Using Dynamic Audio Feedback to Support Peripersonal Reaching in Young Visually Impaired People

Graham Wilson & Stephen A. Brewster
Glasgow Interactive Systems Group
School of Computing Science, University of Glasgow, G12 8RZ UK
{first.last}@glasgow.ac.uk

ABSTRACT

Blind children engage with their immediate environment much less than sighted children, particularly through self-initiated movement or exploration. Research has suggested that providing dynamic feedback about the environment and the child's actions within/against it may help to encourage reaching activity and support spatial cognitive learning. This paper investigated whether the accuracy of peripersonal reaching (space within arm's reach) can be improved by the use of dynamic sound from both the objects to reach for and the reaching hand itself (via a worn speaker). We ran two studies that tested the efficacy of static and dynamic audio feedback designs with blind and visually impaired young people, to identify optimal feedback designs. Study 1 was with young adults aged 18 to 22 and Study 2 involved children aged 12 to 17. The results showed that dynamic audio feedback helps to build spatial connections between the objects and the reaching hand and participants were able to reach more accurately, compared to unchanging feedback.

Keywords

Sound perception; reaching; visual impairment.

1. INTRODUCTION

Children who are congenitally or early blind can be less engaged with objects in their immediate environment [4,22] due to a lack of location awareness and a slower cognitive development of object existence/permanence [27]. Sound and touch can be used to inform the child of the object's existence/location and encourage him or her to reach for it [14]. However, these activities typically require a parent or caregiver to facilitate the interaction, by moving the object, touching the child or making sounds. Providing a means with which children could, of their own accord, learn of the existence of objects, and their own position relative to them, might encourage more self-initiated movement [22,27]. Sounding objects have been used several times (e.g., [11,29,34]) but they require adult activation. A computer-based system that can control the playing of environmental sounds based on the child's activity could provide more complex and engaging feedback to encourage the child to be "more active against the world" [27]. As these children get older, even into adulthood, environmental sound may also be a way to support the development of accurate reaching within peripersonal space (i.e., space within arm's reach [19,23]).

The ABBI project [12] is developing technologies and procedures to (re)habilitate spatial cognition in visually impaired children through natural audio-motor associations. The primary focus is developing the ABBI bracelet, which is placed on the wrists of children (and adults) with visual impairments: it detects movement (acceleration and orientation) and produces sound based on the nature of that movement. Thus it uses the auditory modality to convey spatial information about the movement of the person's own body within personal and peripersonal space, in a natural and direct manner similar to feedback provided by the visual modality in sighted people. These sound sources may then allow the visually impaired child to build a better representation of his/her movement in space and, ultimately, a representation of space itself.

The research in this paper investigated extending the audio-motor association beyond the limb to build associations with objects in the environment. Environmental sounds can encourage reaching [22,29,34] and wrist-based sound can improve spatial movements [8,12]. Therefore, this research looked at whether the combination of environmental sounds and wrist-based sound can guide and improve peripersonal reaching in blind and visually impaired people. The paper describes two studies that measured reaching accuracy and got subjective responses about the feedback design, including which was preferred and whether it established a connection between the hand and the reached-for object. The studies were similar other than the age of the participants taking part. Access to young (< 5) blind children is very limited, as is the amount of time that it is possible to engage them in experimental situations [2]. Older children and young adults can provide more detailed and reasoned feedback than young children to the the design process. Therefore, as a first step, Study 1 was run with blind and visually impaired young adults (aged 18 to 22) to determine which of three audio designs best supports reaching. The results showed a benefit of dynamic audio feedback but no clear advantage for either sound design and so we continued to Study 2 involving blind and visually impaired children aged 12 to 17.

2. RELATED RESEARCH

2.1 Reaching in Peripersonal Space

In a series of experiments, Brungart and colleagues looked at the localisation of sound sources within 1 metre of the head [6,7]. Blindfolded sighted participants sat in an anechoic chamber and placed their chin on a rest to immobilise the head. The experimenter placed a sound source on the end of a curved tube and manually placed the source in positions around the participant's head from 0° to 180° azimuth, -90° to +90° elevation and distances of ~15cm to 1m, all relative to the chin. Participants closed their eyes and, after the sound played, they used the end of a 30cm wooden rod to point to the perceived sound location.

They first tested localisation of white noise (200Hz – 15kHz) pulses (5 x 150ms, separated by 30ms) with amplitude-based distance cues removed [6]. Azimuth error averaged at 12.6°, with error decreasing as distance increased. Azimuth error was higher

ASSETS '16, October 23 - 26, 2016, Reno, NV, USA

ACM 978-1-4503-4124-0/16/10…$15.00
DOI: http://dx.doi.org/10.1145/2982142.2982160

for targets at lower elevations (below head-level) and more so for near distances. Average elevation error was 11.3°, with higher error in front of the participant, than at the sides. As distance cues were removed, distance error was unsurprisingly high (30-40%). A follow-up study, using the same experimental setup, tested the effects of different stimulus characteristics on localisation [7]. They used the same fixed amplitude to provide distance cues and compared the same previous white noise stimulus with two filtered versions (low and high pass) and a monaural stimulus, by blocking the sound in one ear. The smallest azimuth error (12.2°) and elevation error (10.5°) came from the fixed-amplitude original broadband noise stimulus, with the low pass filter resulting in the highest azimuth error (14.7°) and both filtered version having equal elevation error (~15.5°). Providing distance cues improved distance perception but azimuth error decreased as distance increased. These results impact the use of sound in peripersonal reaching, as 1) they show that the sound design will impact localisation and 2) objects are commonly nearby and below head-level, so their sound will be more difficult to localize than distant sources, or those nearer head-level.

Brungart *et al.* [6,7] gave participants a wooden rod with which to point to the perceived location, but we are interested in reaching with an open hand, as that is how children will reach for and interact with objects, and research suggests the method used to indicate perceived direction influences accuracy [16]. One of few studies to examine the accuracy of pointing with the finger in reaching space was conducted by Macé *et al.* [19]. Blind and blindfolded sighted participants sat in front of a semi-circular horizontal array of loudspeakers, arranged in 5 columns of 5 speakers placed at 30° intervals through 120°, starting straight in front. They studied the effect of varying the length and number of white noise (20Hz to 20kHz) bursts: single bursts of 10, 25, 50 or 200ms and multiple bursts of 2, 3 or 4 25ms bursts (separated by 30ms silence). Following stimulus presentation, the participant was to point to the perceived location and then return to facing straight ahead. They found that blind participants were significantly more accurate along the azimuth, although only by 0.5°, but were significantly *less* accurate in judging distance, but again only by ~10mm. Azimuth and distance errors decreased as the length of single bursts increased, but all multiple burst stimuli were more accurately localised than single burst stimuli, although only significantly so compared to the shortest single burst. Multi-burst error sat at around 10.5-11°. This suggests that multi-burst stimuli may better support reaching, over single sounds.

Parseihian *et al.* [23] extended the study by Macé *et al.* [19] looking at the effect of dominant vs. non-dominant hand and using physical touching (with tip of the fingers) instead of pointing to indicate the perceived location of the sound. This is important, as the physical contact is a more direct correlate of reaching for objects. They found no effect of reaching hand on azimuth or distance errors. They found that, potentially due to biomechanical limitations, placing the hand at nearby objects (33cm) was difficult and so led to significantly worse azimuth judgements, compared to distant objects. This may also be related to the poorer localisation of sounds nearer the head [6,7].

2.2 Sound Localisation

Reaching accuracy is inherently influenced by the perceived location of the sound source. We are primarily interested in the accuracy of the reaching movement itself (we do not measure localisation alone) and a full discussion of localisation research is outside the scope of this paper. Therefore, we present a discussion only of particularly relevant research, focusing on the effects of sound parameters on localisation, as we are interested in identifying suitable audio designs to support reaching. Localisation accuracy is typically measured by the ‘minimum audible angle’ (MAA) [24] or ‘localisation blur’: the smallest perceivable change in location. Localisation in sighted people is more accurate straight ahead (and behind) compared to directly at the sides [3]. In a series of experiments, Hartmann & Rakerd looked at various stimulus (and room) effects on localisation of sound sources [17,25]. Removing attack transients (very slow onsets) led to only chance-level localisation accuracy for 200Hz and 500Hz sine tones, but broadband noise stimuli were better localised. In general, the greater the spectral complexity of a stimulus, the better localised it will be [17].

Having shown that slow onsets disrupt localisation, Rakerd & Hartmann [25] identified the threshold onset length for poor localisation as somewhere between 100ms and 500ms, depending on room acoustics. Accuracy improves as onset reduces to 0ms. They also found no effect of stimulus duration, and little effect of stimulus offset on localisation. Perrott [24] also found that the MAA increased as duration increased from 1ms to 500ms, but only for 2kHz stimuli, not low (500Hz) or high (5kHz) stimuli. However, research is currently split on whether blind and visually impaired people have better or worse auditory localisation than sighted people [15,19,26]. Macé *et al.* [19] found that longer stimulus durations led to better pointing accuracy, although they used white noise, rather than the sine tones used in these other studies. Overall, the research suggests that sounds we use should have fast onsets (<=100ms) and should avoid simple sine tones.

2.3 Environmental Engagement

Congenitally or early blind children develop motor skills more slowly than sighted peers and can be less engaged with objects in their immediate environment [4,22] due to a lack of awareness of their locations and a slower cognitive development of object existence or permanence [27]. Millar [22] provides recommendations on how to support a blind child’s understanding of, and interaction with, the space around them. She stresses that simple audible ‘lures’ are an insufficient substitute for the visual feedback that attracts attention in sighted children. Instead, she suggests, “what is needed…is some means whereby sounds can be *systematically connected* with more than one other source of information” (Ch. 9, emphasis added). Hearing should be connected with reaching in the same way that vision is. The ABBI bracelet can provide information about the position and movement of the arm [12,13], but dynamic feedback coming from the arm and/or external objects will provide information about 1) arm location relative to the person, 2) object location and 3) the proximity of the hand to the object. This may add some redundancy of information, but this is recommended, particularly for young children [22]. Millar also suggests some ways to encourage reaching, including placing objects nearby that make noises when touched.

Ross & Tobin [27] also discuss the need for more complex sounds than simple ‘lures’. They recommend that an infant be provided with access to changes in sound over time, sounds that inform the child about the object’s existence or action. They use the word “Flux” to refer to these changes and that perceptual information comes from how objects change over time. In particular, to counter a blind child’s lack of self-initiated action, they recommend the adoption of a rehabilitative program that forces the child to be “more active against the world”. The earlier children are engaged in rehabilitation, the more benefit is gained from them [27].

The Royal National Institute of Blind People (RNIB) in the UK recommends the use of “sensory resource boxes” [28], wooden structures that hang objects with different sounds and textures for a child lying underneath to interact with (see Figure 1). The RNIB guide says these boxes "motivate children to notice the changes in their environment and then to begin to explore these changes". The boxes initially rely on the accidental contact of the child's limb with the objects, or the parent/carer moving the child's hand into contact with them. After this the child may remember the location to deliberately reach for the objects. The ABBI device can inform the child of the location of his/her arm, and a sensory resource box augmented with sounding objects could use the spatialised sound to inform the child of the location of objects they might enjoy interacting with. Tracking hand position relative to the objects could mean that the sounds could provide more information, support or encouragement for reaching.

Figure 1: Blind children reaching for hanging objects in a "sensory resource box" (L) and for sounding chains (R)

The research in this paper sought to design a system that *systematically connects* [22] a person's movement to objects nearby that, when applied to children, might encourage them to be "active against the world". It is "systematic" as the feedback was designed to vary in predictable and logical ways, and the feedback designs were tested with target user groups to ensure they were useful. The dynamic changes in feedback based on proximity are aimed to "connect" the hand and the object and encourage action.

3. EXPERIMENTAL APPARATUS

The apparatus consisted of 7 small speakers connected to a Windows 8.1 PC via USB soundcards and a Microsoft Kinect v2 depth camera, which tracked hand and speaker positions. Six of the speakers (KitSound [18] Mini Buddy "Magic 8 Ball") were used as the target objects and a variant of the same model was used for the wrist speaker ("Owl") (see Figure 2). The back surface houses the speaker (which always faced the participants) with the following characteristics: 100Hz to 20kHz frequency response, 36mm driver, S/N of 85dB and output of 2W RMS @ 4 ohms. Hand tracking was achieved using Kinect's in-built API with a One-Euro Filter [10] applied for smoothing.

Figure 2: Speakers used as targets ("8 Ball", left) and placed on the wrist ("Owl", right). Images from [18].

An ABBI device was not used at this time because the dynamic audio feedback required was not yet incorporated into the device. The KitSound speakers were chosen for several reasons. The ABBI device is deliberately designed to be small, light and cheap to produce, to maximise its validity and usefulness in clinical and home settings with children as young as several months. The use of large or high quality speakers would not reflect the capabilities of the ABBI device, nor likely those of sounding toys, so the results may not be transferrable. These KitSound speakers are similar in size and weight to ABBI so can be held on the wrist with a strap, and they are representative of the intended final sound capabilities. They also allowed us to test more complex feedback in preparation for that functionality being added. While the speakers were connected to audio cables, the cable length and low speaker weight (50g) meant there was no movement restriction or encumbrance. As we are working with blind and visually impaired children, these speakers were also chosen because they include fun and colourful designs (for low-vision), and have different tactile qualities, which would provide a constant base across qualitative/quantitative research opportunities.

The "Magic 8 Ball" was chosen for its smooth and uniform shape, to provide no distractions, obstacles or haptic cues during reaching. The "Owl" was chosen as the protruding 'ears' provided purchase for the elastic bands holding the speaker to a rubber watchstrap and the 'feet' helped keep the otherwise spherical speaker from rotating during use, which would result in the speaker facing the wrong direction. The 6 target speakers were placed on top of 10cm high plastic beakers, to give them greater size and have them at a height suitable for reaching (see Figure 3).

Figure 3: Experimental setup (left), speaker on wrist (right)

4. FEEDBACK DESIGNS

The purpose of the research was to determine which of several potential sound designs best supports accurate reaching and best conveys an *interconnection* between the hand and the object being reached for. While both are desirable, a design that only accomplishes one could still be of benefit in helping a child (or adult) relate his/her own movements to available objects nearby. The two studies presented below looked at both *dynamic* audio designs, which change based on the proximity of the hand to objects, but also constant, unchanging feedback. This was to establish if dynamic feedback, which would require more complex and expensive computing hardware in order to track the hands and object locations, provides any extra benefit in reaching accuracy over more basic feedback, which could be provided more simply and cheaply by a continually playing toy.

Three feedback designs were compared: 1) a dynamic Geiger counter, 2) a dynamic pitch and 3) a constant (unchanging) design that remained the same regardless of proximity. A Geiger counter is a common design that has been used successfully in navigation [21], and changes in pitch have strong effects on perceptual *streaming* [5]. Also, changes in frequency (pitch) and tempo (similar to Geiger counter) are commonly associated with changes in size [30] (in our case distance changes in size). We wanted to measure the effect of feedback emanating from the object alone, the wrist alone and the object and wrist combined, and so each feedback design had two aspects: an *Individual*, single speaker design and a two-speaker *Coalescent* design.

Both designs, but particularly the *Coalescent* design, were based on perceptual streaming [5], where sounds are perceived to be from one or multiple sources. Alternating notes that are close together in time or pitch are perceived as coming from the same source, while more distant notes are separated into multiple streams [5]. Therefore, the feedback designs changed in a way that, at a distance, the hand and object may be perceived as separate, but as proximity increases, they may be perceived as a single source, increasing the association of hand to object. The designs provided multiple interconnected pieces of information [22] and may provide "Flux" information [27].

For the interaction to be accepted by children, the sounds need to be enjoyable, or at least pleasant. A workshop with blind and visually impaired children revealed that synthetic sounds were as preferable as natural sounds, and were more easily tracked in 3D space [33]. To minimise experimental complexity as much as possible, all the sound designs were made using 0.2s synthetic "pluck" tones (similar to a guitar string) of different pitch generated from the free Audacity application [1]. They are composed of a single pitch (e.g., 523.25Hz) but with a non-sinusoidal waveform which, in combination with an onset of 0ms and rapid decay, should increase localisation accuracy over pure sine tones [25]. Regardless of the condition, the target speaker always played C^5 (523.25Hz) to provide a constant reference. The target speaker was always indicated by a burst of five C^5 tones, one every 200ms. A burst was used over single tones to increase localisation accuracy, as per Mace *et al*. [19].

It should be noted that ABBI habilitation activities are deliberately short and periodic, to help children develop their own skills, rather than replacing them or providing a crutch on which they depend. Therefore, the feedback designs were intended to be simple, minimalistic and so (potentially) easily understandable.

4.1 Geiger Counter

In this design, a pluck tone of pitch C^5 was played at an increasing rate/rapidity as the hand approached the object. For the *Individual* speaker design, the sound was produced by either the object or wrist, depending on the condition. Research has suggested that there is an inherent association between increasing the rapidity of sounds and decreasing proximity [32]. Therefore, the feedback began when the hand was 50cm from the target, with a tone playing every 900ms. As the hand approached the object, inter-tone delay decreased by 100ms for every 5cm advance, to a minimum delay of 100ms (10 per second) when the hand was within 7.5cm of the target centre (see Table 1). The two-speaker *Coalescent* design was similar, but the tones alternated between playing on the object and the wrist. The sound from the object and wrist speakers changed in rapidity (based on proximity) the same way as the *individual* design, but the tones were played in alternation, up to a maximum rapidity of 50ms (100ms alternated). To increase the perceptual distinction [35] between object and wrist, the object played C^5 and the wrist played G^4 (392Hz) and both tones played together when touching the target.

4.2 Increasing Pitch

Bregman [5] suggested that rapid notes close in pitch are grouped into a single stream, while notes distant in pitch are perceived as two separate streams. Therefore, the *Individual* design played the 8 tones in a C major scale that increased in pitch from C^4 (261.63Hz) up to C^5 as proximity increased (Table 1). A discrete mapping was chosen over a continuous function to provide perceptually clear changes but also to provide potentially more pleasant feedback, using harmonious musical steps and feedback that was not constantly changing. The notes were played at a constant rate of 200ms. The mapping of *increasing* pitch (notes) to a decrease in distance is based on previous research that found it more in line with mental models of size compared to *decreasing* pitch for decreasing size (distance) [30]. For the *Coalescent* design, both the object and wrist speakers played tones at 100ms rapidity, but played alternately, resulting in an overall frequency of 50ms. In this case, the target always played the highest C^5 tone and the wrist tone increased based on proximity as in the *Individual* design, so that the tones matched when the target was touched.

Table 1: The mapping of Geiger inter-tone delay (ITD) and Pitch to the distance of the hand to the centre of the target

Distance (cm)	**50**	**45**	**40**	**35**	**30**	**25**	**20**	**15**	**7.5**
ITD (ms)	900	800	700	600	500	400	300	200	100
Pitch Note	C^4	C^4	D^4	E^4	F^4	G^4	A^4	B^4	C^5

4.3 Constant Design

The constant audio design was a continual repetition of the target object sound throughout the entire reaching action: a C^5 tone repeated constantly every 200ms. It either played from the target object, the wrist or both the object and the wrist. When both the object and wrist played, the tones were alternated at 100ms, like the Geiger counter and the wrist played G^4 (392Hz).

5. STUDY 1: REACHING IN ADULTS

As access to blind children is very limited we first used blind and visually impaired young adults, with the intention that the best sound design identified during the study would be used in future research with children. Six participants aged 18 to 22 (mean = 18.8) took part (1 female). All were legally blind: four totally blind, one had minimal light vision (Leber congenital amaurosis) and one had retinal dysfunction (the latter two wore a blindfold). Participants took part in all conditions and were paid £10.

The experiment had a 3 x 3 (*Feedback Design x Sound Source)* within-subjects design, giving 9 conditions, but a *Control* condition was added, where only the initial target sound was played (no other signals during the reaching action). Participants completed these 10 conditions over 60 minutes in a random order. Within each condition, each object speaker was the "target" to be reached for 3 times in a random order, following 6 practice trials. At the start of each trial the "target" object was indicated by five C^5 pluck tones played every 200ms. The participant was then free to start reaching for the target with the right hand. They were free to touch multiple objects in their hunt for the perceived target but were instructed to be as fast and accurate as possible. To "select" a speaker, the participant was instructed to place the palm of their right hand on top of it (see Figure 3) and press the space bar on a keyboard with the left hand.

5.1 Speaker Positions

Due to the number of conditions, two different speaker layouts were used to minimise the likelihood that participants would learn the positions, and so reach for a known location, rather than the location of the sound. The two layouts of the targets relative to the participant are shown in Figure 4. As it was not possible to fully counterbalance speaker position with the experimental conditions, speaker layouts were alternated between successive conditions.

Research has suggested that reaching accuracy is similar for the dominant and non-dominant hands [23] and so, to limit the length of the experiment, only the right hand was tested. The participant sat at the left-edge of a desk, facing forward. As distance and relative position can influence peripersonal pointing accuracy

[6,7,23] the speakers were placed so that they varied in both, while reachable with an out-stretched arm. In both layouts, two speakers were each placed on an arc 55cm, 61cm and 67cm from the participant. The azimuth angles for each speaker in each layout, relative to speaker 1 at 0° straight ahead, were 7°, 14°, 36°/31° (layout 1/layout 2), 44°/50° and 52° for speakers 2, 3, 4, 5 and 6 respectively. This had the effect of varying the azimuth angle between speakers, to determine if this had an effect. Distances of 50cm+ were chosen due to the difficulty in localising [23] and reaching for [6,7] nearer sounds.

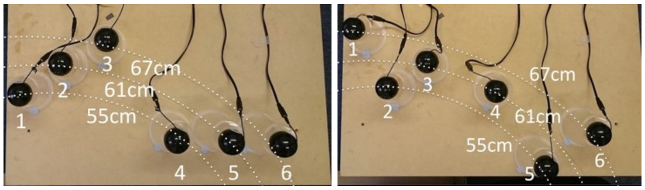

Figure 4: Target speaker layouts 1 (left) and 2 (right), including distance arcs. Participant sat facing speaker 1.

5.2 Variables & Measures

The Independent Variables were *Feedback Design* (Geiger, Pitch and Constant), *Sound Source* (Object, Wrist and Both), *Speaker Layout* (1 and 2) and *Speaker Position* (1 to 6). The Dependent Variables were: *Accuracy* (whether the correct speaker was chosen) and *Task Time* (time between the "target" sound finishing and participant selection, in milliseconds). We also recorded the participants' *Movement Trace* (all their hand movements during the trial in metres) to measure if different audio designs supported more direct and less exploratory reaching. This was done by summing the Euclidean distances between all Kinect hand positions. After each *Feedback Design* we asked three questions rated on 7-point Likert scales from "Strongly Disagree" (1) to "Strongly Agree" (7): *"The sounds created a connection between the hand and the object"*, *"The combination of the hand and object speakers was beneficial"*, *"The sound was pleasant"*. We also asked for *Feedback Design* preferences. While subjective views are important, Walker *et al.* suggest that experimentally derived mappings are better than only using "what sounds good" [30,31].

5.3 Results

5.3.1 Accuracy

No significant learning effect on accuracy was observed ($F_{(2, 538)}$ = 0.65, p > 0.05) and so the accuracy data were analysed using a 3 x 3 repeated-measures ANOVA and is summarised in Figure 5. There was a significant main effect of *Feedback Design* ($F_{(2, 214)}$ = 13.838, p < 0.001), Bonferroni pairwise comparisons showed that the *Constant* design had a significantly higher error rate (24.1%) than both the *Geiger* counter (12.0%) and *Pitch* (12.7%) designs. There was no significant main effect of the *Sound Source*: the *Wrist* speaker resulted in 18.5% error, *Both* speakers were 14.5%, and the *Object* speaker was 15.7%. There was no significant interaction effect between *Feedback Design* and *Sound Source*.

Each individual *Feedback x Sound Source* condition was compared to the *Control* condition, to look at the effect of adding feedback. A one-way repeated-measures ANOVA with the 10 conditions was run, followed by Bonferroni-corrected paired comparisons. There was a significant effect of Condition ($F_{(9, 963)}$ = 7.672, p < 0.001): the *Control* condition resulted in significantly more errors (36.1%) than all *Dynamic* feedback conditions except *Geiger + Wrist*. There were no differences between the *Control* condition and the *Constant* conditions. A 6 x 2 repeated-measures ANOVA found no significant effect of either *Speaker Position* (6) or *Speaker Layout* (2) and no interaction effect. Based on the pattern of selection errors and the azimuth positions of the speakers, the average angular error was 11.4°, with means for the two layouts of 15.5% (layout 1) and 13.3% (layout 2).

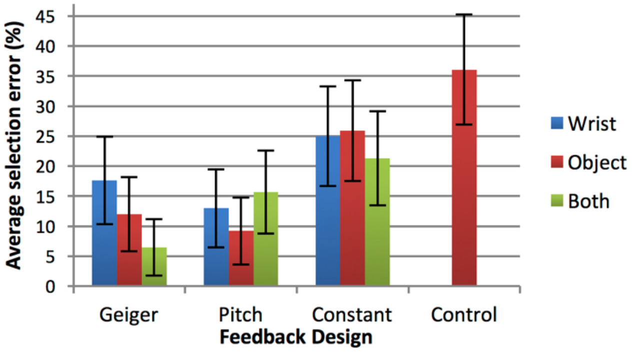

Figure 5: Mean Study 1 target selection error for all Feedback Designs and Speaker Locations. Error bars = 95% CI.

5.3.2 Selection Time

The selection time data violated the normality assumption, and so non-parametric analyses were carried out. Two separate Friedman tests were run to analyse *Feedback Design* and *Sound Source*. Following a significant effect, Bonferroni-corrected Wilcoxon pairwise comparisons were carried out (the necessary value for significance was p = 0.05/3 = 0.0167). A summary of the data can be seen in Figure 6. There was a significant effect of *Feedback Design* on reaching time (χ^2 (2)=13.16, p=0.001), with the *Pitch* design being significantly faster (mean = 3237ms), than both the *Constant* (3488ms) and *Geiger* (4103ms) designs.

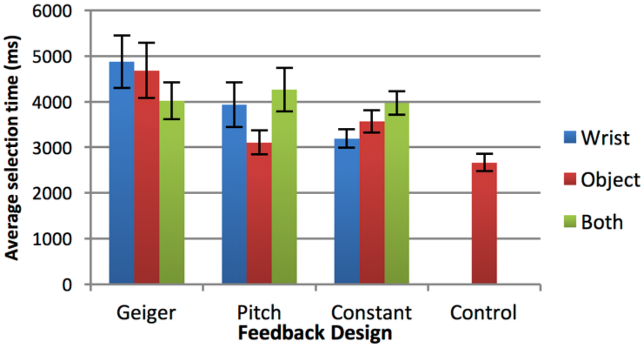

Figure 6: Mean Study 1 target selection times for all Feedback Designs and Speaker Locations. Error bars = 95% CI.

There was also a significant effect of *Sound Source* on reaching time (χ^2 (2)=22.80, p<0.001). Sound from the *Object* alone resulted in significantly faster times (mean = 3784ms) than sound from *Both* speakers (4079ms). The *Wrist* condition had a mean time of 4011ms, between the other two. Similar to the accuracy data, we compared movement time during the *Control* condition to the other conditions. The *Control* condition resulted in significantly faster selection times (χ^2 (9)=106.43, p<0.001; 2666ms) than all other conditions, except for *Pitch* with the *Wrist Speaker*.

A Friedman test found a significant effect of *Speaker Position* on Movement Time (χ^2 (5)=13.45, p<0.05), as Position 1 took significantly longer (mean = 4237ms) than Positions 4 (3705ms) and 5 (3697ms). A Wilcoxon comparison found no significant effect of *Speaker Layout*: mean Times were 3936ms for Layout 1 and 3619ms for Layout 2.

5.3.3 Movement Traces

We looked at the total distance the reaching hand travelled during each trial (in metres) to see if any feedback design led to more direct movement towards the target (a shorter distance would suggest less searching behaviour). As the data were not normally distributed, we used Friedman tests in the same way as for the time data. There was a significant main effect of *Feedback Design* on the total movement distance (χ^2 (2)=15.63, p<0.001). The Constant feedback resulted in significantly less movement (mean = 1.23m) than both the *Geiger* (1.62m) and *Pitch* (1.56m) designs. There was no main effect of *Sound Source,* with means of 1.52m, 1.45m and 1.67m for the *Wrist*, *Object* and *Both*, respectively.

5.3.4 Subjective Responses

In response to the statement *"The sounds created a connection between the hand and the object"*, the average response was 6 for Geiger and 6.3 for Pitch (between "Agree" and "Strongly Agree") but only 4.7 for the Constant design (between "Neutral" and "Slightly Agree"). Based on the average responses to *"The combination of the hand and object speakers was beneficial"*, providing sounds from both the Wrist and Object did not really provide any benefit over one speaker alone. All ratings were <4, between "Slightly Disagree" and "Neutral". Two participants said expressly that providing sounds from both was unnecessary, as sound from only one (particularly the Object) was sufficient. One participant also commented that having two different Pitches in the Coalescent design made it more difficult to process the feedback. However, two participants said having the two sounds meant that comparing the two could help to guide the hand. Two participants expressed that having sounds coming only from the wrist made reaching more difficult than having it from the Object, or having sounds from Both. All Feedback Designs were "pleasant" (> 4), with the Dynamic designs having slightly higher ratings.

When participants were asked which of the audio designs they preferred, one expressed no preference, one preferred the Geiger counter, three preferred the Pitch design and one expressed preference for different aspects of both the Geiger and Pitch designs. When asked why they chose these, the participant that preferred the Geiger counter, and the one who preferred aspects of both, said it was because they thought the changes in sound (based on proximity) were more "intuitive" or "instinctual" than the changes in pitch, as people who are not musical may struggle to make use of the pitch changes. In contrast, one of the participants who preferred the Pitch feedback believed *it* to be the more "intuitive" design, while the other two preferred it because the changes were more obvious and so easier to distinguish, making it more helpful.

The participant who preferred aspects of both designs said that he appreciated the immediate feedback in the Pitch design, as the Geiger counter inter-tone delay is long (900ms) when the hand is far away, so it requires more time to hear any feedback, compared to the Pitch design. This may be why Pitch was significantly faster than the Geiger design. He is the participant who stated that having two Coalescent Pitches made them more difficult to process, which may have been exacerbated by his difficulty in making use of Pitch changes in general. Finally, two participants said that dynamic changes were "necessary" in order to help reaching accuracy, as the constant feedback was not helpful.

5.4 Initial Discussion

Both Dynamic feedback designs led to greater accuracy and more positive subjective responses than the Constant design, and Pitch-based feedback was also significantly faster than the other designs. Constant feedback provided no benefit in accuracy over providing only a brief cue (Control), despite the Constant condition taking significantly longer than the Control. We interpret the shorter movement distance and faster movement time for the Constant/Control conditions as meaning that, in the absence of useful feedback, participants are left with only their initial location perception to rely on, and so make direct movements to the perceived (but more often incorrect) source.

Comparing the two Dynamic designs, the Pitch design had a significantly faster Movement Time than the Geiger counter, but there was no significant difference in accuracy, suggesting that participants were more hesitant in moving using the Geiger counter, potentially due to the long initial ITD in the Geiger, as the lack of a difference in reaching distance suggests they were not searching or investigating during this time. The subjective responses were more strongly in favour of the Dynamic designs, based on the higher average ratings of 6-6.3 compared to 4.7 for the Constant design, as well as individual comments from participants, such as the perceived *requirement* of dynamic changes to reach accurately. Therefore, the results from Study 1 suggested that dynamic audio feedback may be useful in helping blind and visually impaired children accurately reach for nearby objects. The reasons given for preferring either the Geiger or Pitch designs were similar: the changes were clearer, more "intuitive" and more helpful. There is a suggestion that pitch-based sonifications are more effective than tempo-based designs [31] (like the Geiger), although musical ability may play a part in how participants understand and make use of pitch sonifications [20].

The position of the Sound Source (Wrist, Object or Both) also had a moderate impact on objective and subjective results. There was no significant effect on accuracy, despite some moderate differences between the values. Providing sound through Both speakers provided little benefit over a single speaker alone. Using Both resulted in the lowest error value, but this was not significant, and using the Object alone was significantly faster than using Both, with the Wrist in between these values. Subjective participant responses generally did not agree that providing sound from both speakers provided a benefit (3.2-3.8 out of 7), perhaps because it did not provide additional information over one speaker.

There was no effect of target azimuth on reaching error, and only Speaker Position 1 vs. Positions 4 and 5 were different in reaching time (but only by ~540ms), suggesting that azimuth had little effect on reaching accuracy. Research that shows an effect of azimuth on auditory localisation tends to use short sound bursts presented to stationary ears [3], whereas the sounds here were more often continuous and the head was free to move. The average azimuth angular error of 11.37° is similar to the figure found by Macé *et al.* [19] (10.8°) using multi-burst sound in their peripersonal reaching study. The two Speaker Layouts equally affected performance and did not interact with Speaker Position.

Overall, the results showed a benefit of dynamic audio feedback but no clear advantage for either sound design or sound source and so we carried out Study 2 involving blind and visually impaired children to gain further insight.

6. STUDY 2: REACHING IN CHILDREN

Study 2 was largely the same as Study 1 but was conducted with blind and visually impaired children to establish if age has an influence in what sound designs are most useful. Eight children (3 female) aged 12 to 17 (mean = 14) were recruited from the visual impairment unit at a local school. The experiment took place in a school office and the distance to the speakers was reduced by approximately 20cm to accommodate the shorter arm reach of the children. The experimental setup and procedure were almost identical to Study 1, with the exception that the number of repetitions was reduced to minimise fatigue in the younger participants. Each child still completed all ten conditions, but only one trial for each speaker (instead of the three in Study 1). The order of *Feedback Designs* was counterbalanced and the order of the *Sound Source* conditions randomised. Due to the limited number of trials, only Speaker Layout 1 was used for all conditions. The same variables and measures as Study 1 were also used, except for Layout.

6.1 Results

6.1.1 Accuracy

The accuracy data were analysed using a 3 x 3 Repeated-measures ANOVA and are summarised in Figure 7. There was a significant effect of *Feedback Design* ($F_{(2,\ 94)} = 3.456$, $p < 0.05$, $\eta_p^2 = 0.07$): Bonferroni pairwise comparisons showed that the *Pitch* design had a significantly higher error rate (36.1%) than the *Geiger* counter (24.3%). The *Constant* design was in between (30.6%). There was no significant main effect of the *Sound Source*: the *Wrist* speaker resulted in 38.9% error, *Object* was 27.1%, and *Both* speakers were 33.3%. There was no interaction effect.

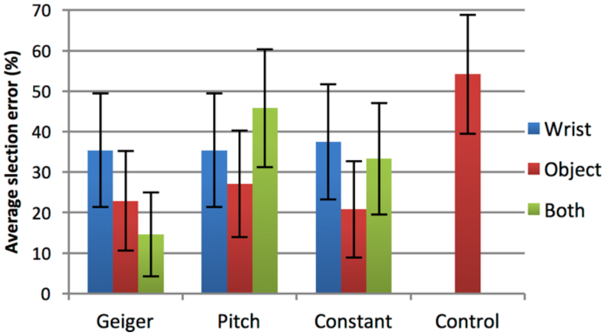

Figure 7: Mean Study 2 target selection error for all Feedback Designs and Speaker Locations. Error bars = 95% CI.

Each individual condition was compared to the *Control* condition. A one-way repeated-measures ANOVA with the 10 Conditions was run and found a significant effect of Condition ($F_{(9,\ 423)} = 3.803$, $p = 0.01$, $\eta_p^2 = 0.07$): the *Control* condition resulted in significantly higher errors (54.2%) than the *Geiger + Both* speakers (14.6%) and *Constant + Object* speaker (20.8%). A one-way repeated-measures ANOVA found a significant effect of *Speaker Position* on accuracy ($F_{(5,\ 395)} = 9.259$, $p < 0.001$, $\eta_p^2 = 0.10$), with mean error values of 15%, 27.5%, 50.0%, 33.8%, 41.3% and 28.8% for Positions 1 to 6, respectively. Position 3 had significantly higher error than Positions 1 and 2, and Position 1 also had significantly lower error than positions 4 and 5. Like the young adults, the majority of confusions were for speakers adjacent to the actual sound source (only 6.6% of responses perceived the sound as further away). The average angular error was 9.88°.

6.1.2 Selection Time

The selection time data again violated normality and so non-parametric analyses were used. There was a significant effect of *Feedback Design* on Selection Time (χ^2 (2)=22.545, $p<0.001$), with the *Constant* design being significantly faster (mean = 3665ms), than both the *Geiger* (4937ms) and *Pitch* (4830ms) designs. There was also a borderline significant effect of *Sound Source* (χ^2 (2)=5.972, $p=0.05$), with *Wrist* sounds resulting in significantly slower times (mean = 5076ms) than *Object* sounds (4211ms). *Both* sounds were in between (4428ms). As with the accuracy data, we compared Selection Time during the *Control* condition and found a significant effect of Condition (χ^2 (9)=41.16, $p<0.001$), with the *Control* condition resulting in significantly faster selection times (3058ms) than all the Dynamic conditions except for *Geiger + Object* and *Pitch + Both*.

6.1.3 Movement Traces

There was no significant main effect of *Feedback Design* on the total movement distance, with mean movement distances of 1.83m, 1.73m and 1.84m for the *Geiger, Pitch* and *Constant* designs, respectively. There was also no main effect of *Sound Source* on movement with means of 2.0m, 1.62m and 1.74m for *Wrist*, *Object* and *Both*, respectively.

6.1.4 Subjective Responses

In response to the "*Connection*" statement the average response was 5.6 for Geiger, 5.7 for Pitch and 4.9 for the Constant design. The average responses to the "*benefit of combination*" statement were below 4 ("Slightly Disagree") for the Geiger and Constant designs but 4.5 for the Pitch ("Slightly Agree"). Most participants said that providing sounds from both locations made the task more difficult, as the extra sounds were distracting and confusing. All Feedback Designs were rated as similarly "pleasant" (> 4). 3 participants preferred the Geiger counter, 2 preferred the Pitch and 3 had no preference.

6.2 Initial Discussion

As there were a different number of data points for the two groups, we did not carry out a statistical comparison of the results. However, there are still some clear differences between the age groups. Accuracy levels were much worse in the children: the young adults had an overall mean error of 17.7% while the children's overall value was 30.3%. In particular, the young adults gained more benefit from the Dynamic feedback designs: error was significantly lower under the Dynamic designs compared to the Constant in young adults. In the children, there was no significant difference between the Dynamic and Constant designs, and the Pitch design was actually significantly worse than the Geiger. Also, while the Control condition led to the highest errors in both groups, it was only significantly worse than two other conditions in young children; it was worse than almost all conditions in young adults. However, comparing Figures 5 (Study 1) and 7 (Study 2) show very similar patterns of reaching error based on condition, even if the absolute values are much higher in children. Reaching times were just under 1 second longer on average in children, which constitutes approximately a 30-40% increase, but is not a high value in real terms, and reaching distance was also slightly higher, but only by around 20-30cm.

Azimuth position had different effects on the groups: there was no effect of Speaker Position on accuracy in young adults, but there was with younger children. Position 3 (Figure 4, left) had significantly higher error than Positions 1 and 2. Position 3 was the most physically distant along the median plane from the children, although the distance was set to be within arm's reach (37cm) and Position 6 had the same distance. Positions 1 and 6 had low error in both groups, as they only have one adjacent speaker, so it may be the combination of physical distance and close azimuthal proximity (to Positions 2 and 4) that made Position 3 hard to find. Despite the differences, the two groups had similar average azimuth errors (11.37° in Study 1 and 9.88° in Study 2) that were similar to Macé *et al.* [19] (10.8°), but lower than in other research [6,7]. It should be noted that, while we deliberately limited the number of trials to minimise fatigue, it is possible that performance could improve over time or more experimental sessions, despite the lack of a learning effect among the Study 1 data.

Subjective views were similar, with the Dynamic designs creating more of a connection between the hand and the object, and providing sound from both the Wrist and Object together was not seen as providing much benefit over just one speaker alone. The feedback preferences were slightly at odds with the objective measurements in Study 2. While the Pitch design resulted in significantly more errors than the Geiger, and also in the longest movement time, it had the highest ratings for all three subjective responses. However, only two participants preferred the Pitch design, with three preferring the Geiger counter (three had no preference). This contrasts with the young adults, three of whom preferred the Pitch design (only one preferred the Geiger). As mentioned earlier, musical ability can influence the usefulness of pitch-based sonifications [20] and so it may be that the two groups had different musical ability, but it could also be that the more developed older adults were better able to make sense and use of the changes in pitch. We will investigate this in the future. Either way, the results strongly suggest that personalisation is an important part of feedback design, and previous research on what kinds of sound blind and visually impaired children like/dislike showed a large degree of variety [33].

Young visually impaired children develop spatial and motor skills more slowly than sighted peers [4,22] and auditory spatial perception is worse at a young age [9]. These findings, along with our own (which show a similar trend), highlight the poorer spatial cognition in younger visually impaired children, bolstering the case for ABBI and using sound to support rehabilitation and improve spatial cognition in young children.

7. GENERAL DISCUSSION

Overall, the results from the two studies suggest that dynamic auditory feedback better supports reaching movements compared to constant (unchanging) feedback or only a brief initial cue (Control condition). Subjective comments from participants also showed a clear preference for dynamic changes, including the view that dynamic changes are "necessary" to guide movement. The results also suggest that sounds from nearby objects are more conducive to accurate reaching than sounds from the wrist, and providing feedback from both the wrist and the object does not appear to provide any benefit over sounds from a single speaker. There was only one significant difference between *Sound Sources:* the *Object*-only condition leading to faster Movement Times than when using *Both* speakers in Study 1.

However, there were some large differences in the results from children compared to young adults. The results from Study 1 (young adults) show a much clearer benefit of dynamic audio feedback over constant (unchanging) feedback, as dynamic feedback did not provide a strong benefit to younger children in Study 2. The control condition had the highest error in both studies, which suggests that these brief signals are not sufficient to inform children or young adults about the location of objects nearby.

7.1 Implications for ABBI & Rehabilitation

Our results have important ramifications for the use of spatialised and on-body sound in learning or habilitation activities. Providing hand guidance information through dynamic changes in sound is a potentially beneficial approach, which suggests that more complex engagement and play setups for children, such as 'sensory resource boxes' augmented with tracking and sounding objects, could be more beneficial than simple adult-activated 'lures' [22]. However, providing sounds from the wrist alone (as in ABBI), or both the wrist and the object, may only partially support accurate reaching, and less so than providing sound through only the object to be reached for. Still, it should be noted that the results from Study 2 suggest that object (e.g., toy) positioning might also be important in encouraging accurate reaching: the high errors suggest objects should be placed more than 10° apart. Also Position 3 was more difficult to find so it may be best to position objects much closer than an arm's reach (even though near sounds are harder to localise in adults [6,7]).

In the case of our feedback designs, the *Coalescent* design does not appear to have added beneficial information over the *Individual* designs. While this is disappointing, we believe our findings can still be useful to other researchers, but there are a few points to consider. Firstly, there is the manner in which sound from the wrist was incorporated into the design. Previous focus group participants suggested that sound from the wrist, together with sound from the environment, would be appreciated in guiding the person to objects of interest in the room. However, it may be that continuous feedback from the wrist distracts attention from the location of the environmental source, and more sporadic or discrete feedback that compliments the environmental sound may be more suitable. Some participants commented that sound from the wrist, physically distant from the target, was distracting or confusing.

A second point is Millar's recommendation to provide redundancy and multiple pieces of systematically interconnected information in the sounds presented to young blind children [22]. We hypothesised that the *Coalescent* feedback may provide these, which may benefit children, but this does not seem to be the case. It may be that the two perceptual streams [5] utilised in the feedback presented additive, and so more perceptually saturating, information rather than repeating (redundant) information. While it is possible that younger children (0-12 years) may make different use of the *Coalescent* feedback, the drop in accuracy with age observed here suggests they may not find benefit in the specific design implemented here. Finally, it should be kept in mind that a key aim for the research, which the Dynamic feedback designs appear to have achieved, was to identify a design that creates a *connection* between the hand and the object: informing the child that there are objects nearby and that they are reachable. In future, we will test if this encourages children to self-initiate engagement with their environment more often. In this case, precise reaching is less important than the elicitation of the reaching itself.

8. CONCLUSIONS

This paper presents the results from two studies involving young visually impaired people aged 12-22 that looked at whether dynamic audio feedback from nearby objects and from a speaker on the wrist can improve peripersonal reaching, compared to unchanging feedback and no online feedback at all. The results showed that children (aged 12-17) were less accurate than young adults (18-22) but, overall, both studies showed a benefit of dynamic feedback on reaching accuracy and sounds from reachable objects are more likely to help reaching than sounds from the wrist or from both the object and wrist together. The dynamic sounds established a perceived connection between the hand and object, which may help young children create an internal connection between his/her movements and the environment, and so support self-initiated engagement.

9. REFERENCES

1. Audacity. Audacity: open source, cross-platform software for recording and editing sounds. Retrieved September 25, 2015 from http://audacityteam.org/
2. Ann E. Bigelow. 1986. The development of reaching in blind children. *British Journal of Developmental Psychology* 4, 4: 355–366.

http://doi.org/10.1111/j.2044-835X.1986.tb01031.x

3. Jens Blauert. 1999. *Spatial Hearing*. MIT Press, Massachusetts.
4. Michael Brambring. 2006. Divergent Development of Gross Motor Skills in Children Who Are Blind or Sighted. *Journal of Visual Impairment & Blindness* 101, 12: 620–634. http://www.afb.org/jvib/jvibabstractnew.asp?articleid=jvib001008
5. Albert S. Bregman. 1994. *Auditory Scene Analysis*. The MIT Press, Cambridge, MA.
6. Douglas S. Brungart, Nathanieal I. Durlach, and William M Rabinowitz. 1999. Auditory localization of nearby sources. II. Localization of a broadband source. *The Journal of the Acoustical Society of America* 106, 4: 1956–1968. http://doi.org/10.1121/1.428212
7. Douglas S. Brungart. 1999. Auditory localization of nearby sources. III. Stimulus effects. *The Journal of the Acoustical Society of America* 106, August 1999: 3589–3602. http://doi.org/10.1121/1.428212
8. Giulia Cappagli, Gabriel Baud-bovy, Elena Cocchi, and Monica Gori. 2014. Audio and proprioceptive space perception in sighted and visually impaired children. *Proceedings of International Multisensory Research Forum (IMRF '14)*.
9. Giulia Cappagli, Elena Cocchi, and Monica Gori. 2015. Auditory and proprioceptive spatial impairments in blind children and adults. *Developmental Science*: n/a–n/a. http://doi.org/10.1111/desc.12374
10. Géry Casiez, Nicolas Roussel, and Daniel Vogel. 2012. 1€ Filter: a Simple Speed-Based Low-Pass Filter for Noisy Input in Interactive Systems. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*, 2527–2530. http://doi.org/10.1145/2207676.2208639
11. Marsha G Clarkson, Rachel K Clifton, and Barbara A Morrongiello. 1985. The effects of sound duration on newborns' head orientation. *Journal of experimental child psychology* 39: 20–36. http://doi.org/10.1016/0022-0965(85)90027-X
12. Sara Finocchietti, Giulia Cappagli, Gabriel Baud-bovy, et al. 2015. ABBI , a new technology for sensory-motor rehabilitation of visual impaired people. *Proceedings of International Conference on Enabling Access for Persons with Visual Impairment (ICEAPVI '15)*, 1–5.
13. Sara Finocchietti, Giulia Cappagli, Lope Ben Porquis, Gabriel Baud-Bovy, Elena Cocchi, and Monica Gori. 2015. Evaluation of the Audio Bracelet for Blind Interaction for improving mobility and spatial cognition in early blind children - A pilot study. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS* 2015-November: 7998–8001. http://doi.org/10.1109/EMBC.2015.7320248
14. Selma Fraiberg. 1977. *Insights from the blind: Comparative studies of blind and sighted infants*. Plenum Press, New York.
15. Monica Gori, Giulio Sandini, Cristina Martinoli, and David C Burr. 2014. Impairment of auditory spatial localization in congenitally blind human subjects. *Brain* 137, 1: 288–293. http://dx.doi.org/10.1093/brain/awt311
16. Lyn Haber, Ralph N. Haber, Suzanna Penningroth, Kevin Novak, and Hilary Radgowski. 1993. Comparison of nine methods of indicating the direction to objects: data from blind adults. *Perception* 22: 35–47. http://doi.org/10.1068/p220035
17. William M. Hartmann. 1983. Localization of Sound in Rooms. *Journal of the Acoustical Society of America* 74, 5: 1380–1391.
18. KitSound. KitSound. Retrieved January 28, 2015 from www.kitsound.co.uk
19. Marc J M Macé, Florian Dramas, and Christophe Jouffrais. 2012. Reaching to sound accuracy in the peri-personal space of blind and sighted humans. *Proceedings of International Conference on Computers Helping People with Special Needs (ICCHP '12)*, 636–643. http://doi.org/10.1007/978-3-642-31534-3_93
20. Lisa M. Mauney and Bruce N. Walker. 2007. Individual Differences and The Field Of Auditory Display: Past Research , a Present Study , and an Agenda For The Future. *Proceedings of ICAD*, 386–390.
21. David McGookin, Stephen A Brewster, and Pablo Priego. 2009. Audio bubbles: Employing non-speech audio to support tourist wayfinding. *Proceedings of International Workshop on Haptic and Audio Interaction Design (HAID '09)*, 41–50. http://doi.org/10.1007/978-3-642-04076-4_5
22. Susanna Millar. 1994. *Understanding and Representing Space: Theory and Evidence from Studies with Blind and Sighted Children*. Oxford Press, Oxford. http://doi.org/10.1093/acprof
23. Gaetan Parseihian, Christophe Jouffrais, and Brian F. G. Katz. 2014. Reaching nearby sources: comparison between real and virtual sound and visual targets. *Frontiers in Neuroscience* 8, September: Article 269. http://doi.org/10.3389/fnins.2014.00269
24. David R Perrott. 1969. Role of signal onset in sound localization. *The Journal of the Acoustical Society of America* 45, 2: 436–445. http://doi.org/10.1121/1.1911392
25. Brad Rakerd and William M. Hartmann. 1986. Localization of Sound in Rooms, III: Onset and Duration Effects. *Journal of the Acoustical Society of America* 80, 6: 1695–1706.
26. Brigitte Röder, Wolfgang Teder-Sälejärvi, Anette Sterr, Frank Rösler, Steven A. Hillyard, and Helen J. Neville. 1999. Improved auditory spatial tuning in blind humans. *Nature* 400, July: 162–166. http://doi.org/10.1038/22106
27. Stuart Ross and Michael J Tobin. 1997. Object Permanence, Reaching, and Locomotion in Infants Who Are Blind. *Journal of Visual Impairment & Blindness* 91: 25–32. http://www.afb.org/jvib/jvibabstractNew.asp?articleid=jvib910105
28. Royal National Institute of Blind People. Sensory Development Resource Boxes. Retrieved March 18, 2015, rnib.org.uk/sites/default/files/sensory_development_resource_boxes[1].doc
29. Dale M Stack, Darwin W Muir, Frances Sherriff, and Jeanne Roman. 1989. Development of infant reaching in the dark to luminous objects and "invisible sounds." *Perception* 18: 69–82. http://doi.org/doi:10.1068/p180069
30. Bruce N. Walker, Gregory Kramer, and David M. Lane. 2000. Psychophysical scaling of sonification mappings. *Proceedings of ICAD 2000*, 99–104. http://dev.icad.org/Proceedings/2000/WalkerKramer2000.pdf
31. Bruce N. Walker and Gregory Kramer. 2005. Mappings and metaphors in auditory displays: an experimental assessment. *ACM Transactions on Applied Perception* 2, 4: 407–412. http://doi.org/10.1145/1101530.1101534
32. Bruce N. Walker and Lisa M. Mauney. 2004. Individual differences, cognitive abilities, and the interpretation of auditory graphs. *Proceedings of ICAD 2004*, 6–10. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.101.1938&rep=rep1&type=pdf
33. Graham Wilson, Stephen A. Brewster, Hector Caltenco, et al. 2015. Effects of Sound Type on Recreating the Trajectory of a Moving Source. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '15) Extended Abstracts*, 1645–1650. http://doi.acm.org/10.1145/2702613.2732821
34. Jennifer G. Wishart, T. G R Bower, and Jane Dunkeld. 1978. Reaching in the dark. *Perception* 7: 507–512. http://doi.org/10.1097/00001756-199010000-00004
35. Eberhard Zwicker, G Flottorp, and Stanley S Stevens. 1957. Critical Band Width in Loudness Summation. *The Journal of the Acoustical Society of America* 29, 5: 548–557.