the real world. Other than the method of locomotion, participants who have used both the virtual prototype and the physical SWAN system do not report any major differences between the experiences (for more on this issue, see Farrell et al., 2003).

Each participant was asked to navigate three different maps (i.e., paths defined in the VE) in a fixed order, with the auditory navigation interface. The VE in which these maps were located was essentially a large empty (virtual) room; visual fidelity is not required. The three maps differed simply in the layout of the waypoints and in overall length. In addition to the starting point, Map 1 had 5 waypoints and Maps 2 and 3 each had 10 waypoints, with map lengths of 100.0, 283.5, and 287.6 units, respectively.

Three different beacon sounds were used to guide participants through the maps. Each participant navigated through the three maps using one of the sounds throughout. The beacon sounds were all approximately 250 ms long and equated for loudness, but they differed greatly in timbre. The first sound was a sonar pulse ("ping"), which was included because Tran et al. (2000) found such a sound to be effective as a navigation beacon. The second was a pure sine wave with a frequency of 1000 Hz, intended to have a narrow spectrum. The third was a burst of pink noise, intended to have a broad spectrum but not to sound as harsh as white noise. The actual sound files are available on the first author's Web site (http://sonify.psych.gatech.edu). From spectral characteristics, we hypothesized that the noise burst and sonar pulse would yield the most localizable beacons (Begault & Wenzel, 1992; Middlebrooks & Green, 1991; Shilling & Shinn-Cunningham, 2002). At the start of a map the beacon sound played in a cyclical on-off pattern in which the sound was played, followed by 1 s of silence, then repeated. 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 preferred mappings between proximity and tempo (Walker, 2004).

Finally, one third of the participants had a small capture radius (0.5 m), one third had a medium capture radius (1.5 m) approximately equal to a human stride, and the final third had a large capture radius (15 m). Capture radius remained constant throughout the maps for a given participant. Thus, the overall experiment had a mixed factorial design with the factors of map number, beacon sound, and capture radius.

Procedure. Participants were randomly assigned to the conditions, subject to the constraint of equal numbers in each. The experimenter explained the task and the salient aspects of the interface – namely, that the beacon sounds would be spatialized to indicate the relative direction of the next waypoint and that tempo would be mapped to distance from that waypoint – and discussed potential front-back confusions that can sometimes occur with artificially spatialized sounds and nonindividualized HRTFs (Middlebrooks & Green, 1991; Wenzel et al., 1993; Wightman & Kistler, 1999). Participants then learned how to use a combination of body rotations and joystick button presses in order to move through the environment. Note that participants were free to rotate or tilt their head, which for some helped localize the sound (see also Wightman & Kistler, 1999). Once the study began, participants moved through the three maps in a fixed order, with a brief rest between maps.
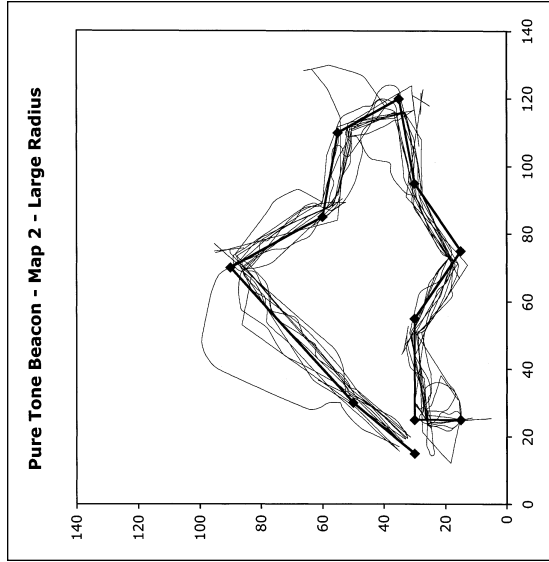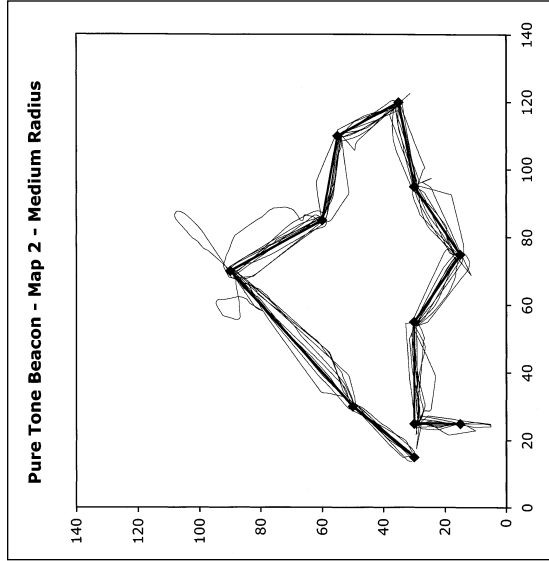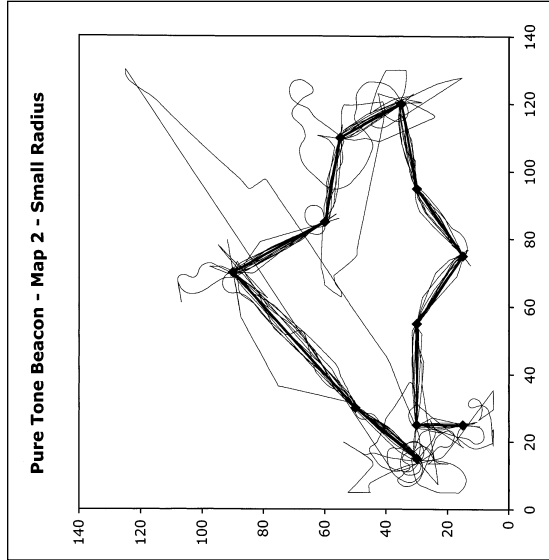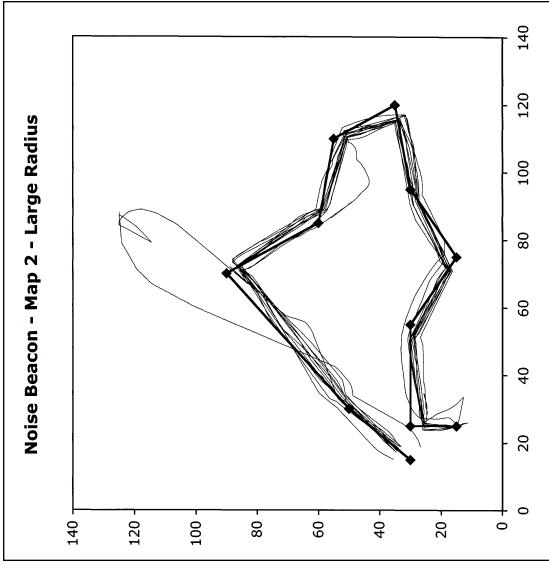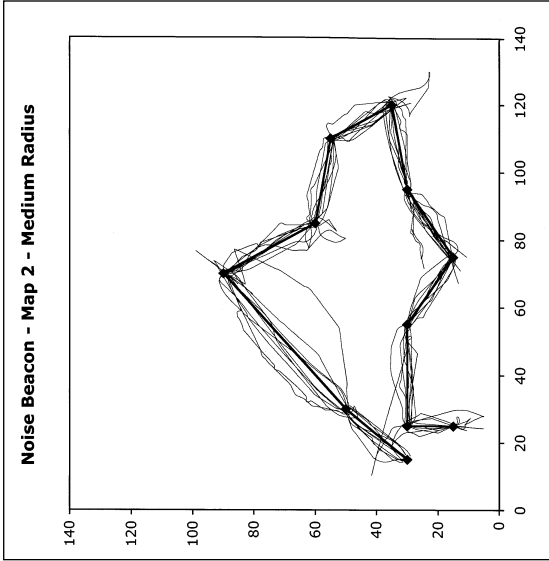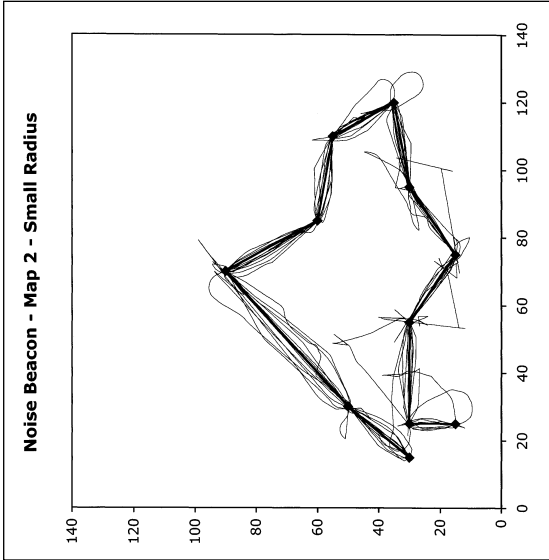
Results

We first considered the global question of whether participants would be able to complete the navigation tasks using only nonspeech auditory cues. Figure 1 presents the movement traces of all participants in Map 2, for each combination of beacon sound and capture radius (results for the first and third maps are very similar to these, and are omitted for brevity). The straight dark solid line between the waypoints represents the scheduled path, and the other lines in each panel represent the actual paths traveled by the participants. The first result to note is the relatively successful navigation through the map by nearly all participants. Nevertheless, it is important to note that in some cases there are substantial departures from the scheduled path. In some cases a star-shaped movement trace around a waypoint indicates overshooting and hunting for the waypoint. The figure

does not depict differences in time to complete the maps, but that is another variable that was considered and will be discussed.

To study both movement path and completion time, we analyzed the data using a three-way mixed factors multivariate analysis of variance (MANOVA). The within-subjects factor was map number. The two between-subjects factors were the beacon sound used and the capture radius. The dependent measures were the participant's *path efficiency* (map length divided by the actual length of the participant's path; thus shorter path lengths are more efficient) and their *time efficiency* (minimum possible time to complete the map divided by the actual time to complete the map; shorter times are more efficient). Because participants typically veered off the shortest path at least sometimes, and in some cases overshot the waypoints, the actual distance the participants traveled was usually (but not always) longer than the scheduled map length. Thus, comparing the distance the person was supposed to travel with the distance they actually moved serves as a useful metric of the movement efficiency afforded by the different beacon sounds and capture radii. This 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), where deviation from the path could potentially be very dangerous.

A path efficiency score near 100% would indicate that the participant had stayed very close to the scheduled map route. Similarly, because there is a minimum time required to complete each map (recall the fixed movement speed and map length), the optimal time efficiency would be 100%. Note that there is usually at least some time spent "standing still" in the map while rotating to get oriented to the beacon, which does not add to path length but does add to completion time (and reduces time efficiency). These two measures are determined from the same performance (hence the MANOVA) but can highlight different aspects of the movement, as is the case for any speed-accuracy trade-off. For example, a person could have a short path but pause frequently, or he or she could have a long and winding actual path but never stop progressing through the map. The latter behavior may be less safe in the real world.

In the multivariate analyses we used Wilks's lambda to determine *F* values, and throughout all

Noise Beacon - Map 2 - Small Radius

Noise Beacon - Map 2 - Medium Radius

Noise Beacon - Map 2 - Large Radius

Pure Tone Beacon - Map 2 - Small Radius

Pure Tone Beacon - Map 2 - Medium Radius
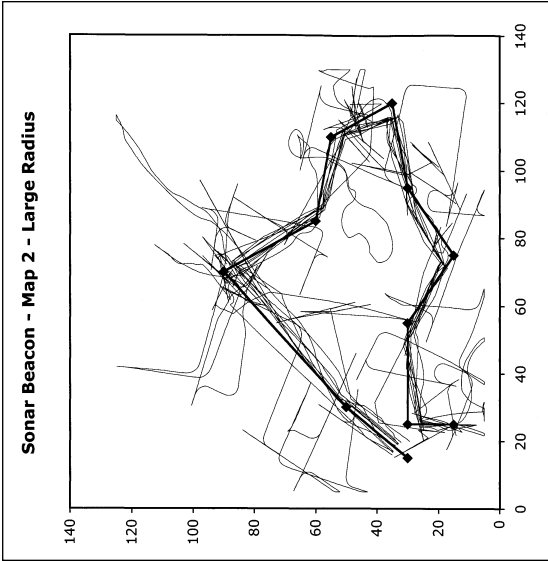
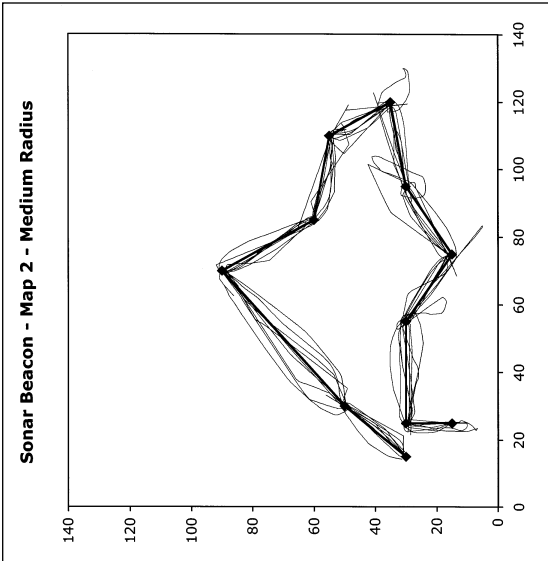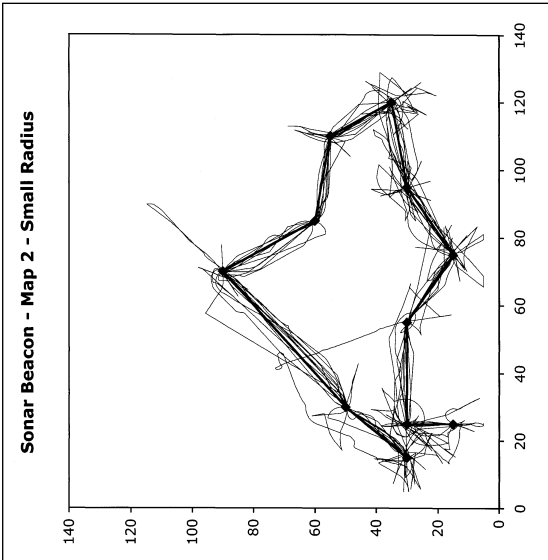Pure Tone Beacon - Map 2 - Large Radius

270

Figure 1. Movement traces for all participants in each combination of beacon sound (noise, pure tone, and sonar sound, in rows) and capture radius (small, medium, and large, in columns) while moving through Map 2. Participants were able to complete the course with little practice and instruction. Some overshoots and bouncing are noted, and this differed across conditions of capture radius and beacon sounds. See text for further explanation.

analyses we set an alpha level of .05. The results of the MANOVA on the combined dependent variables revealed a significant effect of map number, $F(4, 96) = 103.03$, $p < .001$, Wilks's lambda = .19, and a significant effect of capture radius, $F(4, 196) = 63.67$, $p < .001$, Wilks's lambda = .19. There was also a significant multivariate interaction of map number and capture radius, $F(8, 192) = 8.95$, $p < .001$, Wilks's lambda = .53. These significant multivariate effects led us to seek further clarification of the results for the two dependent variables considered separately. Before contemplating the univariate results, we checked

the data for outliers and for violations of the assumptions of normality, linearity, homogeneity, and multicollinearity, with no serious violations noted. We did apply the Greenhouse-Geisser correction to the degrees of freedom in significance tests for effects on path efficiency, in order to correct for violations of sphericity (Mauchley's $W = .822$, $p < .001$). This was unnecessary in the case of time efficiency ($W = .957$, $p = .12$).

As shown in Figure 2, there was a significant increase in time efficiency as participants completed Maps 1, 2, and 3, with means of 63, 103, and 108 percent, respectively, $F(2, 198) = 219.96$, $p <$

Practice Effect for Time Efficiency

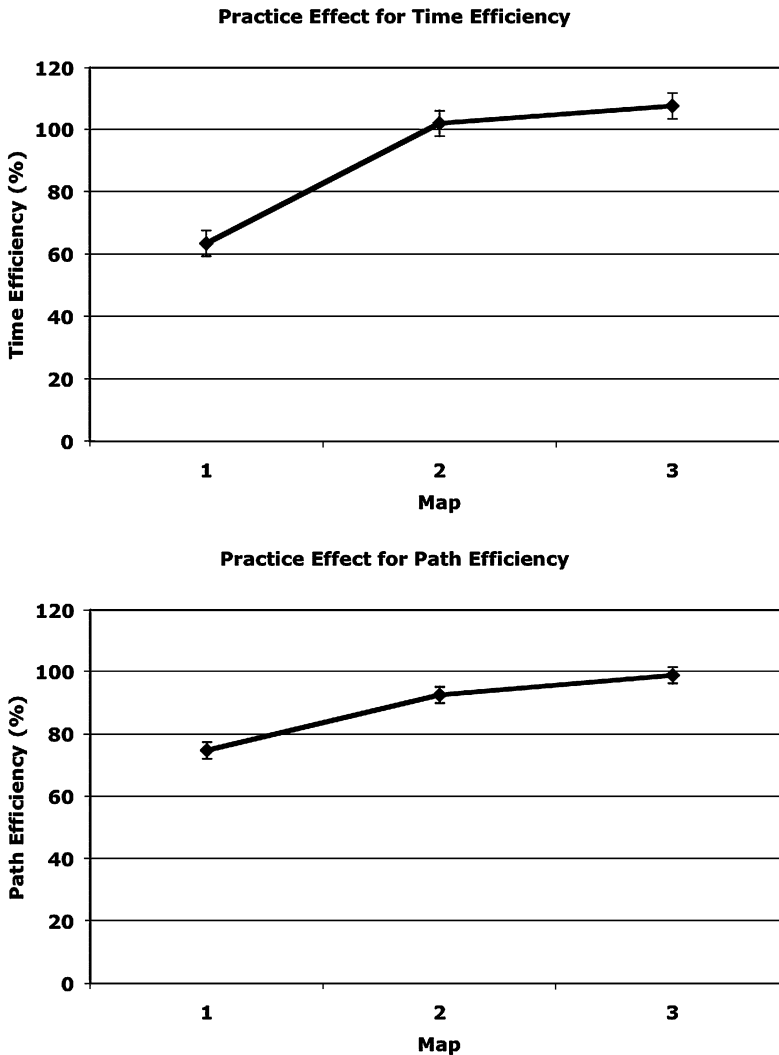**Practice Effect for Path Efficiency**

Figure 2. Main effect of map on time efficiency (top panel) and path efficiency (bottom panel). The significant effects indicate an overall improvement in performance from map to map, namely a practice effect. Note that error bars in all figures represent standard errors.

.001. There was also an increase in path efficiency for subsequent maps, with mean efficiencies of 75, 93, and 99 percent, $F(1.7, 189.8) = 99.99$, $p < .001$. Both of these results reflect an overall main effect of practice.

There was a marginal main effect of beacon sound for path efficiency, $F(2, 99) = 2.54$, $p = .08$, with an ordering in terms of path efficiency of the noise beacon (highest, at 94%), then the pure tone (90%), and finally the sonar ping (83%). The order of performance was the same for time efficiency, namely noise (98%), pure tone (91%), and sonar (83%), though the effect did not reach statistical significance, $F(2, 99) = 1.95$, $p = .15$. There was, however, a significant interaction of beacon sound and map for time efficiency only, $F(4, 198) = 3.10$, $p = .017$. As shown in Figure 3, the noise beacon led to the greatest time efficiencies as well as to the greatest increase in time efficiency with practice. The pure tone and sonar ping beacons led to lower time efficiencies and to less improvement with practice.

There was also a main effect of capture radius on both time efficiency, $F(2, 99) = 29.07$, $p < .001$, and path efficiency, $F(2, 99) = 24.16$, $p < .001$ (see Figure 4). In the case of time efficiency (Figure 4, top panel), overall the largest capture radius yielded the fastest completion (120% time efficiency), the medium capture radius led to the slowest completion (63%), and the smallest capture radius led to an intermediate performance (90%). In the case of path efficiency, however, the results are quite different (see Figure 4, bottom panel). The largest capture radius led to a moderate path efficiency (88%), the medium capture radius led to the greatest efficiency (106%), and the smallest radius led to the lowest efficiency (73%). Taken together, these results show a speed-accuracy trade-off. For example, in the case of the medium capture radius the participants were slow but had short paths. Participants using the large capture radius were fast but somewhat inefficient in terms of path length. That is, they spent less time pausing to orient themselves to the beacon sounds and subsequently traversed a longer path than necessary. However, the large capture radius was very "forgiving," so the participants were still able to complete the maps quickly. The importance of these various strategies will be discussed shortly.

In addition, the main effects of map and capture radius were moderated by a Map × Capture Radius interaction, but only for time efficiency, $F(4, 198) = 9.82$, $p < .001$, not for path efficiency, $F(3.4, 168.0) = 10.3$, $p = .39$. For the sake of comparison, results for the interaction of map and capture radius are shown for both time and path efficiency in Figure 5. For time efficiency
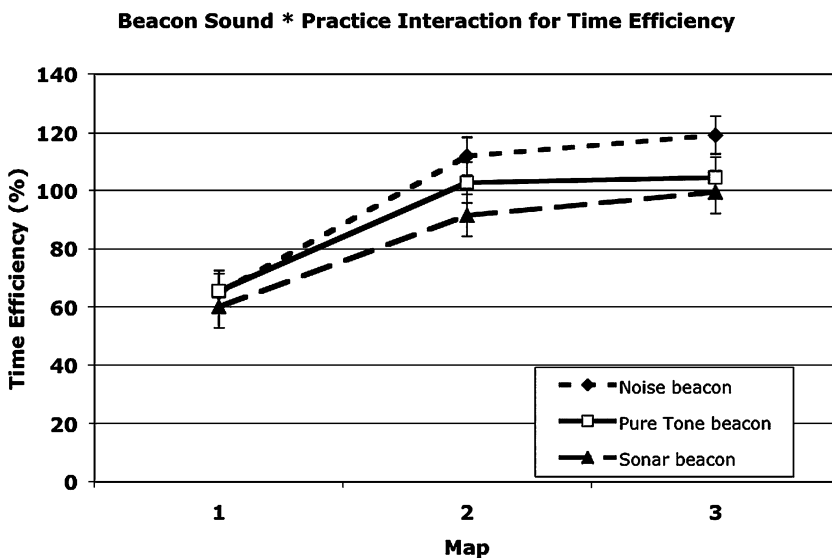

Beacon Sound * Practice Interaction for Time Efficiency

Figure 3. Interaction of beacon sound and map on time efficiency. The noise beacon led to the highest time efficiency as well as to the largest gain in time efficiency across maps. The sonar beacon led to the lowest time efficiency and least improvement, whereas the pure tone beacon led to intermediate results for time efficiency.
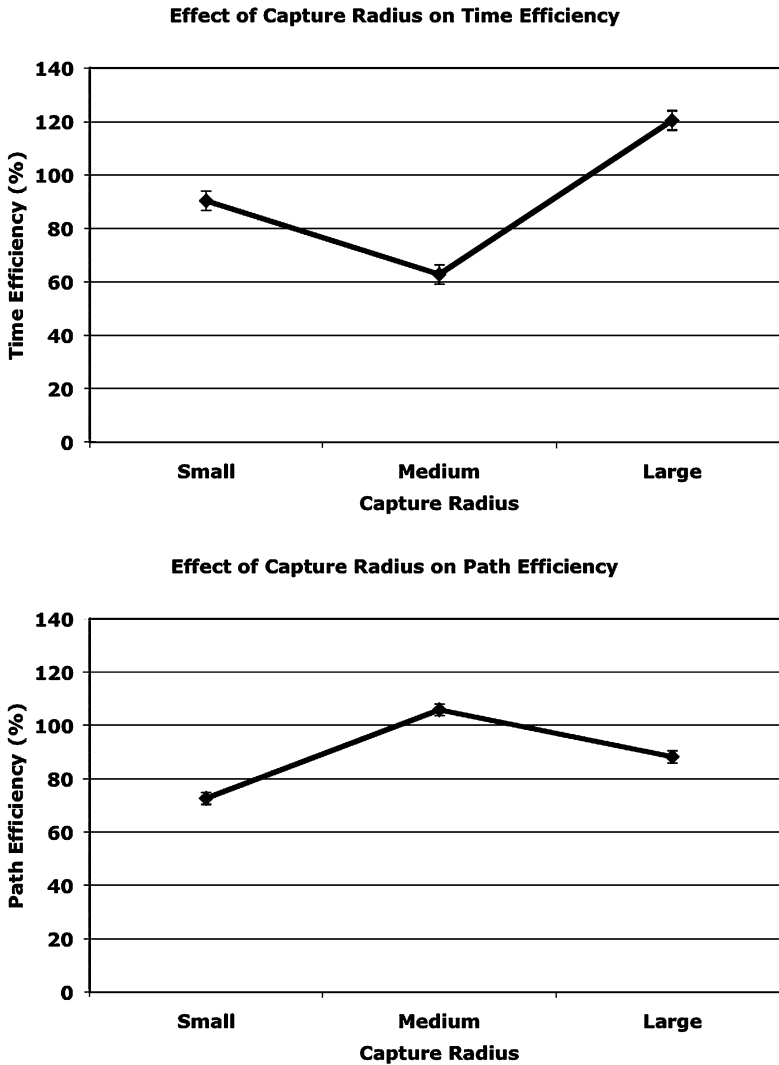
Effect of Capture Radius on Time Efficiency

Effect of Capture Radius on Path Efficiency

Figure 4. Effect of capture radius on time efficiency (top panel) and path efficiency (bottom panel). The main effects indicate a performance trade-off between speed (time efficiency) and accuracy (path efficiency).

(Figure 5, top panel), participants using the largest capture radius started with the highest efficiency and then also improved the most. The medium capture radius led to the lowest time efficiency in Map 1 and to the smallest improvement over the course of the experiment. The smallest capture radius led to an intermediate time efficiency on Map 1 and to an intermediate level of improvement with practice. It is the difference in improvement among the three capture radius groups that leads to the significant interaction. Note, too, that because the large capture radius can lead to a shorter-than-planned path, the completion time can be shorter than the "minimum" time required

to traverse the actual scheduled path. This can explain how the largest capture radius leads to the greatest time efficiency – its benefit is attributable essentially to cutting corners.

Finally, there was an interaction of beacon sound and capture radius for path efficiency only, $F(4, 99) = 2.62$, $p = .04$ (see Figure 6), which comes from the fact that there were differences in path efficiency for the three beacon sounds in the large capture radius condition but no differences among beacon types at the medium and small radius conditions. This indicates that the choice of beacon sound affects path efficiency, but only for the very large capture radius.
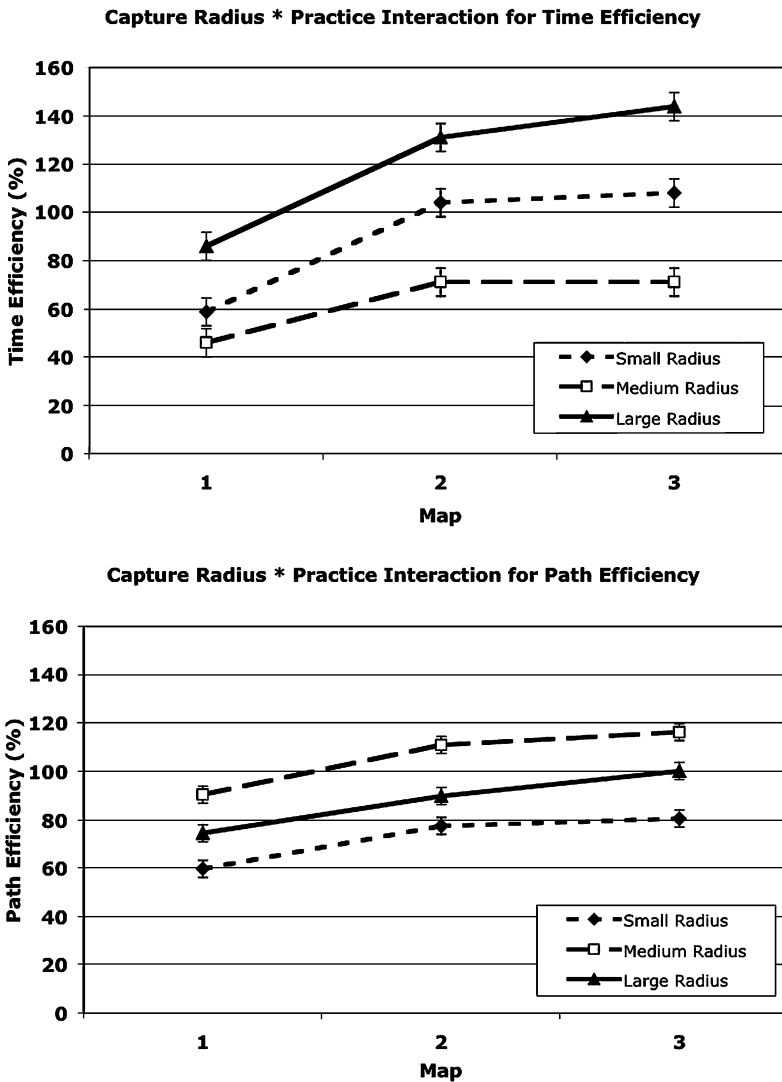
Capture Radius * Practice Interaction for Time Efficiency

Capture Radius * Practice Interaction for Path Efficiency

Figure 5. Interaction of map and capture radius on time efficiency and path efficiency. The interaction reached statistical significance only for time efficiency (top panel), not for path efficiency (bottom panel). The effect for time efficiency indicates a differential practice effect for the three capture radius groups.

DISCUSSION

In terms of the hypotheses set out earlier, the most fundamental conclusion is that the nonspeech auditory interface can definitely be used for successful navigation. Even in the least effective cases most participants strayed relatively little from the path, and all were able to complete the maps. Of course, there were some combinations of beacon sound and capture radius that led to more overshooting and hunting for the waypoints, so completing the map was more challenging and

time consuming. These effects were systematic, though, allowing us to arrive at practical design decisions, discussed in this section. Fortunately, in the physical world the additional navigation cues already present in the acoustic ecology – as well as the additional sensory information from the ground, a cane, or wind – will only make the informational environment richer, leading to even better performance. The next point is that performance improved significantly with practice. Despite exhibiting the predicted slowdown in the rate of improvement, there was still improvement in
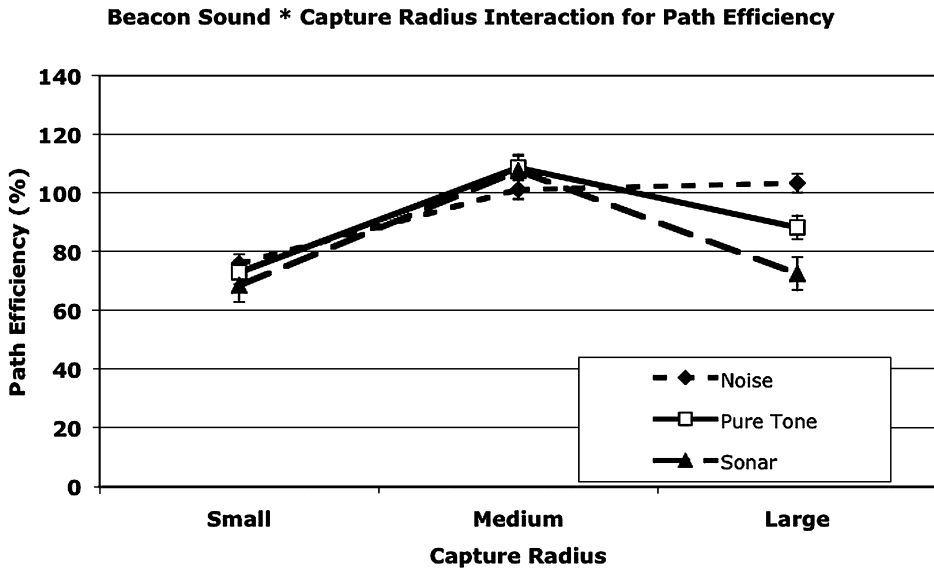
Beacon Sound * Capture Radius Interaction for Path Efficiency

Figure 6. Interaction of beacon sound and capture radius on path efficiency. The three beacons led to different patterns of performance across the three capture radius conditions.

Map 3, suggesting that the limit to gains based on practice was not reached. As is often true when developing any nontraditional interface, further study is needed on the relationship between continued practice and performance.

The next hypothesis was that the broader-spectrum beacon sounds would be more localizable, leading to better navigation. Overall, performance was marginally better with the noise beacon, followed by the pure tone and then the sonar ping. The success of the noise beacon is thus generally as expected, but the relatively good performance of the pure tone (better than the sonar ping, overall), was somewhat surprising. (We have verified that the output of the headphones yields a relatively clean pure tone.) It is clear that performance with this system cannot be predicted solely on the basis of the spectrum of the beacons. Looking more closely at the Beacon Sound × Map interaction (as in Figure 3), it is also interesting to note that time efficiency begins to level off after two maps for the pure tone beacon, whereas performance continues to improve over all three maps for both the noise and sonar beacons. Clearly, then, the results of continued practice need to be studied to assess the longer term performance results. It may be that after continued practice the noise and sonar beacons would outpace the pure tone, as was initially predicted. On the other side of the

same coin, there is also a need for evaluation involving studies outside the lab, to assess the interface in real-world navigation environments.

The final hypothesis was that a capture radius approximately the length of a user's stride would outperform smaller or larger radii. Although there was a speed-accuracy trade-off, the medium radius did, indeed, lead to near-optimal path efficiency. This is clear in both the numerical values (e.g., path efficiency of 106%) and in the movement traces (see Figure 1, middle column). The paths for the smaller and larger radii were numerically inferior, overall, and resulted in either more overshoots and hunting behavior (Figure 1, left column) or in substantial corner cutting (Figure 1, right column). Either of these situations could be dangerous for the user. Granted, there is a trade-off in terms of completion time, but there are two points to be made. First, the time efficiency might be better in some cases, such as the large capture radius, simply because of corner cutting. Thus, it is an essentially unsafe navigation path that leads to faster completion. The second point relates to the broader issue of the difference between theoretical and practical considerations. Given that safety (i.e., staying on the sidewalk) is the greatest concern for us, we are compelled to look first at the conditions that lead to the tightest adherence to the path, and only then consider other factors

that may increase movement rate. 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.

In this light, all things considered, we would conclude that a capture radius of approximately 1.5 m should be optimal for auditory navigation (see Figure 5, bottom panel). With that choice of medium radius, the sonar ping and pure tone led to slightly faster and more efficient paths than did the noise beacon (which performed better only in the large capture radius condition; see Figure 6). In that case, the sonar ping is likely preferred, being a more complex tone and less subject to masking from environmental noises.

In summary, we have shown the effectiveness of nonspeech auditory beacons in guiding listeners along a path and have presented time and path efficiency as useful metrics to be considered jointly when designing and evaluating auditory navigation interfaces. The actual beacon sounds employed in the interface are important to consider, but more so is the way the user interacts with the sounds, such as the implementation of a capture radius about the waypoints on a path. Practice effects need continued evaluations. Applications such as the System for Wearable Audio Navigation (SWAN) need to place path efficiency (accuracy) ahead of speed, in order to maximize safety from the outset. Careful implementations can then begin to provide effective technologies for users who are unable to navigate via vision, leading to enhanced mobility, safety, situational awareness, and quality of life.

ACKNOWLEDGMENTS

REFERENCES

Begault, D. R., & Wenzel, E. M. (1992). Techniques and applications for binaural sound manipulation in human-machine interfaces. *International Journal of Aviation Psychology, 2,* 1–22.

Blasch, B. B., Wiener, W., & Welsh, R. (1997). *Foundations of orientation and mobility* (2nd ed.). New York: American Foundation for the Blind.

Busboom, M., & May, M. (1999). Mobile navigation for the blind. In *Proceedings of the International Conference on Wearable Computing* (pp. 19–23). Vienna, Austria: ISWC.

De l'Aune, W. (2002). *Legal blindness and visual impairment in the veteran population 1990–2025.* Decatur, GA: Veterans Administration Rehabilitation R&D Center.

Farrell, M. J., Arnold, P., Pettifer, S., Adams, J., Graham, T., & MacManamon, M. (2003). Transfer of route learning from virtual to real environments. *Journal of Experimental Psychology: Applied, 9,* 219–227.

Golledge, R., & Stimpson, R. J. (1997). Spatial cognition, cognitive mapping, and cognitive maps. In R. Golledge & R. J. Stimpson (Eds.), *Spatial behavior: A geographic perspective* (pp. 224–266). New York: Guilford Press.

Goodrich, G. L. (1997). *Growth in a shrinking population: 1995–2010.* Palo Alto, CA: Palo Alto Health Care System.

Guth, D., & LaDuke, R. (1995). Veering by blind pedestrians: Individual differences and their implications for instruction. *Journal of Visual Impairment & Blindness, 89,* 28–37.

Helal, A., Moore, S. E., & Ramachandran, B. (2001). Drishti: An integrated navigation system for visually impaired and disabled. In *Proceedings of the Fifth International Symposium on Wearable Computers (ISWC'01)* (pp. 149–156). Zurich, Switzerland: ISWC, Zurich.

LaGrow, S., & Weessies, M. (1994). *Orientation and mobility: Techniques for independence.* Aukland, New Zealand: Royal New Zealand Foundation for the Blind/Dunmore Press.

Kessler, D., Kooper, R., & Hodges, L. F. (1998). *The Simple Virtual Environment Library: Version 2.0 user's guide.* Retrieved 29 March, 2006, from ftp://ftp.cc.gatech.edu/pub/gvu/tr/1998/98-13.pdf

Lane, N. E. (1987). *Skill acquisition rates and patterns: Issues and training implications.* New York: Springer-Verlag.

Leonard, R. (2002). *Statistics on vision impairment: A resource manual.* New York: Arlene R. Gordon Research Institute of Lighthouse International.

Levy, S. B., & Gordon, A. R. (1988). Age-related vision loss: Functional implications and assistive technologies. *International Journal of Technology and Aging, 1,* 16–25.

Loomis, J., Klatzky, R., & Golledge, R. (2001). Navigating without vision: Basic and applied research. *Optometry and Vision Science, 78,* 282–289.

Massof, R. W. (2003). Auditory assistive devices for the blind. In *Proceedings of the International Conference on Auditory Display ICAD2003* (pp. 271–275). Boston, MA: ICAD.

Middlebrooks, J. C., & Green, D. M. (1991). Sound localization by human listeners. *Annual Review of Psychology, 42,* 135–159.

Mowbray, G. H. (1953). Simultaneous vision and audition: The comprehension of prose passages with varying levels of difficulty. *Journal of Experimental Psychology, 46,* 365–372.

Mowbray, G. H., & Gebhard, J. W. (1961). Man's senses as informational channels. In H. W. Sinaiko (Ed.), *Human factors in the design and use of control systems* (pp. 115–149). New York: Dover.

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition.* (pp. 1–55). Hillsdale, NJ: Erlbaum.

Petrie, H., Johnson, V., Strothotte, T., Raab, A., Michel, R., Reichert, L. & Schalt, A. (1997). MoBIC: An aid to increase the independent mobility of blind travellers. *The British Journal of Visual Impairment, 15,* 63–66.

Shilling, R. D., & Shinn-Cunningham, B. (2002). Virtual auditory displays. In K. M. Stanney (Ed.), *Handbook of virtual environments: Design, implementation, and applications.* (pp. 65–92). Mahwah, NJ: Erlbaum.

Smith, D. R., & Walker, B. N. (2002). Tick-marks, axes, and labels: The effects of adding context to auditory graphs. In *Proceedings of the 8th International Conference on Auditory Display* (pp. 362–367). Kyoto, Japan: ICAD.

Stokes, A., Wickens, C. D., & Kite, K. (1988). *Display technology: Human factors concepts.* Warrendale, PA: Society of Automotive Engineers.

Tran, T. V., Letowski, T., & Abouchacra, K. S. (2000). Evaluation of acoustic beacon characteristics for navigation tasks. *Ergonomics, 43,* 807–827.

Walker, B. N. (2002). Magnitude estimation of conceptual data dimensions for use in sonification. *Journal of Experimental Psychology: Applied, 8,* 211–221.

Walker, B. N. (2004). *Consistency of magnitude estimations with conceptual data dimensions used for sonification.* Manuscript submitted for publication.

Walker, B. N., & Lane, D. M. (2001). Psychophysical scaling of sonification mappings: A comparison of visually impaired and sighted listeners. In *Proceedings of the 7th International Conference on Auditory Display* (pp. 90–94). Espoo, Finland: ICAD.

Welsh, R., & Blasch, B. (1997). *Foundations of orientation and mobility* (2nd ed.). New York: American Foundation for the Blind.

Wenzel, E. M., Arruda, M., Kistler, D. J., & Wightman, F. L. (1993). Localization using nonindividualized head-related transfer functions. *Journal of the Acoustical Society of America, 94,* 111–123.

Wickens, C. D. (1992). *Engineering psychology and human performance* (2nd ed.). New York: HarperCollins.

Wightman, F. L., & Kistler, D. J. (1999). Resolution of front-back ambiguity in spatial hearing by listener and source movement. *Journal of the Acoustical Society of America, 105,* 2841–2853.

Bruce N. Walker is an assistant professor in the School of Psychology and the College of Computing at the Georgia Institute of Technology. He received his Ph.D. in psychology in 2001 from Rice University.

Jeffrey Lindsay is a graduate student of engineering psychology at the Georgia Institute of Technology, where he received his B.S. in psychology in 2002.