Research Note

A Two-Interval, Forced-Choice, Observer-Based Procedure for Evaluating Hearing Sensitivity in Children With Motor and Developmental Impairments

Heather L. Porter,^a Emily Buss,^b Jenna Browning,^a and Lori J. Leibold^a

Purpose: It can be challenging to collect reliable behavioral responses to sound from individuals with significant motor or developmental impairments, the most common types of comorbid disability found in children with hearing loss (e.g., Gallaudet Research Institute, 2011). The purpose of this study was to test the feasibility of using a 2-interval, forced-choice, observer-based method for individuals considered to be difficult-to-test using behavioral audiometric assessments.

Method: Participants were 5 children with motor and developmental impairments, ages 5–15 years ($M = 11.6$, $SD = 4.6$). The functional abilities of all participants were greater than 2 SD s below the mean, as measured by the Vineland-II Parent Caregiver Rating Form. Participants listened to either a male talker saying the word “playground”

or a 1000-Hz warble tone, presented via an insert earphone or a sound field speaker. An observer, blind to signal presentation, selected 1 of 2 temporal intervals, determining which contained the signal based only on participant behavior. Criterion was reached when the observer correctly identified the interval containing the signal for 8 of the last 10 trials.

Results: An 80%-correct criterion was met for all participants, suggesting feasibility for use in children with motor or developmental impairment. Two participants were tested using an adaptive tracking procedure; a reliable threshold estimate was obtained for both children. This method offers promise for children who have difficulty performing behavioral audiometric assessments currently in use clinically.

The prevalence of additional disabilities in children with hearing loss is approximately 30%–60%, most commonly involving motor or developmental impairments (Darin, Hanner, & Thiringer, 1997; Gallaudet Research Institute, 2011; Van Naarden Braun et al., 2015; Van Naarden, Decoufle, & Caldwell, 1999). The most recent Gallaudet Research Institute Regional and National Survey, including data from over 37,000 children with hearing loss, lists mental retardation (8.3%), learning disability (8%), attention-deficit disorder or attention-deficit/hyperactivity disorder (5.4%), developmental delay (5.3%), and orthopedic impairment (4.4%) as the five most common

additional disabilities specified for children with hearing loss (Gallaudet Research Institute, 2011). Epidemiological data for children in the United Kingdom ($n = 17,160$) indicate developmental delay and cerebral palsy as the third and fourth most prevalent disabilities experienced by children with hearing loss, respectively (Fortnum, Marshall, & Summerfield, 2002). Similar observations have been made longitudinally on a smaller scale. For example, ongoing estimates from the metropolitan Atlanta area report intellectual disability and cerebral palsy as the two most common disabilities specified in children with hearing loss, with autism spectrum disorder rapidly increasing in prevalence (Van Naarden Braun et al., 2015; Van Naarden et al., 1999).

It is well established that early identification and intervention promotes optimal speech, language, and educational outcomes for all children with hearing loss (e.g., Moeller, 2000). Importantly, similar benefits associated with early identification and intervention have been specified for children with hearing loss who have additional disabilities and their counterparts with no additional disabilities (e.g., Cupples et al., 2018; Yoshinaga-Itano, Sedey, Coulter,

^aCenter for Hearing Research, Boys Town National Research Hospital, Omaha, NE

^bDepartment of Otolaryngology/Head and Neck Surgery, School of Medicine, University of North Carolina at Chapel Hill

Correspondence to Heather L. Porter: heather.porter@boystown.org

Editor-in-Chief: Sumitrajit Dhar

Editor: Ann Eddins

Received October 26, 2018

Revision received February 1, 2019

Accepted March 19, 2019

https://doi.org/10.1044/2019_AJA-18-0162

Disclosure: The authors have declared that no competing interests existed at the time of publication.

& Mehl, 1998). Accordingly, the Joint Committee on Infant Hearing recommends universal newborn hearing screening by 1 month of age, complete diagnostic testing by 3 months of age for infants who do not pass the screening, and intervention by 6 months of age for infants with confirmed hearing loss (Joint Committee on Infant Hearing, 2007). In addition, ongoing timely audiological assessment is recommended for infants and children with risk factors for progressive hearing loss, such as cytomegalovirus infection, diagnosis of neurodegenerative disorder, and receipt of extracorporeal membrane oxygenation or chemotherapy (Joint Committee on Infant Hearing, 2007). For children with hearing loss and additional disabilities, missed or under-managed hearing loss can have cascading effects on speech, language, and educational outcomes. Insufficient auditory input may also prevent optimal outcomes for other therapeutic interventions relying on verbal instruction and reinforcement (e.g., physical therapy, occupational therapy; Rawool, 2010). The ability to accurately assess and manage hearing loss is an essential precursor to initiation of any effective intervention strategy for children with hearing loss and other disabilities.

Behavioral responses to sound are the most reliable and accurate source of information we have about hearing in human listeners, representing the functioning of the entire auditory pathway, from the auditory periphery to cortex (Dworsack-Dodge et al., 2012; Widen, 1993). However, not all listeners are able to provide the conventional responses to sound that are expected with current behavioral audiologic techniques. These techniques require reliable and timely listener action in response to sound such as pressing a button (conventional audiometry), putting a toy block into a bucket (conditioned play audiometry), or head turning (visual reinforcement audiometry; VRA). These types of behavioral responses may be infeasible or delayed for listeners with significant motor or developmental impairments.

An alternative option to estimate auditory sensitivity is through objective measurements, such as auditory brainstem response (ABR) testing, which utilizes electrodes that record the auditory nerve's response to sounds. Although ABR testing may be useful for populations that are challenging to test behaviorally, the data obtained reflect auditory function at or below the brainstem; this provides only partial information about the integrity of the auditory pathway and functional hearing abilities. Another limitation is that ABR is not appropriate for estimating auditory sensitivity in children with hearing disorders affecting neural synchrony (e.g., auditory neuropathy). Finally, ABR testing is highly sensitive to listener movement and noise, requiring participants to remain in a deeply restful or sleep state for the duration of the assessment, which can take 1–2 hr. As a consequence, sedation may be required to maximize the possibility of successfully obtaining results. This introduces potential risk for significant medical complications and may not be justifiable for children with motor or developmental impairments who have other complicating medical conditions such as cardiovascular abnormalities

or respiratory disease and infection (Krauss & Green, 2006; Tobias & Leder, 2011).

Observer-based methods offer potential alternatives for testing individuals who are unable to reliably respond using traditional behavioral audiometric assessments. These methods have been used to assess a wide range of auditory abilities in infants and toddlers, including tone detection in quiet, localization, gap detection, and frequency discrimination (Dasika, Werner, Norton, Nie, & Rubinstein, 2009; Grieco-Calub, Litovsky, & Werner, 2008; Morrongiello, Fenwick, & Chance, 1990; Olsho, Koch, Carter, Halpin, & Spetner, 1988; Olsho, Koch, & Halpin, 1987; Werner & Gillenwater, 1990; Werner, Marean, Halpin, Spetner, & Gillenwater, 1992). The single-interval observer-based psychophysical procedure (OPP) developed by Olsho, Koch, Halpin, and Carter (1987) uses operant conditioning to reinforce infant responses to auditory stimuli. In OPP, the observer is blind to whether or not a signal has occurred on a given trial. The observer must determine the presence or absence of a signal during a trial based only on listener behavior. Depending on the age of the listener, typical behaviors include head turns, eye darts, cessation of sucking or other movement, and eye widening (Watrous, McConnell, Sitton, & Fleet, 1975). Visual and social reinforcement is provided to the listener when the observer correctly identifies trials in which a signal was presented (e.g., activation of a lighted mechanical toy within a plexiglass box).

VRA and OPP are similar, in that they both rely on operant conditioning. Moreover, prior research has demonstrated that OPP and VRA yield similar pure-tone detection thresholds across a range of frequencies (Olsho, Koch, Halpin, & Carter, 1987). However, there are several important distinctions between VRA and OPP. One essential difference between the two methodologies is that VRA is influenced by examiner bias because the observer initiates the presentation of signals. Knowing when a signal is presented may influence the observer's decision regarding whether or not the listener made a response. In OPP, the observer initiates each trial but is blinded to whether or not a signal is presented. This provides opportunity to accept a variety of listener responses to sound, provided those responses can be reliably observed. Although VRA is of limited utility in listeners with a developmental age below 6 months (Gans & Gans, 1993), OPP can be effective in infants as young as 1 month of age (Werner & Gillenwater, 1990). The flexibility of OPP in terms of accommodating a variety of behavioral responses to sound is of obvious utility for assessing children with motor and developmental impairments, as their responses may be subtle, delayed, or inconsistent. The fact that the observer is blinded to trial type protects against observer response bias, adding confidence to estimates of the reliability of behavioral responses for a given child.

A challenge associated with the single-interval OPP is that training is required to familiarize inexperienced observers with common behaviors associated with responses to sound and reduce observer false alarm rate (Olsho, Koch, Halpin, & Carter, 1987). A modification of the single-interval OPP that could make it easier for observers to detect

subtle responses to auditory stimuli is to use a two-interval forced-choice (2IFC) procedure. In the single-interval OPP, the observer must indicate “yes” or “no” regarding their observation of whether or not a signal was presented. In the 2IFC observer-based procedure, a signal is presented on each trial in one of two temporal intervals; the observer must decide which interval contained the signal.

Browning, Buss, and Leibold (2014) evaluated the feasibility and efficiency of a 2IFC observer-based procedure using a 500-ms low-pass filtered broadband noise for 7- to 9-month-old infants with normal hearing. The signal was presented in one of two 500-ms temporal intervals, visually marked in real time, separated with a 300-ms inter-stimulus interval (ISI). Observers were given feedback after every trial indicating the signal interval. Participants were reinforced with activation and illumination of a mechanical toy when the observer correctly identified trials in which a signal was presented. This toy reinforcement was timed to maintain a consistent temporal relationship between the signal and reinforcement, regardless of whether the signal occurred in the first or second interval. The results of that study demonstrated that a 2IFC observer-based procedure was feasible for use with infants who had normal hearing. Moreover, fewer trials were required to reach a criterion of 80%-correct detection for the 2IFC observer-based procedure compared to previous studies using the single-interval OPP (e.g., Leibold & Werner, 2006; Werner & Boike, 2001). This increased test efficiency for the 2IFC observer-based method may be associated with a reduction in the effort required to observe a change in infant behavior between two listening intervals compared to the effort required to determine a behavior associated with signal detection based on task feedback over a series of trials, as in the single-interval OPP.

The 2IFC observer-based procedure could substantially improve our ability to evaluate hearing in listeners with significant developmental or motor impairments because it accommodates a wide range of listener responses, protects against examiner bias, and may be more efficient than single-interval procedures (Browning et al., 2014). The purpose of this study was to test the feasibility of a 2IFC observer-based procedure for behavioral evaluation of hearing in children with motor and developmental impairment. Feasibility was assessed by determining the success rate of participant/observer pairs for achieving criterion, defined as the observer's correct identification of the interval containing the signal for eight of the last 10 trials in a single test session. Reliability for estimating hearing threshold was qualitatively examined in a subset of children through visually inspecting adaptive tracks, determining the range of reversals, assessing probe hit rate, and comparing threshold estimates across runs within an individual.

Method

Participants

Children between 5 and 15 years of age with significant motor or developmental impairment were recruited for

participation over a 2-year period at two test sites (University of North Carolina and Boys Town National Research Hospital). Children with and without hearing loss were invited to participate. Recruitment criteria were broad to maximize opportunities for participant enrollment. Study procedures were approved by the institutional review boards at the University of North Carolina and Boys Town National Research Hospital. Informed consent was obtained following a process that included a formal review of consenting paperwork, an informal tour of the laboratory space, and guided familiarization to the sound booth and study materials. Parents of children unable to indicate assent (e.g., provide a signature, verbally agree) provided consent for their children, and the reason a child assent form was not completed was documented and submitted to the institutional review board. Parents were counseled on the right to withdraw, and research staff monitored children closely for signs of discomfort or distress.

Five children were consented to participate, referred to as C1 through C5 for this report. Participant diagnoses included cerebral palsy ($n = 3$), autism ($n = 1$), and metabolic disorder accompanied by motor and developmental delay ($n = 1$). The average age of participants was 11.6 years ($SD = 4.6$ years, range: 5–15 years). Three children had known hearing loss and wore bilateral amplification or were in the process of obtaining amplification. Prior to testing, caregivers of all participants indicated some degree of motor and developmental delay resulting in difficulty or failure to obtain reliable information regarding hearing sensitivity via behavioral audiometry in a clinical setting. Additional detail is provided about individual participants below.

Parent Survey

During testing, a parent or caregiver completed the Vineland-II Parent/Caregiver Rating Form (Vineland-II; Sparrow, Cicchetti, & Balla, 2005), designed to assess the development and functioning of individuals with and without disabilities. The Vineland-II provides an estimate of adaptive behavior, defined as performance of activities required for personal and social sufficiency. This construct is bound by acknowledgement that adaptive behavior is age related, defined by the expectations and standards of those in one's community, modifiable, and defined by typical performance, not ability.

Stimuli

The target signal was either a recording of the word “playground” produced by an adult male talker, with 815-ms duration, or a 1000-Hz warble tone that was 500 ms in duration with 10-ms rise and fall times. The rationale for using speech and tonal stimuli was to provide test flexibility in order to maximize opportunities for success. The target signal was chosen by the research staff based on the observed responsiveness of the child before the listening session and conversations with parents based on typical child behaviors

outside the laboratory setting. Prior research has shown that speech is more likely to elicit an unconditioned response than pure tones (e.g., Hoversten & Moncur, 1969). Once chosen, the same target was used for all phases of testing.

On each trial, the target signal was presented in quiet in one of two listening intervals. The time between the two intervals ranged from 900 to 1200 ms, adjusted for each participant based on their response time. ISIs of 300 ms have been successfully implemented with typically developing infants in a 2IFC observer-based paradigm in our laboratory (Browning et al., 2014). A range of longer ISIs than previously used was implemented in this study to accommodate an anticipated delayed participant motor response. Pilot testing indicated a maximum ISI of 1200 ms in order for participants to successfully establish and maintain a relationship between the target signal and response reinforcement. Similar to the method used for target signal determination, the duration of ISI was initially chosen based on observation of child behaviors and conversations with parents based on typical child behaviors outside the laboratory setting. Adjustments were made to ISI during the first phase of testing (i.e., the conditioning phase, described below), but no further adjustments were made once an appropriate ISI was determined.

Custom software was used to control stimulus generation and presentation (MATLAB, MathWorks). Stimuli were amplified (Applied Research and Technology, SLA4) and presented to three participants through an ER-1 insert earphone (Etymotic Research, Inc.). Two participants listened to stimuli through a sound field speaker (RadioEar SP90, RadioEar) at 45° azimuth and 0° elevation. One participant who was tested in the sound field wore personal amplification during sound field testing.

Participant Response Reinforcement

Caregivers were asked about their child's needs and preferences prior to the test session, and visual reinforcement was tailored to meet participant needs on an individual basis, as described below. Social reinforcement was provided for all participants.

Procedure

Test sessions lasted approximately 60–90 min, including the child's listening session and caregiver completion of the Vineland-II. The listening session lasted approximately 30 min, including up to three phases: (a) conditioning, (b) establishing criterion, and (c) threshold estimation. An observer, blinded to stimulus presentation, was seated at a computer in front of a window looking into a sound-attenuated test booth. The observer's role included managing the software used for experimental control, initiating trials, and monitoring participant behaviors. Listening intervals were indicated visually on the computer monitor. The observer received feedback regarding which interval contained the stimulus after providing a response at the end of a trial. Consistent with previous research using single-interval

and 2IFC observer-based paradigms, reinforcement was provided to the listener when the observer correctly identified trials in which a signal was presented (e.g., Browning et al., 2014; Olsho, Koch, Halpin, & Carter, 1987). An assistant sat inside the booth performing up to two primary duties. The assistant provided visual reinforcement to participants with visual impairments (i.e., activating a lighted toy in close proximity to the participant) and supplementary response reinforcement to all participants using encouraging phrases and praise when the observer correctly identified trials in which a signal was presented. They also collaborated with the observer to determine when a response to the signal occurred for most listening sessions. Although this was a collaborative relationship, the observer maintained responsibility for final determination of the interval in which the signal occurred. The assistant and observer communicated through a remote microphone system during most listening sessions, allowing the assistant to know when the two temporal intervals were occurring. For one listener (C5), the assistant was not blind to signal presentation. In this case, the remote microphone system was not used, and the assistant refrained from collaborating with the observer to determine the response interval.

Observer/Assistant Pairs

Data were collected by six individuals, four of whom either completed or were in the process of completing a clinical doctorate in audiology program with a focus in pediatric audiology. The other two research staff had at least 1 year of experience working with children in our laboratory. Specific training in the 2IFC observer-based method within our laboratory includes graduated participation from passive observer of data collection, to test assistant, to primary observer; training is supplemented by didactic instruction and journal article discussions. The amount of training research staff received prior to participation in data collection for this study varied based on their previous experience with pediatric audiology and observer-based test procedures.

Multiple observer/assistant pairs were used for this study to assess feasibility for use in test environments made up of various potential testers (e.g., high-volume pediatric audiology clinics, large research laboratories). Observer/assistant pairs varied for each participant based on staff availability at the time of testing. To maximize possibility for successful data collection, the observer in the observer/assistant pair was always an individual with training in pediatric audiology. The observer and assistant were not blinded to the type of target stimuli used during a test session. Knowledge of the target stimuli was used to describe the task to parents and participants and during participant response reinforcement.

Conditioning Phase

The conditioning phase included stimulus presentation at 70 dB SPL (C1, C2, C4, and C5) or 80 dB SPL (C3) in the second of the two temporal intervals. This presentation level was chosen with the expectation that it would be clearly

audible to participants. During this phase, the relationship between the presentation of the stimulus and reinforcement was established. This phase also allowed the observer to become familiar with the participant's responses to the stimuli. The conditioning phase ended when the observer subjectively determined that the child was successfully conditioned and consistent responses were observed, after a minimum of three trials.

Criterion Phase

In the criterion phase, the stimulus was presented in one of the two listening intervals, selected randomly on each trial. Recall that the observer was blind to which temporal interval contained the signal, requiring them to base their decision solely on the participant's behavior. The observer received feedback regarding which interval contained the stimulus after every trial. Reinforcement was provided to the child in the booth only if the observer correctly identified the interval containing the stimulus. Criterion was reached when the observer scored eight out of 10 consecutive responses correct.

Threshold Estimation Phase

Threshold estimation was attempted only if the participant/observer pair achieved criterion and participants maintained attentiveness, as subjectively determined by the observer. Similar to the criterion phase, the probability of the stimulus being presented in either interval was 0.50, and the observer was blind to the temporal interval containing the signal, requiring a decision to be made based on participant behavior. Starting levels were chosen to be clearly audible to participants based on known audiometric thresholds or assumption of normal hearing sensitivity based on parent report. A two-down, one-up stepping rule was used to obtain an estimate of the signal level required to achieve 70.7% correct detection (Levitt, 1971). The step size was 4 dB for the first two reversals and 2 dB thereafter. Each run included eight reversals; the last six reversals were used to estimate threshold. Probe trials were randomly presented twice out of every 12 trials, always at 80 dB SPL. Probe hit rates needed to be greater than 80% for data to be considered valid. The adaptive track was based on participant response to signal trials only; participant responses to probe trials did not affect the adaptive track.

Results

This feasibility study relies on a descriptive approach to data analysis to determine the extent to which the 2IFC observer-based method can be used for children with significant developmental or motor impairment. This method includes up to three phases: (a) conditioning, (b) criterion, and (c) threshold estimation. The primary outcome of interest for this feasibility study was the success rate of participant/observer pairs for achieving criterion in a single test session. The number of trials needed to reach criterion was recorded as a descriptive supplement to the primary outcome of interest. Threshold estimation was only attempted for a

subset of children who were subjectively determined to be well suited for success within a single testing visit. These threshold data are also considered qualitatively.

All participants successfully achieved criterion in a single testing session, indicating that the participant/observer pair reliably and accurately identified behavioral responses to suprathreshold sounds. The number of trials required for the conditioning phase ranged from three to 11 (see Table 1), illustrating the diversity in responsiveness to auditory stimuli in this small group of children with developmental and motor impairments. The number of trials required to meet the 80%-correct criterion ranged from eight (the minimum requirement) to 44 (see Table 1). Similar to the number of trials required for conditioning, the wide range of trials needed to reach criterion also demonstrates the diversity of participant responsiveness. This range also provides insight into the possible benefit of observer experience, described in detail below (see C3). Two participants maintained attentiveness, such that threshold was successfully estimated immediately after criterion was reached (i.e., C2 and C4).

Method specifications and individual participant characteristics are described below to inform feasibility for potential adaptation in other research or clinical settings. Participant response type fit into two categories: variable and delayed hand raise.

Response Type: Variable

C1

Diagnosed with cerebral palsy, this 14-year-old participant presented with a hearing health history including chronic ear infections, pressure equalization tube placement in early childhood, and no parental concern for permanent hearing loss. This child's Adaptive Behavior Composite standard score was 49 (95% confidence interval = ± 7), indicating low general adaptive functioning. This score is less than the 0.1 percentile for age, meaning that this child's scores are lower than 99.9% of other 14-year-olds.

The observer and assistant relied on eye flutters, localization toward visual reinforcement, and smiling as responses to auditory stimuli. Visual reinforcement consisted of

Table 1. Number of trials required for each test phase.

Participant	Test phase		
	Conditioning	Criterion	Threshold (dB SPL)
C1	11	9	n/a
C2	3	8	35
C3	6	44 ^a	n/a
C4	3	8	53
C5	5	8	n/a

Note. n/a = not applicable.

^aThis result includes trials from an unsuccessful attempt at reaching criterion ($n = 30$) and a successful attempt at reaching criterion ($n = 14$) occurring immediately following the first attempt. For the second attempt, the more experienced observer was seated inside the booth with the participant.

mechanical toys housed within dark plexiglass boxes. The stimulus was a male talker saying the word “playground,” presented through an insert earphone. The time between listening intervals was adjusted from 900 ms to 1200 ms to accommodate prolonged participant response time. After this adjustment, 11 and nine trials were needed for the conditioning and criterion phases, respectively. Following criterion achievement with speech stimuli, testing was attempted with 1000-Hz warble tone stimuli; however, testing was discontinued during the conditioning phase due to participant fatigue.

C3

Diagnosed with cerebral palsy, this 7-year-old participant presented with known bilateral moderate hearing loss. Testing was completed using an insert earphone; no amplification was worn during the test session. This child’s Adaptive Behavior Composite standard score was 48 (95% confidence interval = ± 5), indicating low general adaptive functioning. This score is less than the 0.1 percentile for age, indicating that this child’s scores are lower than 99.9% of other 7-year-olds.

The observer and assistant relied on subtle changes in body movement and occasional eye darts as responses to auditory stimuli. Visual reinforcement was provided with lighted handheld toys, based on the participant’s limited visual acuity. The stimulus was a 1000-Hz warble tone presented through an insert earphone; a 1200-ms ISI was used throughout the test session. Six trials were needed for the conditioning phase. Two criterion phases were undertaken. During the first criterion phase, the participant/observer pair failed to reach 80% correct after 30 trials. A second criterion phase was undertaken using the assistant and observer in opposite locations. During the second criterion phase, 14 trials were needed to reach criterion. Additional testing was not attempted based on participant fatigue.

Of interest, the assistant in the booth with the participant during the second criterion phase was an audiologist with clinical training in pediatric audiology, including experience with children with motor and developmental delays and research experience using observer-based methods during undergraduate and graduate studies. This suggests that experience in conjunction with close proximity to the participant may be beneficial to distinguishing subtle responses to auditory stimuli in individuals with developmental or motor impairment.

C5

Diagnosed with a rare metabolic disorder including motor and developmental delays, this 5-year-old participant presented with recently diagnosed mild bilateral hearing loss via ABR testing under sedation and was in the process of obtaining amplification. Testing was first attempted using an insert earphone, but the transducer used for stimulus presentation was changed to a sound field speaker due to listener discomfort with the insert earphone. This child’s Adaptive Behavior Composite standard score was 49 (95% confidence interval = ± 5), indicating low general adaptive

functioning. This score is less than the 0.1 percentile for age, indicating that this child’s scores are lower than 99.9% of other 5-year-olds.

The observer relied on head turns, changes in facial expression, and decreased body movement as indication of response to auditory stimuli. Mechanical toys and social praise were used as response reinforcement. The stimulus was a 1000-Hz warble tone; a 1200-ms ISI was used throughout the test session to accommodate for participant motor delay. Five trials were used for the conditioning phase. Eight trials were needed to reach criterion, the minimum number possible. Additional testing was not attempted due to participant restlessness and indication of desire to discontinue the listening session.

Response Type: Delayed Hand Raise

C2

Diagnosed with cerebral palsy, this 14-year-old participant presented with known bilateral hearing loss. Testing was completed using a sound field speaker for stimulus presentation, and the participant wore personal hearing aids during the test session. This participant’s Adaptive Behavior Composite standard score was 65 (95% confidence interval = ± 7), indicating low general adaptive functioning. This score is at the 1.0 percentile for age, indicating that this child’s scores are lower than 99.0% of other 14-year-olds.

The observer relied on a delayed hand raise to indicate response to auditory stimuli. Social praise was used as reinforcement for correct responses. The stimulus was a 1000-Hz warble tone; a 1200-ms ISI was used throughout the test session to accommodate for delayed response times. Three trials were used for the conditioning phase. Eight were needed to reach criterion, the minimum number possible. Following criterion achievement, two adaptive threshold estimation tracks were successfully completed; each track comprised 35 trials. Thresholds for the first and second tracks were 38 and 37 dB SPL, respectively. Probe hit rate was 100% for both tracks. Adaptive tracks were orderly and indicated consistent responses throughout the testing session (see Figure 1, left panel). The standard deviations for reversals in the first and second tracks were 1.8 and 1.7, respectively.

C4

Diagnosed with autism, this 15-year-old participant presented with no parental concern for permanent hearing loss or history of chronic ear infections. This child’s Adaptive Behavior Composite standard score was 61 (95% confidence interval = ± 7), indicating low general adaptive functioning. This score is at the 0.5 percentile for age, indicating that this child’s scores are lower than 99.5% of other 15-year-olds.

The observer and assistant relied on a delayed hand raise to indicate response to auditory stimuli. Social praise was used as reinforcement for correct responses. The stimulus was a 1000-Hz warble tone presented through an insert earphone. A 1200-ms ISI was used throughout the test session.

Figure 1. Adaptive track(s) for C2 (left panel) and C4 (right panel) using a two-down, one-up stepping rule to obtain an estimate of the signal level required to achieve 70.7% correct detection.

Similar to C2, three trials were used for the conditioning phase, and the minimum number of trials was needed to reach criterion (eight). Following criterion achievement, threshold was estimated successfully in 53 trials (threshold = 15 dB SPL). The probe hit rate was 90%. The adaptive track was orderly and indicated consistent responses throughout the track (see Figure 1, right panel). The standard deviation of reversals was 2.1 dB.

Discussion

The 2IFC observer-based procedure has the advantage of providing an estimate of performance that is not biased by the observer's prior expectations. Observer bias is prevented by blinding the observer to the stimulus presentation interval, such that the observer's decision is based solely on participant behavior. This approach allows for increased tester confidence in behavioral estimation of hearing sensitivity compared to methods in which the observer is responsible for both stimulus presentation and determining whether a response to sound has occurred. For example, an audiologist might obtain an estimate of hearing sensitivity using behavioral audiometry but rate the reliability as "poor" because the listener's responses to sound were delayed or weakly executed. Using the 2IFC observer-based procedure, the audiologist would have confidence that reliable responses were observed.

Another important advantage to the 2IFC observer-based procedure is that it provides the observer with sufficient information to adapt response criteria to changes in participant behaviors, which can vary throughout a test session. This flexibility is based on the fact that the observer can use any change in behavior that corresponds to one of the listening intervals. The ability to accommodate changes in participant behavior increases the range of participants who can be tested using these methods.

Consistent with these advantages, there was a 100% yield rate for criterion achievement, suggesting that the 2IFC observer-based procedure has potential use for behavioral

assessment of children with developmental or motor impairments. Although not a primary outcome of interest for this study, this method was also successfully used to reliably estimate hearing threshold in both participants for whom an attempt to obtain threshold was made, indicating potential efficacy for obtaining behavioral hearing thresholds in children traditionally considered to be difficult to test. Although response type was not a primary factor in the decision to attempt to estimate hearing threshold, neither of the participants for whom a threshold was obtained were categorized as having variable response types. It is reasonable to consider that attempts to estimate threshold from a larger group of children that includes those with variable response types could have a lower success rate than that observed in this study. Moreover, follow-up studies are needed to determine the accuracy of threshold estimation using this procedure by comparing thresholds obtained at multiple frequencies with the 2IFC method to those obtained clinically (e.g., ABR testing). In addition, it would be informative to evaluate overall test duration, comparing the 2IFC procedure and standard clinical assessments. Future studies are needed to determine the utility of incorporating the observer-based 2IFC into clinical assessment of children who are hard to test using traditional methods.

Another unresolved question is whether the 2IFC observer-based method requires less training for inexperienced observers than single-interval methods, such as OPP or VRA. In theory, comparison of behaviors between two listening intervals separated by a short duration should reduce observer memory load and provide more opportunity to become familiar with subtle participant responses than single-interval methods. However, observer experience with clinical pediatric audiology and children with developmental disabilities seemed to be a benefit for assessing at least one participant included in this report (i.e., C3). As such, questions regarding the amount of experience and training required for successful use require further investigation.

The efficiency of the 2IFC observer-based procedure for children with motor and developmental impairments

can be estimated by comparing the number of trials required to reach criterion (i.e., 80% correct) to other studies using this method. One of the children in this study required nine trials to reach the 80%-correct criterion (C1), whereas three others required eight trials (C2, C4, and C5). Browning et al. (2014) reported that an average of 11 trials were required to reach the 80%-correct criterion for 7- to 9-month-old infants, indicating that this method may yield similar efficiency for children with motor and developmental impairments as observed for 7- to 9-month-old infants. The 2IFC observer-based procedure seems to be more efficient than the single-interval observer-based procedure for most participants. The single-interval observer-based procedure has been shown to require 20 or more trials to reach the 80%-correct criterion for infants (e.g., Leibold & Werner, 2006; Werner & Boike, 2001). In contrast to the relatively small number of trials required to reach criterion for four of the five listeners tested in the current study, one participant required a total of 44 trials to reach criterion (C3). Recall that the observer team members switched places during the criterion phase for this child; the observer and assistant exchanged locations after the first observer failed to reach criterion after 30 trials. The researcher who relocated into the booth had more experience working with the observer-based method and with children with motor and developmental delays, which may have been a critical factor in the ability to reach criterion after 14 trials once locations were reversed.

Clinical Implications and Future Considerations

The 2IFC observer-based method shows potential for clinical use for children with motor and developmental impairments in at least two ways. First, it is sensitive to variable participant responses. This differs from traditional behavioral testing that requires robust participant responses such as pressing a button, putting a block into a bucket, or making a clear head turn in response to sound. Responses to auditory stimuli offered by C1, C3, and C5 included eye flutters, eye darts, smiling, and changes in body movement—responses that fall outside those typically accepted as responses in a conventional clinical setting. The 2IFC observer-based procedure reduces bias by blinding the observer/examiner to signal presentation; the observation that all five participants tested in this study met a response criterion of 80%-correct indicates that each child provided reliable behavioral responses. Second, this procedure exhibits potential for clinical utility for children who are capable of providing to sound that may not be time-locked to presentation of auditory stimuli. A delayed hand raise in response to auditory stimuli in a clinical setting may result in examiner uncertainty and decreased test reliability. Two participants reported here responded to auditory stimuli with considerably delayed hand raises, yet threshold was successfully determined for both. Moreover, probe hit rates were 100% (C2) and 90% (C4), quantifiably verifying test reliability.

Identification of subtle participant responses and determination of adequate ISIs for delayed hand raises can be challenging. Interviewing parents and caregivers prior to the session regarding their child's responses to sound can help observers recognize subtle behaviors made in association to target stimuli. Spending time interacting with the child and family prior to the listening session may also benefit response identification during testing and provides an opportunity to informally gauge the average time it takes the participant to respond. The success of 2IFC observer-based procedure for participants with delayed responses to sound relies on selection of an adequate ISI. A brief ISI may prohibit identification of a slow response to a signal presented in the first temporal interval; an extended ISI may unnecessarily extend the duration of the listening session, risking participant fatigue or habituation resulting in discontinuation of the session. Although it was not the case for participants included in this report, successful use of the 2IFC observer-based procedure might require multiple visits to allow for testers to become familiar with the response behaviors of participants and establish an appropriate ISI.

The positive results in this preliminary study using 2IFC observer-based procedures to evaluate auditory sensitivity warrant further investigation to verify feasibility for use with participants with motor and developmental impairment and to determine the potential for application in clinical settings. Recruitment efforts for this report spanned 2 years and two research settings (University of North Carolina and Boys Town National Research Hospital), as well as extensive community outreach with local schools, day care centers, and hospitals. Nonetheless, only five participants with motor and developmental impairments were successfully recruited to participate. Caregivers of children who were contacted but chose not to participate often cited time and difficulty getting to the laboratory as deterrents. Caregivers of children with motor and developmental impairments frequently manage schedules that include multiple appointments with therapists, specialists, and educators. For this reason, it may be advantageous to conduct future work on this procedure within a center that provides care to large numbers of participants with motor and developmental delays. Specifically, testing immediately following a previously scheduled clinical audiology visit would eliminate potential burdens associated with coordinating travel and participation in a separate appointment and potentially the need to move to a different location within the clinical center as this paradigm can readily be performed in most clinical audiology settings. Feasibility of such an arrangement is high because this method employs hardware typically used in pediatric audiology test suites (i.e., computer, sound transducers, toy reinforcers), custom-made software that readily transfers across sites, and a user-friendly software interface. Investigation of a 2IFC observer-based method within a center providing audiologic care to a substantial population of individuals with motor and developmental delays could yield data from which conclusions regarding the clinical applicability of these methods can be made.

Participants in this study completed testing separately from clinical visits, so the sample could include a bias toward children with less complicated care requirements, greater family resources, and/or potentially higher levels of responsivity and likelihood for success than children whose families chose not to participate. However, the Adaptive Behavior Composite standard scores of all participants reported here were at least 2 *SDs* below the mean, indicating significant impairment in general adaptive functioning. These results broadly suggest that this sample of children was inclined toward modest and unreliable behavioral responsiveness to sound. It is possible that the 2IFC observer-based method success rate may differ for a larger sample of participants with motor and developmental impairment. It is not unreasonable to speculate that a larger sample of children could yield a more diverse range of adaptive behavior skills, potentially indicating utility for children with adaptive behavior skills more closely matched to age-based normative data than the participants included in this report.

In addition to potential objectivity confounds, the Vineland-II may not be the ideal questionnaire for use in this population because parents occasionally reported feelings of frustration or sadness when answering questions related to their child's capabilities. Parents reported feeling that this questionnaire highlighted their child's limitations and that they would have preferred an assessment that allowed them to focus on their child's strengths. This consequence could be avoided in future work by obtaining estimates of development through investigator assessment or clinical report.

Conclusion

Current behavioral audiometric procedures are the result of decades of study (Widen & Keener, 2003). However, the 2IFC observer-based method has the potential to improve the reliability of hearing threshold estimation of children who are often considered to be difficult-to-test using traditional behavioral audiometric assessments. All participants met an 80%-correct criterion, suggesting feasibility for use in children with motor or developmental impairment. Moreover, well-formed adaptive tracks, low reversal standard deviations, and high probe hit rates suggest that reliable estimates of threshold can be obtained using this procedure. This method offers promise for children who have difficulty performing behavioral audiometric assessments currently in use clinically. Future studies are needed to determine feasibility in a larger sample of participants with motor or developmental impairment and compare threshold estimates obtained with the 2IFC observer-based procedure and those obtained during the course of clinical care.

Acknowledgments

Data collection and article preparation were supported by the National Institute on Deafness and Other Communication Disorders Grant R01 DC014460 (awarded to Emily Buss). The

authors are grateful to the members of the Human Auditory Development Lab, particularly Hannah Hodson McLean who provided organizational input and recommendations across testing sites.

References

- Browning, J., Buss, E., & Leibold, L. J. (2014). Preliminary evaluation of a two-interval, two-alternative infant behavioral testing procedure. *The Journal of the Acoustical Society of America*, 136(3), EL236–EL241.
- Cupples, L., Ching, T. Y. C., Button, L., Leigh, G., Marnane, V., Whitfield, J., . . . Martin, L. (2018). Language and speech outcomes of children with hearing loss and additional disabilities: Identifying the variables that influence performance at five years of age. *Language and speech outcomes of children with hearing loss and performance at five years of age. International Journal of Audiology*, 57, S93–S104.
- Darin, N., Hanner, P., & Thiringer, K. (1997). Changes in prevalence, aetiology, age at detection, and associated disabilities in preschool children with hearing impairment born in Goteborg. *Developmental Medicine & Child Neurology*, 39, 797–802.
- Dasika, V. K., Werner, L. A., Norton, S. J., Nie, K., & Rubinstein, J. T. (2009). Measuring sound detection and reaction time in infant and toddler cochlear implant recipients using an observer-based procedure: A first report. *Ear and Hearing*, 30(2), 250–261. <https://doi.org/10.1097/AUD.0b013e3181986dfe>
- Dworsack-Dodge, M. M., Gravel, J. S., Grimes, A. M., Hunter, L., Johnson, K., Neault, M., . . . Yellin, W. (2012). *Audiologic guidelines for the assessment of hearing in infants and young children*. Fairfax, VA: The American Academy of Audiology.
- Fortnum, H. M., Marshall, D. H., & Summerfield, A. Q. (2002). Epidemiology of the UK population of hearing-impaired children, including characteristics of those with and without cochlear implants —Audiology, aetiology, comorbidity and affluence. *International Journal of Audiology*, 41, 170–179.
- Gallaudet Research Institute. (2011). *Regional and national summary of data from the 2009–10 annual survey of deaf and hard of hearing children and youth*. Washington, DC: GRI, Gallaudet University.
- Gans, D., & Gans, K. D. (1993). Development of a hearing test protocol for profoundly involved multi-handicapped children. *Ear and Hearing*, 14(2), 128–140.
- Grieco-Calub, T. M., Litovsky, R. Y., & Werner, L. A. (2008). Using the observer-based psychophysical procedure to assess localization acuity in toddlers who use bilateral cochlear implants. *Otology & Neurotology*, 29, 235–239.
- Hoversten, G. H., & Moncur, J. P. (1969). Stimuli and intensity factors in testing infants. *Journal of Speech and Hearing Disorders*, 12, 687–702.
- Joint Committee on Infant Hearing. (2007). Year 2007 position statement: Principles and guidelines for early hearing detection and intervention programs. *Pediatrics*, 120, 898–921.
- Krauss, B., & Green, S. M. (2006). Procedural sedation and analgesia in children. *The Lancet*, 367(9512), 766–780.
- Leibold, L. J., & Werner, L. A. (2006). Effect of masker-frequency variability on the detection performance of infants and adults. *The Journal of the Acoustical Society of America*, 119(6), 3960–3970.
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, 49(2), 467–477.
- Moeller, M. P. (2000). Early intervention and language development in children who are deaf and hard of hearing. *Pediatrics*, 106(3), e43.

- Morrongiello, B. A., Fenwick, K. D., & Chance, G.** (1990). Sound localization acuity in very young infants: An observer-based testing procedure. *Developmental Psychology*, 26(1), 75–84.
- Olsho, L. W., Koch, E. G., Carter, E. A., Halpin, C. F., & Spetner, N. B.** (1988). Pure-tone sensitivity of human infants. *The Journal of the Acoustical Society of America*, 84(4), 1316–1324.
- Olsho, L. W., Koch, E. G., & Halpin, C. F.** (1987). Level and age effects in infant frequency discrimination infants. *The Journal of the Acoustical Society of America*, 82(2), 454–464.
- Olsho, L. W., Koch, E. G., Halpin, C. F., & Carter, E. A.** (1987). An observer-based psychoacoustic procedure for use with young infants. *Developmental Psychology*, 23(5), 627–640.
- Rawool, V. W.** (2010). Invisible hearing loss among obvious multiple disabilities, part 1: Ensuring auditory care. *Hearing Review*, 17(1), 18–21.
- Sparrow, S., Cicchetti, D., & Balla, D.** (2005). *Vineland adaptive behavior scales—Second Edition (Vineland-II), survey interview form/caregiver rating form*. Livonia, MI: Pearson Assessments.
- Tobias, J. D., & Leder, M.** (2011). Procedural sedation: A review of sedative agents, monitoring, and management of complications. *Saudi Journal of Anaesthesia*, 5(4), 395–410.
- Van Naarden Braun, K., Christensen, D., Doernberg, N., Schieve, L., Rice, C., Wiggins, L., . . . Yeargin-Allsopp, M.** (2015). Trends in the prevalence of autism spectrum disorder, cerebral palsy, hearing loss, intellectual disability, and vision impairment, Metropolitan Atlanta, 1991–2010. *PLOS ONE*, 10, e0124120.
- Van Naarden, K., Decoufle, P., & Caldwell, K.** (1999). Prevalence and characteristics of children with serious hearing impairment in metropolitan Atlanta, 1991–1993. *Pediatrics*, 103(3), 570–575.
- Watrous, B. S., McConnell, F., Sitton, A. B., & Fleet, W. F.** (1975). Auditory responses of infants. *Journal of Speech and Hearing Disorders*, 40(3), 357–366.
- Werner, L. A., & Boike, K.** (2001). Infants' sensitivity to broadband noise. *The Journal of the Acoustical Society of America*, 109(5), 2103–2111.
- Werner, L. A., & Gillenwater, J. M.** (1990). Pure-tone sensitivity of 2- to 5-week old infants. *Infant Behavior and Development*, 13(3), 355–375.
- Werner, L. A., Marean, G. C., Halpin, C. F., Spetner, N. B., & Gillenwater, J. M.** (1992). Infant auditory temporal acuity: Gap detection. *Child Development*, 63, 260–272.
- Widen, J. E.** (1993). Adding objectivity to infant behavioral audiometry. *Ear and Hearing*, 14(1), 49–57.
- Widen, J. E., & Keener, S. K.** (2003). Diagnostic testing for hearing loss in infants and young children. *Developmental Disabilities Research Reviews*, 9(4), 220–224.
- Yoshinaga-Itano, C., Sedey, A. L., Coulter, D. K., & Mehl, A. L.** (1998). Language of early- and later-identified children with hearing loss. *Pediatrics*, 102(5), 1161–1171.