

A Critical Review of Audiological Outcome Measures for Infants and Children

Trends in Amplification
15(1-2) 23-33
© The Author(s) 2011
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1084713811412056
http://tia.sagepub.com


Marlene P. Bagatto, AuD¹, Sheila T. Moodie, MCISc¹,
Richard C. Seewald, PhD¹, Doreen J. Bartlett, PhD²,
Susan D. Scollie, PhD¹

Abstract

Outcome evaluation is an important stage in the pediatric hearing aid fitting process, however a systematic way of evaluating outcome in the pediatric audiology population is lacking. This is in part due to the need for an evidence-based outcome evaluation guideline for infants and children with hearing loss who wear hearing aids. As part of the development of a guideline, a critical review of the existing pediatric audiology outcome evaluation tools was conducted. Subjective outcome evaluation tools that measure auditory-related behaviors in children from birth to 6 years of age were critically appraised using a published grading system (Andresen, 2000). Of the tools that exist, 12 were appraised because they met initial criteria outlined by the Network of Pediatric Audiologists of Canada as being appropriate for children birth to 6 years of age who wear hearing aids. Tools that were considered for the guideline scored high in both statistical and feasibility criteria. The subjective outcome evaluation tools that were ultimately chosen to be included in the guideline were the LittEARS Auditory Questionnaire (Tsiakpini et al., 2004) and the Parents' Evaluation of Aural/Oral Performance of Children (PEACH) Rating Scale (Ching & Hill, 2005b) due to the high grades they received in the critical review and their target age ranges. Following this critical review of pediatric outcome evaluation tools, the next step was for the Network Clinicians to evaluate the guideline (Moodie et al., 2011b).

Keywords

outcome measures, outcome evaluation, infant, child, hearing loss, hearing aids

Background

Pediatric audiologists share a common goal of providing infants and children who have permanent hearing loss appropriate access to early intervention. One component of intervention for many infants and children is access to sound through the use of hearing aids. Suitable technology and evidence-based hearing aid fitting protocols support accurate and safe hearing aid fittings (i.e., American Academy of Audiology [AAA], 2003; Bagatto, Scollie, Hyde, & Seewald, 2010; College of Audiologists and Speech Language Pathologists of Ontario [CASLPO], 2002; Early Hearing Equipment Advisory Group (2006; British Columbia Early Hearing Program [BCEHP]); King, 2010; "Guidelines," 2005]). This supports infants and children identified with permanent childhood hearing impairment (PCHI) in developing language and literacy skills. The aim of providing hearing aids is to improve functional auditory capacity and participation in hearing- and communication-specific situations. The provision of amplification is a process that includes the calculation of prescriptive targets based on accurate hearing assessment information, the selection of the physical and electroacoustic elements of a hearing aid, verification that the specified acoustical prescriptive targets

have been achieved, and outcome evaluation of device effectiveness in daily life. Of these stages, outcome evaluation does not currently have a systematic approach described in many pediatric hearing aid fitting protocols. The development of spoken language depends on the reception and transmission of information through the auditory channel. For a child with PCHI, this channel is impaired; therefore, the function of the auditory system with acoustic input should be monitored closely. There is little research related to what a typical outcome might be for an infant who wears hearing aids or how to track the child's auditory development and performance over time. This is in part due to the lack of well-developed outcome measures available for use with infants

¹National Centre for Audiology, Faculty of Health Sciences, University of Western Ontario, London, Ontario, Canada

²School of Physical Therapy, Faculty of Health Sciences, University of Western Ontario, London, Ontario, Canada

Corresponding Author:

Marlene P. Bagatto, AuD, National Centre for Audiology, University of Western Ontario, 1201 Western Road, Elborn College Room 2262J, London, Ontario, Canada, N6G 1H1
Email: bagatto@nca.uwo.ca

and children who wear hearing aids. Early steps in the hearing aid fitting process effect later steps and if not followed in a systematic way, they could impact the child's auditory, speech and language development. Receptive and expressive language development as well as speech perception and production are important aspects of outcome evaluation. Most pediatric hearing aid fitting protocols do, however, mention the importance of monitoring overall outcome even when specific strategies for doing so are not provided (e.g., AAA, 2003; Bagatto et al., 2010). Additionally, monitoring outcomes for infants at high risk of developing late-onset or progressive hearing impairment or those with PCHI who do not wear hearing aids (i.e., due to family choice) is an important aspect of pediatric audiology services. Both of these tasks would be supported by well-validated, clinically feasible monitoring protocols to track auditory development. Known clinical tools with good normative properties, validity, feasibility, and utility would support the development of an evidence-based outcome evaluation guideline for the pediatric audiology population. This purpose of this article is to review the current status of such tools, thereby identifying a subset that will be considered within a suggested guideline for their implementation (Moodie et al., 2011a). The sections below will present the various types of outcome measurements available, consider the properties to be appraised, and finally provide a critical review of available outcome evaluation tools within the category of caregiver-report questionnaires.

Types of Outcome Measures

Monitoring the hearing-related outcomes of infants and children with hearing loss can be accomplished both objectively and subjectively. One example of an objective measure is the use of aided sound field thresholds (ASFT). ASFT can be conducted in the sound field with the child wearing his or her hearing aids. This measures the child's aided ability to detect low-level sounds, and is considered an *objective* measure. Limitations of ASFT include the impact of room and hearing aid circuit noise, off-frequency listening with steeply sloping hearing losses, and patient responses to low-level sounds do not provide an indication of performance to moderate levels (Hawkins, 2004). Other examples of aided sound field testing are speech-sound discrimination and early measures of speech recognition which require the use of age-appropriate tests. Speech stimuli (e.g., Ling 6 sounds) can be included to obtain information about the infant's speech sound detection thresholds. Later, the child can be conditioned to discriminate between various speech sound patterns (i.e., "ahhhh" vs. "ah ah ah") at suprathreshold levels and ultimately perform speech recognition testing. This hierarchy of functional auditory assessment will provide more objective information about the infant's auditory skills. In contrast, questionnaires, diaries, and structured interviews

are examples of *subjective* ways to assess a child's auditory behaviors in real world environments. A combination of objective and subjective outcome evaluation tools may provide a multidimensional approach to tracking a child's auditory-related performance over time. A test battery of outcome evaluation tools provides caregivers and clinicians with a way to measure the auditory performance of an infant or child during the early months as well as later years of hearing aid use or nonuse (i.e., if the child has a known hearing loss but does not wear a device).

One advantage of objective measures is that they provide a direct measure of the child's hearing while wearing hearing aids and can therefore be used as a way to determine the impact of the intervention. In cases in which the child's ability to make use of aided sound is in question, for example children with auditory neuropathy spectrum disorder (ANS), this may provide critical information for the management of the child. Disadvantages of objective speech recognition testing are that the specific measurement technique and stimuli that are appropriate to use with a child of a given age and developmental level vary considerably. For an infant, early measurement techniques described in the literature focus on gross abilities such as detection or discrimination of large contrasts (e.g., visual reinforcement assessment of the perception of speech pattern contrasts [VRASPAC]; Martinez, Eisenberg, Boothroyd, & Visser-Dumont, 2008); later measures may focus on more complex tasks such as word or sentence recognition in closed or open set tasks (e.g., Bamford-Kowal-Bench Sentences in Noise [BKB-SINTM]; Etymotic Research, 2005). Although the need to increase the complexity of speech tasks is encouraging because it reflects the child's progress and development, it also means that an age-appropriate protocol for the use of objective measures requires careful consideration of the hierarchy of tasks, including how this hierarchy should be applied to children with typical development versus developmental delays. Objective measures may be difficult to obtain in cases of children with complex factors (e.g., difficult to test, speak languages other than those of the tests used, and so on). These same children may also present assessment and/or management difficulties more generally. In the early years, clinicians expend exorbitant efforts to obtain an audiogram from some children. Objective outcome measurement occupies the same equipment (e.g., test booth), the same child state (e.g., alert, cooperative, responsive), and the same clinician state (e.g., at the equipment, engaged with the child in a structured test procedure). Objective speech tests overlap with the basics of getting a full test of hearing sensitivity and getting the hearing aid fitting individualized, for example. Focusing on objective strategies as the primary strategy for outcome evaluation, therefore, is not likely to be successful on those very cases in which outcome measures are needed the most.

In contrast, caregiver reports can be done while caregivers are sitting and waiting for the clinician to execute hearing tests

or simulated real-ear verification procedures and therefore hold the possibility of adding information to the process without fully adding time and space requirements to the situation. Therefore, subjective measures may seem like less of a barrier in some instances. Finally, objective measures of speech detection and recognition only tell us about performance within the highly controlled acoustic conditions of a sound booth. They do not indicate how the caregiver perceives the auditory abilities of his or her child, or how the child performs in real world environments that include competition, distance, and interactive communication. Subjective measures focus on the child's responses to various sounds in real-life situations, as reported by the caregiver. Practically speaking, some administration barriers may arise with caregiver reports. For example, questionnaires are more appropriately administered in the native language of the family and there may be challenges for caregivers who have literacy issues (Johnson & Danhauer, 2002). These barriers can be overcome through the use of questionnaires in various languages or administering the tool interview style. Overall this type of outcome measurement provides rich and important information that can support the more objective tests that clinicians perform as well as being more applicable to children with complex needs. Therefore, this critical review focused on the evaluation of subjective outcome evaluation tools that assess auditory-related behaviors in infants and children.

As previously noted, there are many clinically relevant tools for the pediatric population with hearing loss that have incorporated rigor in their design, have compelling face validity, and/or that have been evaluated for reliability and validity, as required for inclusion in an evidence-based guideline. A *critical review* is characterized by an extensive review of the literature and critical evaluation of its quality (Grant & Booth, 2009). It goes beyond a simple description to include the degree of analysis and a conceptual innovation resulting in a hypothesis or model (Grant & Booth, 2009). Therefore, the development of an outcome evaluation guideline involved a review of the literature related to pediatric subjective outcome evaluation tools. This was followed by an assessment of the relevant tools, using a specific grading system, to support the inclusion of the chosen measures in a guideline. This article describes the review of the literature including the grading system that was used, the tools that were graded, and the outcome of the critical review. The subjective outcome evaluation tools chosen from the critical review are included in a guideline that will be described in detail in the final article of this issue (Bagatto et al., 2011).

Characteristics of a Good Outcome Evaluation Tool

Several researchers have described criteria for assessing the quality of outcome evaluation tools in rehabilitation (Andresen, 2000; Cox et al., 2000; Hyde, 2000). For example, a good

outcome evaluation tool should have *conceptual clarity* to ensure that it covers the relevant domains intended to be measured. Additionally, *normative data* for comparison purposes are a valuable aspect of any outcome evaluation tool. Published norms allow the clinician to compare the results obtained from the tool to standards for normal hearing and hearing impaired children. The *measurement model* of a good quality tool should be able to capture the true breadth and detail of the differences in the group being measured. Tools that consistently result in responses at the bottom (i.e., floor) or top end (i.e., ceiling) of the scale are not measuring the true range of the population being assessed. The outcome evaluation tool should *not have bias* either within the items or the instrument as a whole; the responses should not be affected by differences in culture or social circumstances. Statistically, the tool should have good test-retest *reliability*, internal consistency, *validity*, and *responsivity*. Of equal importance is the feasibility and utility of the outcome evaluation tool so that it is more likely to be implemented in clinical practice (Andresen, 2000; Graham et al., 2006). Therefore, excessive *respondent and administrative burden* should be avoided; the length and the content should be acceptable to the respondent and the tool should be reasonable to administer, score, and interpret by the clinician. In addition, the tool should have alternative modes of administration (i.e., electronically, brail) and/or language adaptations for different cultures, if possible.

With these characteristics in mind, subjective outcome evaluation tools for infants and children with PCHI were examined. Based on a system developed by Andresen (2000), operational definitions of grades were used in appraising a variety of auditory-related pediatric outcome evaluation tools. This system has been used to evaluate disability outcome evaluation tools for children and youth, such as the ABILITIES index and the Gross Motor Function Measure (Lollar, Simeonsson, & Nanda, 2000). The result of this analysis was a report card, in which each outcome evaluation tool received a grade, on each appraisal criterion, of A, B, C, or U (unknown). This type of analysis provides a brief yet detailed comparison of outcome evaluation tools across appraisal criteria, supporting a critical review. Such information is not currently available for outcome evaluation tools used to assess the performance of children with permanent hearing impairment. A detailed description of the appraisal criteria, as well as the grading system for each criterion as it applies to pediatric audiology is presented in Table 1.

Critical Review Objectives

Although there are several outcome evaluation tools available for the pediatric population, the intention was to evaluate tools that met the needs of the population identified by the Network of Pediatric Audiologists of Canada: birth to 6 years of age who wear hearing aids (see Moodie et al.,

Table 1. Appraisal Criteria as well as the Grading System for Each Criterion as it Applies to Pediatric Audiology

Characteristic	Description	Grade criteria
Conceptual clarity	Tool covers relevant domains intended to be measured (e.g., detection, localization, speech understanding)	A = Completely covered B = Adequately covered C = Inadequately covered
Norms and standard values	Large scale normative data for infants and children with normal hearing and PCHI. Experimental data collected using the tool is also considered given the lack of large scale norms available	Published data are available from: A = A large number infants and children with normal hearing and with PCHI who wear hearing aids B = A large number of infants and children with normal hearing C = Experimental data using the tool with infants and children with normal hearing and PCHI who wear hearing aids
Measurement model	There should not be ceiling or floor effects in measurement, particularly when used to measure the abilities of children with hearing loss	A = No issues B = Few or marginal evidence of skewing C = Substantial skewing
Item/instrument bias	The tool, and items within it, must not show evidence of bias when used with children who have PCHI. Bias-free tools have been evaluated on population subgroups and/or have had the response scale of the tool evaluated with Rasch analysis	A = Tool/items have been reviewed by parents of children with PCHI and acceptability is published OR Rasch analysis is good B = Adequate face validity to support low bias OR factor analysis is good/Rasch analysis shows some issues C = Bias is evident or tested OR inadequate statistical analysis
Respondent burden	The tool should be brief and clear enough for the caregiver to complete. The terminology used should not be offensive to those with hearing loss or deafness	The tool is: A = Brief (≤ 15 min) and has high acceptability for caregiver B = Either appropriately longer or some reported problems of acceptability C = Lengthy and acceptability is problematic
Administrative burden	The tool should be easy to administer, score, and interpret	The tool is: A = Scored by hand and the resulting metric is relevant and interpretable for the clinician and caregiver B = Scored by a computer and interpretation is obscure C = Costly and complex scoring; interpretation by another professional required
Reliability	The tool should give consistent results, within itself, and across time and testers	Internal consistency coefficient alpha: A $\geq .80$; B $< .80, > .70$; C $< .70$ Retest intraclass correlation coefficient: A $\geq .75$; B $> .40, < .75$; C $\leq .40$
Discriminant validity	The scores should differ for two subgroups of the population who would be expected to have different scores (e.g., normal hearing vs. hearing impaired children, on some items related to hearing)	A = Strong, expected direction, supported by clinical evidence B = Moderate or conflicting evidence C = Weak or based solely on statistical evidence
Convergent (criterion-related) validity	The tool should have been validated against a gold-standard measure, and/or the subscale structure of the tool has been statistically evaluated	A = Correlation of $\geq .60$; confirmed factor structure B = Correlations of $> .30, < .60$; few problems with factor structure C = Correlation of $\leq .30$; weak or not confirmed factor structure
Ecological validity*	The tool evaluates the child's responses within the context of specific, realistic environments and assesses the child as an active participant	A = Specific, realistic environments assessed B = Some situations are applicable and realistic for the child C = Environments are unrealistic and nonspecific

(continued)

Table 1. (continued)

Characteristic	Description	Grade criteria
Responsiveness	The scores on this tool have been shown to change, in the expected direction, when important changes are made to hearing status, hearing aid intervention, or therapy	Criteria for change are: A = Strong, supported by patient and clinical evidence B = Moderate or conflicting evidence C = Weak or based solely on statistical evidence
Alternate/accessible forms	The tool has been experimentally evaluated for use with different administration formats (e.g., paper and pencil vs. computer-assisted vs. interview-format administration)	A = Appropriate or varied modes are available and have been tested B = Some accommodations or testing among caregiver of children with PCHI C = No accommodations or mode information for special groups
Culture/language adaptations	The tool has been adapted and reevaluated for use with different languages and/or cultures (e.g., translations, use within deaf culture, with those who are deaf/blind)	A = Evidence of testing and applicability for cultural subgroups and interpretations B = Evidence of translations or testing with subgroups; some problems C = No evidence of testing or applicability to groups

Source: Adapted from Andresen (2000).

*Not from Andresen (2000) criteria

2011b). In addition, administration of the outcome evaluation tools by the audiologist to the caregiver at follow-up appointments will be an important aspect of this guideline. This will facilitate the caregivers becoming good observers of their child's listening behaviors while also allowing them to share a common language with their audiologist. The outcome evaluation tools will assist with reevaluating the previous stages of the amplification process, evaluating the overall impact of the hearing aid fitting, and sharing this outcome with the family in a systematic way. The following section will describe the procedure used to grade each outcome evaluation tool with the goal being to identify the best tools for inclusion in a guideline for the population identified.

Data Collection and Critical Review

Search Strategy

Subjective outcome evaluation tools that measure auditory-related behaviors for the pediatric population were located in several domains including health-related electronic databases (CINAHL, PubMed; 2008 and 2009), visually scanning reference lists from relevant studies, hand-searching key journals and conference proceedings, searching relevant internet resources, contacting experts in the area including the Network of Pediatric Audiologists of Canada, and citation searching. Key words used for searching included outcome evaluation, pediatric, infant, child, questionnaires, checklists, auditory development, auditory performance, hearing, hearing loss, and hearing aids. Various combinations of these key-words were used in the search domains. When a relevant tool or reference was obtained, the selection criteria listed below

were applied. If the tool met the criteria, it was included in the review.

Selection Criteria

As noted, early hearing detection and intervention (EHDI) programs are in need of high-quality outcome evaluation tools for infants and children from birth to 6 years of age. With this in mind, the following selection criteria were applied to the available pediatric outcome evaluation tools prior to including them in the review:

- Age range = birth to 6 years
- Questionnaire- or interview-based
- Parent/caregiver respondent
- Audiologist administered and scored
- Auditory-related outcomes measured
- Application to infants and children who wear hearing aids

Tools were selected by the first author based on the stated criteria. The tools selected for critical review along with a brief description of each are listed in Table 2.

Critical Evaluation

The outcome evaluation tools identified through the review process were graded for each characteristic listed in Table 1 using the grading system described by Andresen (2000). The first author carried out all grading and presented the results to the second author and modifications were made when necessary to come to agreement. As specified in Table 1, a grade of "A" is the highest and was assigned only when high-quality

Table 2. Subjective Outcome Evaluation Tools Selected for Critical Review Along With a Brief Description of Each

Outcome evaluation tool	Description					Reference
	Number of items	Response format	Scoring format	Age range	Factors assessed	
Auditory Behavior in Everyday Life (ABEL)	24	7-point scale	Subscale and overall averages	4 to 14 years	Aural-Oral, auditory awareness, social/conversational	Purdy et al (2002)
Children's Home Inventory for Listening Difficulties (CHILD)	15	8-point scale	Total and overall average	3 to 12 years	Understanding sound at home	Anderson and Smaldino (2000)
Children's Outcome Worksheet (COW)	5	5-point scale	Overall average	—	Individually defined needs and outcomes	Williams (2004)
Client Oriented Scale of Improvement – Child Version (COSI – C)	3 to 5	5-point scale	Degree of change, overall average	>0	Parent-defined goals	National Acoustics Laboratories
Developmental Index of Audition and Listening (DIAL) / Family Expectations Worksheet (FEW)	3 to 5	5-point scale	Degree of change, overall average	Birth to 22 years	Auditory behaviors, organized in a developmental hierarchy	Palmer and Mormer (1999)
Early Listening Function (ELF)	12	Yes/maybe/no	Complex	Birth to 3 years	Furthest distance at which the child consistently responds in real life	Anderson (2000)
Functional Auditory Performance Indicators (FAPI)	31	Not present/emerging/in process/acquired	Sum score per category	Birth to childhood	Seven categories of auditory behaviors, in developmental order	Stredler-Brown and Johnson (2003)
Hearing Aid Benefit Scale for Infants / Toddlers (HABIT)	10	3-point scale	Not specified	Birth to 3 years	Hearing aid benefit	Geier (1998)
Infant-Toddler Meaningful Auditory Integration Scale (IT-MAIS)	10 probes	Parental observation plus structured interview	Overall score (based on examples given)	Older infancy through childhood	Vocalization behavior Alerting to sounds Meaning from sound	Zimmerman-Phillips et al (2000)
LittLEARS Auditory Questionnaire	35	Yes/no	Total of "Yes" responses	Birth to 24 months	Three categories of auditory behaviors, organized in a developmental hierarchy	Tsiakpini et al (2004)
Parents' Evaluation of Aural/Oral Performance of Children (PEACH)	Diary	Parental observation plus structured interview	Subscale and overall percentages (based on examples given)	Infancy through childhood	Hearing aid use, loudness discomfort, communication in quiet and noise, phone use, environmental sounds	Ching and Hill (2005a)
	Rating scale	5-point rating scale	Subscale and overall percentages			Ching and Hill (2005b)

evidence existed that the tool met the accepted standards for good performance. This was followed by Grades “B” and “C”, or Grade “U” if published data for evaluation did not exist. The results of the evaluation of each tool are summarized in Table 3.

Results

Twelve auditory-related subjective pediatric outcome evaluation tools were identified through the search process and subjected to the grading process (Table 2). Of these tools, seven use a rating scale or yes/no response format (e.g., ABEL, CHILD, ELF, FAPI, HABIT, LittEARS, PEACH Rating Scale); three use a goal-setting and assessment format (e.g., COW, COSI-C, DIAL); and two use a caregiver interview response format (e.g., IT-MAIS, PEACH Diary). Each of these tools were evaluated against the appraisal criteria shown in Table 1. The evaluations are discussed in further detail below, within the general categories of conceptual clarity, norms, measurement model, item/instrument bias, respondent and administrative burden, reliability, different types of validity, responsiveness, alternate/accessible forms and language adaptations.

Conceptual Clarity

The majority of the tools received an “A” or “B” grade on the conceptual clarity domain, indicating that the relevant domains intended to be measured were covered by the tool. The tools that received an “A” grade (i.e., CHILD, DIAL, FAPI, LittEARS) covered the relevant content domains well by containing many items that thoroughly cover auditory-related content. Those that received a “B” grade (i.e., ABEL, COW, ELF, HABIT, IT-MAIS, PEACH Diary, PEACH Rating Scale) were rated to have not adequately covered the relevant content domains because they had fewer items that did not completely address as much auditory-related content. The COSI-C (National Acoustics Laboratories) received a “C” grade due to the fact that the goals are set collaboratively by the audiologist and caregiver and there were no examples provided as with the COW (Williams, 2004).

Normative Values

Normative values gathered from a large group of infants and children with normal hearing and PCHI who wear hearing aids are available for the PEACH Diary (Ching & Hill, 2005a), therefore the tool was assigned a grade of “A” for normative values. The LittEARS Auditory Questionnaire (Tsiakpini et al., 2004) received a grade of “B” because the authors gathered norms from 218 normal hearing infants and children from German-speaking families to create their normative data. Many of the tools did not have normative values gathered from a large scale study with which to compare individual children’s scores for clinical interpretation and

utilization of the tool (e.g., ABEL, CHILD, COW, COSI-C, DIAL, ELF, FAPI, PEACH Rating Scale¹). Both the HABIT (Geier, 1998) and the IT-MAIS (Zimmerman-Phillips et al., 2000) received a “C” grade for reporting on experimental rather than large scale clinical data gathered using the tool on children with normal hearing and PCHI with a hearing device.

Measurement Model and Item/Scale Bias

Information regarding the measurement model and item/scale bias was typically not available for the outcome evaluation tools that were reviewed (e.g., ABEL, CHILD, COW, COSI-C, DIAL, ELF, FAPI, PEACH Rating Scale¹). The HABIT, IT-MAIS, LittEARS and PEACH Diary received grades of “A” or “B” for their data regarding ceiling or floor effects (i.e., measurement model) within these tools and the LittEARS and PEACH Diary received “A” grades for reporting good acceptability and/or Rasch analysis of the items (i.e., no item/scale bias) within the questionnaire (Ching & Hill, 2005a; Tsiakpini et al., 2004).

Respondent and Administrative Burden

Respondent and administrative burden were assessed either through publications, the current authors’ clinical experiences with the tool, and/or expert reports from members of the Network of Pediatric Audiologists of Canada. During a focus group meeting of the Network Audiologists many reported that time was one of the main barriers to routine outcome evaluation in their clinical practice. They preferred tools that did not take up too much of the caregiver’s or clinician’s time, and discussed that a 10-min duration for this procedure may be feasible. In addition to time, interview-based scoring can contribute to administration and respondent burden and therefore variability with scores. A study looking at the relationship of cortical evoked potentials and functional measures in infants with hearing loss found the results of the PEACH Diary to be highly variable (Golding et al., 2007). The authors indicated that the caregiver’s ability to observe their child varied and may have been limited by competing factors in the household (i.e., number of children, wellness of the child, lifestyle). Golding and colleagues (2007) also noted that an inexperienced interviewer may have had difficulty extracting useful examples from the parents even though the interviewer received instructions on how to administer the PEACH. This observation was also noted in a research study conducted in the UWO Child Amplification Laboratory (CAL; S. Scollie, personal communication, Ching et al., 2010). Therefore, tools that required lengthy interviews and/or scoring were given a “C” grade because they were too lengthy and not widely accepted either by the caregivers or clinicians (i.e., IT-MAIS, PEACH Diary). Outcome evaluation tools that performed well in terms of their lack of respondent and administrative burden were the ABEL (Purdy, Farrington, Moran, Chard, &

Table 3. Grade Report for Each Outcome Evaluation Tool Assessed in This Critical Review. A Grade of “A” is the Highest and was Assigned Only When High-Quality Evidence Existed That the Tool Met the Accepted Standards for Good Performance. This was Followed by Grades “B” and “C”, or Grade “U” if Data for Evaluation did not Exist

	Outcome evaluation tool										PEACH	PEACH Rating
	ABEL	CHILD	COW	COSI-C	DIAL	ELF	FAPI	HABIT	IT-MAIS	LittLEARS	Diary	Scale
Conceptual clarity	B	A	B	C	A	B	A	B	B	A	B	B
Normative data	U	U	U	U	U	U	U	C	C	B	A	U
Measurement model	U	U	U	U	U	U	U	A	B	A	B	U
Item/Scale bias	U	U	U	U	U	U	U	U	U	A	A	U
Respondent burden	A	B	B	B	B	C	B	A	C	A	C	A
Administrative burden	A	A	B	B	C	C	B	A	C	A	C	B
Retest reliability	A	B	U	U	U	U	U	A	B	A	A	U
Discriminant validity	U	U	U	U	U	U	U	A	U	B	B	U
Convergent validity	A	C	U	U	B	C	C	A	A	A	B	U
Ecological validity	A	B	A	A	A	A	A	A	B	A	A	A
Responsiveness	A	B	U	U	U	U	U	A	B	U	A	U
Alternate/accessible forms	C	B	B	B	C	B	C	C	B	C	B	B
Other languages	C	C	C	C	C	C	C	C	C	A	B	C

Hodgson, 2002), CHILD (Anderson & Smaldino, 2000), HABIT, LittLEARS and PEACH Rating Scale. These tools had a reasonable number of items with either a yes/no or rating response format that was scored in a straightforward manner and did not require lengthy interviews to complete the tool.

Reliability, Validity and Responsivity

The authors of the ABEL, CHILD, HABIT, IT-MAIS, LittLEARS and PEACH Diary reported good reliability of their outcome evaluation tool and the grades in Table 3 reflect this. Discriminant validity was either strong or moderate on the HABIT, LittLEARS and PEACH Diary and were assigned either a grade of “A” or “B”. The remaining tools did not have data available for this characteristic and were assigned a “U” grading. Other than the goal-setting tools (e.g., COW, COSI-C), the majority of the tools evaluated had good to excellent convergent validity. Ecological validity was also good to excellent for the outcome evaluation tools assessed in this critical review. The responsiveness of the ABEL, CHILD, HABIT, IT-MAIS, and PEACH Diary were assessed and received an “A” or “B” grade. The remaining tools did not have responsiveness data available at the time of this review.

Alternate/Accessible Forms and Language Adaptations

Alternate and/or accessible forms were available for a good portion of the questionnaires as many are now available online or in computer software format. The final category that was evaluated was availability in other languages. The LittLEARS and PEACH Diary received the highest grades for having the tools available in other languages; the LittLEARS is

available in 19 languages and the PEACH Diary is available in six.

Overall Grades

Overall, the HABIT, IT-MAIS, LittLEARS and PEACH Diary received “A” or “B” grades for the majority of the reviewed characteristics. Although the HABIT is applicable for the infant population, has low respondent and administrative burden and high reliability, validity and sensitivity, the main limitations are that the normative data are lacking and the questionnaire is an unpublished doctoral dissertation rendering it virtually unknown to the clinical community. The IT-MAIS is more widely available, however large scale norms are not provided for English-speaking normal hearing or hearing impaired infants with hearing aids. Additionally, the interview format of the IT-MAIS increases the respondent and administrative burden which may influence the feasibility and utility of the questionnaire which may ultimately impact the clinical uptake of the tool (Andresen, 2000; Graham et al., 2006). The LittLEARS received high grades on most characteristics and is accessible to the clinical community for a fee. The PEACH Diary has large scale normative values for normal hearing and hearing impaired infants, which increases the clinical utility of the tool. However, the PEACH Dairy’s interview-style format introduces the same clinical feasibility and utility concerns as the IT-MAIS. For this reason, the PEACH Rating Scale may be more successfully used in a clinical setting provided the statistical characteristics from the PEACH Diary can be applied to the items in the PEACH Rating Scale. The items in the two PEACH tools are extremely similar, but the administration format of the tool (interview/diary vs. ratings only) differs significantly.

In light of this critical review, the LittleEARS Auditory Questionnaire and the PEACH Rating Scale scored most favorably in the majority of the review categories. To ensure the target age range from birth to 6 years is properly represented for the outcome evaluation guideline, both the LittleEARS and PEACH Rating Scale were chosen to be included. The LittleEARS targets children from birth through the first two years of hearing and the PEACH items appear to target toddlers and older children. Therefore, it is possible that a guideline could provide a two-stage process whereby the LittleEARS is used with caregiver of young infants until they reach a ceiling score and/or age on the tool. This would indicate a certain level of auditory development has occurred within the infant and the child will be developmentally ready to be evaluated by the items in the PEACH Rating Scale. These and other administration issues will be further addressed in the description of the guideline and supporting data provided in the final article in this issue (Bagatto et al., 2011).

Conclusions

A critical review of auditory-related pediatric subjective outcome evaluation tools was completed as part of the development of an outcome evaluation guideline. Although there are many subjective tools available for the pediatric population, few have the relevant psychometric and/or feasibility characteristics necessary to promote clinical uptake within a guideline. Prior to considering a caregiver report questionnaire within a guideline, a review of the existing outcome evaluation tools for infants and children aged birth to 6 years followed by a systematic grading of the tools was necessary. Twelve outcome evaluation tools with specified criteria were identified prior to assigning grades for thirteen psychometric and feasibility characteristics (Andresen, 2000). Results indicated that four out of the 12 tools received high grades in most of the characteristics and of these four, only two would be considered clinically feasible within an outcome evaluation guideline for infants and young children. Based on these results, the LittleEARS Auditory Questionnaire and the PEACH Rating Scale were considered for inclusion in an outcome evaluation guideline. The next step in the guideline development process was to consult with the Network of Pediatric Audiologists of Canada and have them systematically evaluate the chosen questionnaires. Moodie and her colleagues (2011) provide the results of this evaluation.

List of Abbreviations:

AAA: American Academy of Audiology
 ABEL: Auditory Behavior in Everyday Life
 ANSD: Auditory neuropathy spectrum disorder
 ASFT: Aided sound field thresholds
 BCEHP: British Columbia Early Hearing Program

BKB-SIN: Bamford-Kowal-Bench Sentences in Noise
 CAL: Child Amplification Laboratory
 CASLPO: College of Audiologists and Speech Language Pathologists of Ontario
 CHILD: Children's Home Inventory for Listening Difficulties
 COSI-C: Client Oriented Scale of Improvement—Child Version
 COW: Children's Outcome Worksheet
 DIAL: Developmental Index of Audition and Listening
 EHDI: Early hearing detection and intervention
 ELF: Early Listening Function
 FAPI: Functional Auditory Performance Indicators
 FEW: Functional Expectations Worksheet
 HABIT: Hearing Aid Benefit Scale for Infants and Toddlers
 IT-MAIS: Infant-Toddler Meaningful Auditory Integration Scale
 PCHI: Permanent childhood hearing impairment
 PEACH: Parents' Evaluation of Aural/Oral Performance of Children
 VRASPAC: Visual Reinforcement Assessment of the Perception of Speech Pattern Contrasts

Acknowledgments

The authors gratefully acknowledge the support of the Early Learning and Child Development Branch and the Children's Corporate Systems Branch of Ontario's Ministry of Children and Youth Services in Canada as well as Kelley Keene for her help in the preparation of this manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported with funding by the Canadian Institutes of Health Research [Marlene Bagatto: 200811CGV-204713-174463 and Sheila Moodie: 200710CGD-188113-171346] and the Ontario Research Fund, Early Researcher Award to Susan Scollie, Siemens Hearing Instruments, Canada and The Masonic Foundation of Ontario, Help-2-Hear Project.

Note

1. It is possible that the PEACH Diary characteristics could be used for PEACH Rating Scale. See Bagatto et al., (2011).

References

American Academy of Audiology. (2003). *American Academy of Audiology pediatric amplification guidelines*. Retrieved from <http://www.audiology.org/resources/documentlibrary/Documents/pedamp.pdf>

- Andresen, E. M. (2000). Criteria for assessing the tools of disability outcomes research. *Archives of Physical Medicine & Rehabilitation*, 81, Suppl. 2, S15-S20.
- Anderson, K. L. (2000). Early listening function. Retrieved from http://www.kandersonaudconsulting.com/uploads/ELF_Questionnaire.pdf
- Anderson, K. L., & Smaldino, J. J. (2000). Children's home inventory of listening difficulties. Retrieved from: http://www.kandersonaudconsulting.com/uploads/child_questionnaire.pdf
- Bagatto, M.P., Moodie, S.T., Malandrino, A.C., Richert, F.M., Clench, D.A. & Scollie, S.D. (2011). The University of Western Ontario Pediatric Audiological Monitoring Protocol (UWO PedAMP), *Trends in Amplification*, 15(1-2)
- Bagatto, M. P., Scollie, S. D., Hyde, M. L., & Seewald, R. C. (2010). Protocol for the provision of amplification within the Ontario Infant Hearing Program. *International Journal of Audiology*, 49, S70-79.
- Ching, T. Y., & Hill, M. (2005a). *The Parents' Evaluation Of Aural/Oral Performance Of Children (PEACH) diary*. Chatswood, New South Wales, Australia: Australian Hearing. Retrieved from http://www.nal.gov.au/outcome-measures_tab_peach.shtml
- Ching, T. Y., & Hill, M. (2005b). *The Parents' Evaluation of Aural/Oral Performance of Children (PEACH) Rating Scale*. Chatswood, New South Wales, Australia: Australian Hearing. Retrieved from <http://www.outcomes.nal.gov.au/LOCHI%20assessments.html>
- Ching, T. Y., Scollie, S. D., Dillon, H., Seewald R, Britton L, Steinberg J, . . . King, K. A. (2010). Evaluation of the NAL-NL1 and the DSL v4.1 prescriptions for children: Paired-comparison intelligibility judgments and functional performance ratings. *International Journal of Audiology*, 49, S35-S48.
- College of Audiologists and Speech-Language Pathologists of Ontario. (2002). *Preferred practice guideline for the prescription of hearing aids to children*. Retrieved from <http://www.caslpo.com/Portals/0/ppg/preshearingaidschild.pdf>
- Cox, R., Hyde, M., Gatehouse, S., Noble, W., Dillon, H., Bentler, R., . . . Hallberg, L. (2000). Optimal outcome measures, research priorities and international cooperation. *Ear & Hearing*, 21(4), 106S-115S.
- Early Hearing Equipment Advisory Group. (2006). *Hearing equipment protocol from British Columbia Early Hearing Program*. Retrieved from <http://www.phsa.ca/NR/rdonlyres/06D79FEB-D187-43E9-91E4-8C09959F38D8/40116/aHearingEquipment-Protocol.pdf>
- Etymotic Research. (2005). *BKB-SINTM speech in noise test version 1.03* (compact disc). Elk Grove Village, IL: Etymotic Research.
- Geier, A. K. (1998). *Development of the Hearing Aid Benefit Scale for Infants and Toddlers*. Unpublished doctoral dissertation, Central Michigan University, Mount Pleasant.
- Golding, M., Pearce, W., Seymour, J., Cooper, A., Ching, T. & Dillon, H. (2007). The relationship between Obligatory Cortical Auditory Evoked Potentials (CAEPs) and function measures in young infants. *Journal of the American Academy of Audiology*, 18, 117-125.
- Graham, I., Logan, J., Harrison, M.B., Straus, S.E., Tetro, J., Caswell, W., & Robinson, N. (2006). Lost in knowledge translation: Time for a map? *Journal of Continuing Education of Health Professionals*, 26, 13-24.
- Grant, M. J. & Booth, A. (2009). A typology of reviews: an analysis of 14 reviews types and associated methodologies. *Health Information & Libraries Journal*, 26, 91-108.
- Guidelines for the fitting, verification and evaluation of digital signal processing hearing aids within a children's hearing aid service. (2005). *Modernising children's hearing aid services*, University of Manchester, 1-3. Retrieved from <http://www.psych-sci.manchester.ac.uk/mchas/guidelines/fittingguidelines.doc>
- Hawkins, D. B. (2004). Limitations and uses of the aided audiogram. *Seminars in Hearing*, 25(1), 51-62.
- Hyde, M. L. (2000). Reasonable psychometric standards for self-report outcome measures in audiological rehabilitation. *Ear & Hearing*, 21, 245-365.
- Johnson, C. E., & Danhauer, J. L. (2002). Outcomes measurement and the audiologist. In C. E. Johnson & J. L. Danhauer (Eds.), *Handbook of outcomes measurement in audiology* (pp.1-14). Clifton Park, NY: Thomson Delmar Learning, Singular.
- King, A. M. (2010). The national protocol for paediatric amplification in Australia. *International Journal of Audiology*, 49, S64-69.
- Loller, D. J., Simeonsoon, R. J., & Nanda, U. (2000). Measures of outcomes for children and youth. *Archives of Physical Medicine & Rehabilitation*, 81(12), S46-52.
- Martinez, A., Eisenberg, L., Boothroyd, A., & Visser-Dumont, L. (2008). Assessing speech pattern contrast perception in infants: Early results on VRASPAC, *Otology and Neurology*, 29, 183-188.
- Moodie, S. T., Bagatto, M. P., Seewald, R. C., Kothari, A., Miller, L., & Scollie, S.D. (2011a). An integrated knowledge translation experience: Use of the Network of Pediatric Audiologists of Canada to facilitate the development of the University of Western Ontario Pediatric Audiological Monitoring Protocol (UWO PedAMP). *Trends in Amplification*, 15(1-2), 34-57.
- Moodie, S. T., Kothari, A., Bagatto, M.P., Seewald, R. C., Miller, L., & Scollie, S.D. (2011b). Knowledge translation in audiology: Promoting the clinical application of best evidence. *Trends in Amplification*, 15(1-2), 5-22.
- National Acoustics Laboratories. Client Oriented Scale of Improvement – Child Version (COSI-C). *Australian Hearing*, Retrieved on June 4, 2011 from <http://www.nal.gov.au/pdf/COSI-C-Questionnaire.pdf>
- Palmer, C. V., & Mormer, E. A. (1999). Goals and expectations of the hearing aid fitting. *Trends in Amplification*, 4(2), 61-71.
- Purdy, S. C., Farrington, D. R., Moran, C. A., Chard, L. L., & Hodgson, S. A. (2002). A parental questionnaire to evaluate children's auditory behavior in everyday life (ABEL). *American Journal of Audiology*, 11, 72-82.
- Stredler-Brown, A. & Johnson, C. D. (2004). *Functional auditory performance indicators: And integrated approach to auditory skill development*. Retrieved from http://www.cde.state.co.us/cdesped/download/pdf/FAPI_3-1-04g.pdf.

- Tsiakpini, L., Weichbold, V., Kuehn-Inacker, H., Coninx, F., D'Haese, P., & Almadin, S. (2004). *LittlEARS Auditory Questionnaire*. Innsbruck, Austria: MED-EL.
- Williams, C. N. (2004). COW—Designed with children in mind. *Hearing Journal*, 57(3), 68.
- Zimmerman-Phillips, S., Osberger, M. J., & Robbins, A. M. (2000). *Infant-Toddler Meaningful Auditory Integration Scale (IT-MAIS)*. Sylmar, CA: Advanced Bionics Corporation. Retrieved from http://www.advancedbionics.com/UserFiles/File/IT-MAS_20brochure_20_2.pdf