between the intervals was the timing and level of the echo, calculated relative to the referent click, that accompanied each click. Between trials, clicks were chosen randomly from the 17 clicks described previously.

3.4.3 Procedure

After providing written consent, participants were informed about the structure of the tasks and were given the opportunity to ask questions. They were also instructed not to focus on any one cue in the stimuli and to close their eyes during the experiment. They performed the experiment while seated at a computer in the aforementioned listening chamber. Participants listened to the stimuli through Sennheiser HD 600 headphones.

There were two main tasks: distance and left/right discrimination. All participants performed the distance task first within each of the 15 hour-long test sessions. Each of these tasks was a 2-interval forced choice (2IFC) task in which the order of the intervals was randomized. During the distance task, each interval of a trial contained a referent click, which was centered (no ILDs were applied) and whose channels were normalized relative to each other, followed by an echo generated according to the aforementioned parameters. The angle of echoes in the distance task was kept constant at 0º. Participants discriminated a click with a close echo from a click with a far echo and reported which interval, 1 or 2, contained the closer echo by pressing the corresponding key. During the left/right task, the echo always had a simulated distance of 1 m whereas the angle varied between -90º and +90º at 10-degree increments. In the left/right condition, participants reported whether the echoes in the two intervals moved from right to left or from left to right by pressing the 1 or 2 key, respectively.

During each trial of the 2IFC tasks, participants were shown a sentence reminding them of the correct key presses. (For example, in the distance condition: “Which click contained the closer echo? (1 2)”.) A trial consisted of two clicks separated by 500 ms of silence. For example, a trial in the distance condition could contain a click with an echo from 1 m away, followed by 500 ms of silence, followed by a click with an echo from 4.5 m away. Prerecorded verbal feedback was given (e.g. “correct!”). The next trial was determined using a three-down-one-up staircase paradigm [18]. The staircase paradigm allowed for the determination of a threshold at which participants responded accurately to about 78% of the trials. At the beginning of all three conditions, the staircases started at the easiest level (distance: 5m, lateralization and left/right: 90 degrees). If participants correctly answered three trials in a row, the subsequent trial increased in difficulty by one level. If participants incorrectly answered a single trial, the subsequent trial decreased in difficulty by one level. The track ended after 11 reversals were observed or if the track lasted for over 70 trials without 11 reversals. Here, reversals are defined to be points during the track where participants answered correctly three times after answering incorrectly on the previous trial, or points where participants answered incorrectly once after answering correctly on the previous trial. Participants performed 2-5 tracks per condition per test session. A condition average for each participant was calculated by computing the average of each of that participant’s tracks. If the track did not contain 6 reversals or if the participant performed more than 3 trials at the easiest level during any given reversal in the last 6 reversals, that track was not included in the participant’s threshold calculation. The staircase adjusted the distance (in depth, or angular distance) between the two intervals of a trial and determined a threshold. After each track, performance was reviewed by an experimenter. If the participant’s performance was good, the level of the echo was decreased in the next track. In this way, the experimenter aimed to keep the threshold relatively constant while gradually decreasing the echo level over time.

4. RESULTS OF GAME PLAY VS. PSYCHOACOUSTIC TRAINING

4.1 Learning during training

Using the EchoExplorer™ game, we measured number of crashes into walls per level, number of echoes requested per level, number of steps taken per level, and active time per level. Because echo level did decrease by 2 dB every 15 levels after the tutorial, in order to look for learning effects, we pulled out performance at a few echo levels to compare the number of crashes per level as training went on. The maze level was cycled through a few times so that we were able to compare performance at similar echo/maze levels over time. Figure 1 shows log fits to the average number of crashes per level as a function of the sequential time of play of each level. Crashes were higher initially for the earliest level with the 21 dB boost because it was the first level encountered but, as expected, the asymptotic

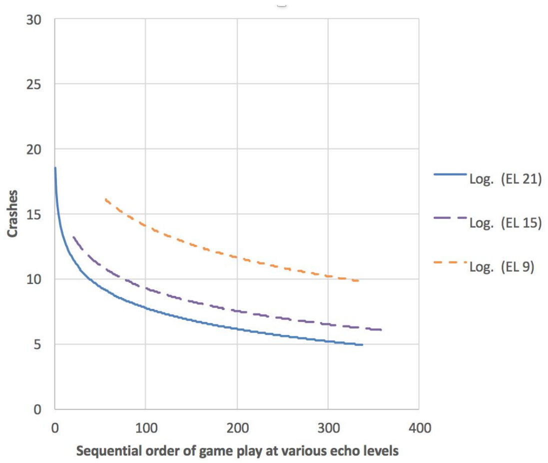

Figure 1. Number of times crashing into maze walls as a function of the sequential order in which the level was completed. The parameter is the level boost of the echo in dB, either 9, 15 or 21 dB. The average data across all App participants are fitted with Log functions.

performance level was best for this condition. Even at the most difficult levels, the echo was still boosted beyond what would typically occur in a real hallway.

There was evidence of learning during the lab training. Across participants in the Lab group, the average simulated distance that supported 78% discrimination from a 1m distance was 4.45 m (SD 0.9). Figure 2 shows the echo level required for discrimination of distance as a function of training hour for four participants. (Varying echo level data were not available for the fifth participant due to a procedural error). Improvement (measured as a decrease in echo level) ranged from 4 to 12 dB over time. Across participants, the average simulated angle difference that supported 78% correct discrimination between right and left lateral positions was 38.4 degrees (SD 9.9). Figure 3 shows the echo level required for discrimination of lateral position (in the left/right task) for all 5 participants in the Lab group. In both graphs, the echo level required decreased over time. The echo was still boosted beyond what would typically occur in natural conditions, but by the end, all participants could reliably discriminate the echo when its level was lower than the source level.

4.2 Improvement in echolocation

Discrimination of lateral position and distance was evaluated in both a pretest and posttest. Dprime was the measure of sensitivity to the different possible locations of the board [19]. The four lateral positions (-90, -45, 45, and 90) yielded a chance level of 25% which would be equivalent to a d' of 0 for a four-alternative forced-choice task. The three possible distances (0.9, 1.8 and 2.7 meters) yielded a chance level of 33% with a d' of 0 for a three-alternative forced-choice task.

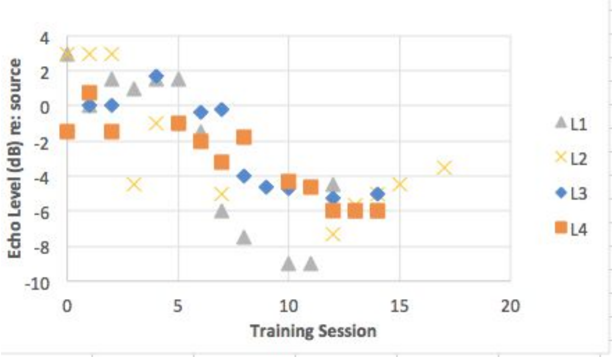

Figure 2. Echo level used relative to the initial outgoing click in order to support average threshold performance on distance discrimination as a function of hours of training in the lab (for individuals in the Lab group).

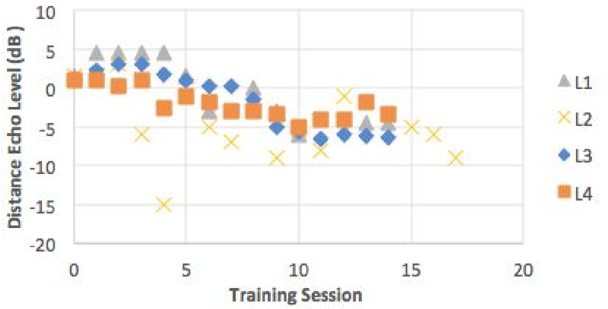

Figure 3. Echo level used relative to the initial outgoing click in order to support average threshold discrimination of left/right lateral position as a function of hours of training in the lab (for the Lab group).

When head position was fixed, there was modest sensitivity to lateral position with an average d' of 0.10 at pretest and 0.43 at posttest, with 9 of 11 trained participants showing improvement. When the head was moved freely, the average lateral position discrimination was 0.15 at pretest and 0.64 at posttest, with 10 of 11 participants showing improvement. For the head moving condition, post test d' versus pretest d' is shown for all observers in Figure 4. No change in performance would be implied by the dashed line, whereas improvement is indicated by all data points above that line. The average sensitivity was relatively low given that a d' value of 1 is typically considered threshold discrimination, similar to the 78% correct

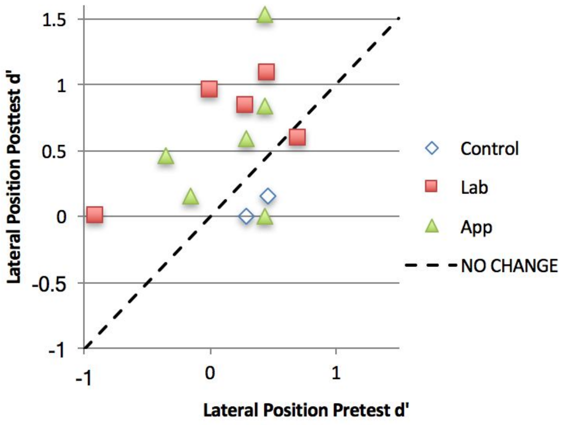

Figure 4. Post test d’ (sensitivity) to lateral position as a function of Pre test d’ in a real-world discrimination task when the head was moved freely. Control participants (open diamonds) did not have any training between tests. Lab participants (red squares) and App participants (green triangles) were trained in the lab, or used the app, respectively, for 15 hours between tests. Data points above the dashed line indicate improvement.

value that we targeted in the lab training. Nonetheless, the average improvement in d’ (0.30 across conditions, 0.11 for fixed, 0.49 for moving) was reliably greater than zero (tested by a mixed ANOVA using a between subject factor of training group and within-subject factors of fixed vs moving head and pre vs post test (F(1,4)=63, p<.001). There no significant main effect of the type of training (Lab and App groups) and no effect of the fixed and moving head conditions, nor were there any interactions. Note that the condition in which the head moved also benefited from more observations because it came after the fixed condition, so we would not draw conclusions about head movement *per se* from this result. Among the 44% of post-test lateral position trials in which trained participants changed their answers between the fixed and head moving conditions, they changed from an incorrect to a correct answer 42.6% of the time (whereas chance would be 33%). The two control participants displayed pretest dprimes within the middle of the range of all other participants and did not improve their sensitivity between pre and post tests.

Average discrimination of distance was low in the 3afc task, but reliably above chance (average of d’=0.215, with 95% confidence intervals for all conditions above zero). There was no learning of distance between pre and post tests as indicated by an ANOVA with factors of training group, pre/post test and fixed/moving head. Average distance d’ did not improve between pretest (0.22) and posttest (.21) and this did not interact with training group or head movement. Data were missing from one participant in the lab group in the fixed head distance discrimination; the resulting App group did have a lower average d’ (0.12) than the Lab group (0.31) (F(1,8)=6.13, p<.038), but because distance discrimination did not significantly change at post-test, this group difference was inconsequential for characterizing training effects. Dprime for distance was not reliably lower when the head was fixed (0.19) compared to moving (0.24) nor did it interact with any other factors.

5. CONCLUSION AND DISCUSSION

At the outset, there were many factors working against the possibility that our first attempt to train people with this game would produce measurable echolocation benefits in the real world. Among these factors were the fact that this was a beta test of an app; our participants were sighted; the training conditions did not match the test conditions in terms of the reflector locations; the referent clicks were not the participant’s own clicks; the HRTFs were not emitted by the participant; and the training echoes were louder than they are in real life. For this reason, we find the small reliable benefits that we measured to be encouraging.

If sighted users can gain some small benefit from this training, the next step after further refinement would be to test blind users. At this point, we can conclude that training with artificial sounds - sounds that do not require the participant to emit clicks or move - can be effective in improving echolocation with self-generated clicks in the real world. At the start of this study it was not known whether a game such as ours could produce a benefit.. Note that this new finding of transfer is distinct from showing that training on echolocation in the real world improves echolocation in the real world. Even so, locating a board while moving one’s head is only one small step towards using echolocation to navigate while moving in a realistic environment. Although it is likely that our artificial training may be less effective than if the same amount of time were spent navigating in the real world while exclusively using mouth clicks, both of our training methods have potential advantages, especially when many hours of solo training are required. Our methods allow the use of enhanced echoes, and our data show that this enhancement was required to achieve reliable discrimination at the start for all participants. We saw that learning in the lab using a standard psychoacoustic method was not substantially superior to learning with the app. However, a key advantage to training with the game is that it is under the user's control and can be used at their convenience; it therefore is more accessible and practical than the customized psychoacoustic training method in the lab. It should also be pointed out that the EchoExplorer™ game was tested as a beta version and is not yet optimized. Therefore, we find these experimental results encouraging for the future of using games to learn new ways to use sound for navigation.

6. ACKNOWLEDGMENT

We thank Sarah Kwan, Tejal Kudav, Kiran Matharu, Jacqueline Hon, Jessica Kwon, Aaron Schwartz, Chieko Asakawa, Catherine Getchell, participants at the Blind and Vision Rehabilitation Services of Pittsburgh, and Art Rizzino for helpful discussions, feedback and pointers. We also thank Google Inc., the NSF Center for Science of Information (CSoI), CMU's SURG program, and the NSF REU program for their generous support.

7. REFERENCES

[1] A. J. Kolarik, S. Cirstea, S. Pardhan, and B. C. J. Moore, "A summary of research investigating echolocation abilities of blind and sighted humans," *Hearing Research*, vol. 310, pp. 60-68, 2014.

[2] B. F. G. Katz and L. Picinali, "Spatial audio applied to research with the Blind.," in *Advances in Sound Localization*, 2011, pp. 225–250.

[3] D. Pelegrin-Garcia, M. Rychtáriková, C. Glorieux, and B. F. G. Katz, "Interactive auralization of self-generated oral sounds in virtual acoustic environments for research in human echolocation," in *Proceedings of Forum Acusticum 2014*, 2014.

[4] H. Wallach, E.B. Newman, and M. R. Rosenweig, "The precedence effect in sound localization," *The American Journal of Psychology*, vol. 62(3), 315-336, 1949.

[5] K. Saberi, and J. V. Antonio, "Precedence-effect thresholds for a population of untrained listeners as a function of stimulus intensity and interclick interval," *J. Acoustical Soc. Am.*, vol. 114, pp. 420-429, 2003.

[6] L. Wallmeier, N. Geßele, and L. Wiegrebe, "Echolocation versus echo suppression in humans," *Proceedings of the Royal Society B.,* vol. 280, 20131428, 2013.

[7] B. C. J. Moore, *An Introduction to the Psychology of Hearing, Sixth Edition.* Bingley, UK: Emerald Group Publishing Limited, 2012.

[8] S. Teng, and D. Whitney, "The acuity of echolocation: Spatial resolution in the sighted compared to expert performance," *Journal of Visual Impairment and Blindness*, vol. 105(1), pp. 20-32, 2011.

[9] B. N. Schenkman, and M. E. Nilsson, "Human echolocation: Blind and sighted persons' ability to detect sounds recorded in the presence of a reflecting object," *Perception*, vol. 39, pp. 483-501, 2010.

[10] A. Tonelli, L. Brayda, and M. Gori, "Depth echolocation learnt by novice sighted people," *PLoS One*, vol. 11, no. 6, pp. 1–14, 2016.

[11] L. Thaler, "Echolocation may have real-life advantages for blind people: an analysis of survey data," *Frontiers in Psychology,* 2013, vol. 4(98), eCollection.

[12] W. Wu, R. Morina, A. Schenker, A. Gotsis, H. Chivukula, M. Gardner, F. Liu, S. Barton, S. Woyach, B. Sinopoli, P. Grover, and L. M. Heller, "EchoExplorer™: A game app for understanding echolocation and learning to navigate using echo cues," in *ICAD Proceedings 2017,* 2017.

[13] C. S. Green, D. Bavelier, "Action video game training for cognitive enhancement," *Current Opinion in Behavioral Sciences*, vol. 4, pp. 103-108, 2015.

[14] R.V. Algazi, R.O. Duda, D.M. Thompson, C. Avendano, The CIPIC HRTF Database, *IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics* In WASSAP '01 (2001).

[15] E. De Sena, N. Kaplanis, P.A. Naylor, T. van Waterschoot, "Large-scale auralised sound localisation expe riment," *AES 60th International Conference*, 2016.

[16] L. S. R. Simon, N. Zacharov, and B. F. G. Katz, "Perceptual attributes for the comparison of head-related transfer functions," *J. Acoust. Soc. Am.*, vol. 140, no. 5, pp. 3623–3632, 2016.

[17] J. A. M. Rojas, J. A. Hermosilla, R. S. Montero, and P. L. L. Espi, "Physical Analysis of Several Organic Signals for Human Echolocation: Oral Vacuum Pulses," *Acta Acustica United with Acustica,* vol. 95, pp. 325-330, 2009.

[18] M. R. Leek, "Adaptive procedures in psychophysical research," *Perception & Psychophysics,* vol. 63(8), pp. 1279-1292, 2001.

[19] N.A. Macmillan and C.D. Creelman, Detection Theory: A Users Guide, 2nd Ed. Psychology Press, 2004.