12-29-2016

Language Assessment in a Snap: Monitoring Progress Up to 36 Months

Jill Gilkerson
LENA Research Foundation

Jeffrey A. Richards
LENA Research Foundation

Charles R. Greenwood
University of Kansas

Judy K. Montgomery
Chapman University, montgome@chapman.edu

Recommended Citation

Gilkerson, J., Richards, J. A., Greenwood, C. R., & Montgomery, J. K. (2016). Language assessment in a snap: Monitoring progress up to 36 months. *Child Language Teaching and Therapy*, 0265659016660599. https://doi.org/10.1177/0265659016660599

Language Assessment in a Snap: Monitoring Progress Up to 36 Months

Language Assessment in a Snap: Monitoring Progress up to 36 Months

Jill Gilkerson[a,b], Jeffrey A. Richards[a],

Charles R. Greenwood[c], and Judy K. Montgomery[d]

RUNNING HEAD: LANGUAGE SNAPSHOT

In press – *Child Language Teaching and Therapy*

Affiliations: [a]LENA Research Foundation, 5525 Central Avenue Suite 100, Boulder, Colorado 80301. Email: jillgilkerson@lenafoundation.org; jeffrichards@lenafoundation.org; [b]University of Colorado at Boulder, Boulder, CO, 80309; [c]University of Kansas, Department of Applied Behavioral Science, Juniper Gardens Children's Project, 444 Minnesota Avenue Suite 300, Kansas City, Kansas 66101. Email: greenwood@ku.edu; [d]Chapman University, Communication Sciences and Disorders Program, One University Drive, Orange, California 92866. Email: montgome@chapman.edu

Address correspondence to: Jill Gilkerson, LENA Research Foundation, 5525 Central Avenue Suite 100, Boulder, Colorado 80301. Phone number: 303-545-9696; Email: jillgilkerson@lenafoundation.org

Abstract

This paper describes the development and validation of the Developmental Snapshot, a 52-item parent questionnaire on child language and vocal communication development that can be administered monthly and scored automatically. The Snapshot was created to provide an easily administered monthly progress monitoring tool that enables parents to better recognize language milestones and offer professionals prompt information to fine-tune intervention strategies. Initial items were piloted by 15 families; refinement and further development of the instrument was conducted with parents of 308 typically developing children. Reliability and criterion validity metrics were examined on subsets of approximately 60 children who completed the Snapshot on a monthly basis and were evaluated on standard SLP-administered assessments. Divergent validity was also examined for samples of children diagnosed with language delays related to ASD (n = 77) or not (n = 49). Results supported the criterion validity (r = .67-.97) and test-retest reliability claims of the Snapshot (r = .95). Sensitivity and specificity for language delay detection also were good at 87%. Potential applications for progress monitoring, fidelity of intervention, and enhancing parent awareness of their child's language and vocal communication changes are discussed.

Language Assessment in a Snap: Monitoring Progress up to 36 Months

Advances in our knowledge of the important role played by the home language environment as a social determinant of very young children's language learning are rapidly leading to interventions designed to reduce the 30 million word gap (Hart and Risley, 1995; Hart and Risley, 1999; Suskind et al., 2015). Based on these and other reports that disadvantaged children on average develop smaller vocabularies and hear vastly fewer words spoken to them from birth to 4 years compared to more advantaged families (Hart and Risley, 1992; Rowe, 2008; Huttenlocher et al., 1991; Hoff, 2003), the landscape of language intervention and assessment is changing (Shankar, 2014). For example, interventions at national and local community levels are focused on public messaging in the media regarding the importance of talking to your baby (Crow and O'Leary, 2015). Other messaging is focused on tips for parents and how to talk to your baby (The Family Conservancy, 2015). At the same time, structured child-level interventions focusing on parents and caregivers are being developed and evaluated (Roberts and Kaiser, 2011), and some are being scaled up for delivery by traditional state and local children's service systems such as home visiting and child care programs (Providence Talks, 2015).

Common to all such efforts is recognition of the critical role played by parents in their child's language environment, the need for parents to become active, informed partners in implementing change in their child's language environment and observing their child's language development over time. A typical expectation for parent intervention partners is that they learn to interleave language enrichment and communicative interaction strategies into daily routines known to accelerate child language growth (Warren and Brady, 2007; Landry et al., 2006). Reports indicate that parents who are coached in using these strategies regularly can achieve relatively rapid changes in child language development, both for at-risk typically developing

children and children with language delays and disabilities (Kaiser et al., 1996; Landry et al., 2006; Landry et al., 2008).

An important element in an informed parent partnership is gathering dependable, frequent information on an infant/toddler's vocal development over time and intervention – this process can help educate and encourage caregivers as well as inform parents' future actions. Further, a frequent monitoring approach can facilitate identifying children with language delays and/or highlight when a specific intervention is not having the intended effect on the child's language development over time. Although there is a need to identify children at-risk for language delays in the earliest stages of language development, and research suggests that early screening can effectively identify very young children who could benefit from intervention (Nelson et al., 2006), pediatricians, for example, have not systematically included language-focused screens in routine check-ups (Halfon et al., 2002). Thus, a need exists for valid and reliable measurement tools that allow parents to report their observations on child progress and development more frequently, that are sensitive to growth and change over time, and that potentially can facilitate identification of children who are at-risk for language delays. Such measures for parents should be brief, scored automatically without professional assistance, producing information that is intuitive and easily understandable by parents so as to support their ongoing efforts.

Many commonly used language development instruments are neither designed nor logistically appropriate for tracking continuous progress because of long administration time, the need for costly and time consuming professional scoring, or the length of time separating administrations (i.e., six months or longer). For example, the *MacArthur-Bates Communicative Development Inventories*, a widely used language assessment that incorporates a comprehensive vocabulary checklist, can take up to 40 minutes to complete (Fenson et al., 2007). The *Child*

Development Inventory, a 300-item parent questionnaire, requires professional expertise for scoring the seven individual subscales (Ireton, 1992) and can take up to 50 minutes to complete. Clinician administered assessments such as the Receptive Expressive Emergent Language Scale, Third Edition (REEL-3) (Bzoch et al., 2003) and the Preschool Language Scale, Fourth Edition (PLS-4) (Zimmerman et al., 2002), have been considered gold standards for early language evaluation, but they require clinical expertise and observation by speech/language professionals, added travel and professional costs, and are not intended for frequent administration. Although the observational progress monitoring Early Communication Indicator (ECI) (Greenwood et al., 2010) is designed to be conducted frequently, it is administered and scored by a trained professional, (e.g., home visitor), and feedback shared with the parent during regular visits (Greenwood et al., 2006; Greenwood et al., 2010).

The Developmental Snapshot (hereinafter Snapshot) was designed to address these issues. Our goal was to develop and validate a new measure of language and vocal communication development in children from birth to three years meeting the following criteria: 1) parent-completed, 2) brief enough to be completed within 15 minutes, 3) compatible with monthly administration and longitudinal tracking, 4) automatic scoring, and 5) generates a language development age and age-referenced standard score to facilitate flexible interpretation. Phase I of this research describes the initial item development, item order refinement and scoring approach. Phase II addresses the following research questions:

1. What was the test-retest reliability of the measure?
2. What was the criterion validity with respect to language and communication proficiency as measured by standardized assessments?

3. What was the divergent validity related to identification of children with language delays?

Method

General Design

Phase I of this research involved item development and refinement as well as initial pilot testing and scoring of the measurement tool, while Phase II addressed the Snapshot's psychometrics (i.e., reliability, validity and identification of language delay). The data examined here were drawn from those collected during various phases of the LENA Research Foundation's *Natural Language Study* (*NLS*), an ongoing research and data collection effort designed to collect information about various aspects of child development and the early language environment (see Gilkerson & Richards (2008); Oller et al. (2010) for more detailed information on the *Natural Language Study*). Participants (see below) were primarily parents of typically developing (TD) children between birth and 36 months of age in the study. During development (Phase 1), parents of TD children older than 36 months were included to determine ceiling effects for the question set. As part of the validation analyses (Phase 2), samples of children diagnosed with language delays (LD) related or not to Autism Spectrum Disorder (ASD) were included as well.

Criterion Validity Measures

Parent report/questionnaires. To measure receptive/expressive vocabulary development, parents of children ages 8-30 months completed the *MacArthur-Bates Communicative Development Inventory - Words and Gestures*/*Words and Sentences* (MBCDI; Fenson et al., 2007), a checklist asking about words the child understands and/or says. The *Child Development Inventory* (Ireton, 1992) is a 300-item parent questionnaire for children over 15 months that

includes 8 subscales (e.g., social, gross motor, fine motor, expressive and receptive language) as well as a general development scale. Since there were no standardized parent questionnaires available for very young infants, parents of children 2-6 months of age completed a phone interview with a certified Speech Language Pathologist (SLP) using the *Receptive Expressive Emergent Language Test, 3rd Edition* (REEL-3; Bzoch et al., 2003); although the REEL-3 is an observational tool traditionally administered in person, the items focusing on early infant development are largely parent report and conducive to an interview format. Receptive and expressive language scores from the REEL-3 and the Child Development Inventory, as well as the verbal production (vocabulary) score from the MBCDI were used for analyses.

Professional language evaluations. Children were evaluated by a certified SLP utilizing up to three standard observational assessments. The number and type administered during each session varied with child age and attention span and with time constraints. The core battery that participants completed typically included the *REEL-3*, the *Preschool Language Scale, 4th Edition* (PLS-4; Zimmerman et al., 2002) , and the *Cognitive Adaptive Test/Clinical Linguistic and Auditory Milestone Scale* (CAT/CLAMS; Accardo and Capute, 2005). Note that the REEL-3 was administered with children older than 6 months of age in its traditional modality, as a professional observational assessment (as opposed to a phone interview). Parents received no feedback on assessment results. PLS-4 and REEL-3 receptive and expressive language scores and CAT and CLAMS cognitive and language scores were used for analyses. SLPs completing these assessments were not informed of the Snapshot results.

Phase 1: Snapshot Development

Item Development Participants

Participants were 15 mothers of children 3-41 months of age recruited through word of

mouth and email invitation. This was a sample of convenience; ten of the participant mothers had a college degree, four had an associate's degree or had attended some college, and one had a high school diploma. All participants were living in monolingual English-speaking households, and 13/15 (87%) were white of non-Hispanic origin.

Item Development Procedures

Item development was led by two of this study's authors: a linguist who specializes in early language acquisition (Gilkerson), and a certified speech language pathologist (SLP) with over 30 years of clinical experience (Montgomery). Domains of focus included: vocal behavior and preverbal communication for infants under 12 months; responsiveness to instruction, spontaneous speech production and vocabulary development for one-year-olds; and conceptual and grammatical development for children over 24 months. To generate the initial items we reviewed ten commonly used professional language evaluation tools (e.g., Bayley-III, Mullen) and parent report questionnaires (MBCDI, Ages & Stages) to identify major concepts repeated across assessments and developmental phases, e.g., nonverbal behavior, preverbal vocalization, vocabulary and grammar development and semantic concept formation.[1]

---

[1] Reviewed assessments included: *The Mullen Scales of Early Learning AGS Edition, Rossetti Infant-Toddler Language Scale, Preschool Language Scale Fourth Edition, Receptive-Expressive Emergent Language Test-Third Edition, Child Development Inventory, Cognitive Adaptive Test/Clinical Linguistic and Auditory Milestone Scale (CAT/CLAMS), Bayley Scales of Infant and Toddler Development Third Edition, The MacArthur-Bates Communicative Development Inventories – Words and Gestures/Words and Sentences, Ages and Stages Questionnaire Second Edition, Child Guidance Center*.

To facilitate parental monitoring, we sought to include items associated with easily observable child behaviors, such as turning one's head in response to the child's name, naming objects in a room, or the presence of a ten-word spoken vocabulary. Fifty-four items were written and arranged in developmental sequence, roughly blocked by age year (i.e., 0-12, 13-24, 25-36 months) and consistent with expected major milestones (e.g., first word around 12 months, vocabulary burst around 18 months, two-word sentences around 24 months, etc.). To help parents complete the Snapshot without professional guidance, explanations and examples for potentially ambiguous items were included as needed. For example, anticipating that the question "Does your child follow simple one-step directions" might be difficult to interpret for some parents, we added: "For example, if you say 'go get your shoes' or 'give me the ball' will your child respond correctly?" To cue parents to observe their child's language development, question responses were coded "Yes" for items they consistently observed or "Not Yet" (rather than "No") for items they have not. The total number of "Yes" items constitutes the raw score. This version of the Snapshot was sent to 15 test families along with an age-appropriate MBCDI and the *Ages and Stages Questionnaire* (ASQ) developmental screen (Squires, Bricker, & Potter, 1997). Scores confirmed participant children's typically developing language skills.

Item Refinement Analysis

We examined parent responses to the Snapshot to identify items to clarify, remove or replace. When items endorsed did not match expected developmental sequence, parents were also interviewed regarding their perception of specific questions and examples. Items were removed or replaced when interviews with parents indicated that they could not understand the question or it was not appropriate for a paper and pencil modality. For example, several parents of children in the canonical babbling stage did not endorse the item "Does your child say 'ahh-

goo' or 'goo'?" even though they were producing more "advanced" babbles. Discussion with parents made clear that this item needed to be accompanied by a verbal example by a professional, and so it was removed. In other cases, clarification was added for certain items. For example, the original question "Do the sounds (not including cries) that your child makes vary from high pitch to low pitch?" was changed to "Does your child engage in "vocal play" by producing a wide variety of sounds? For example: does your child produce sounds that range from very high pitch (squeals) to very low pitch (growls)?" Modifications based on this qualitative evaluation yielded a 63-item version of the Snapshot.

Pilot Testing/Scoring

Pilot Testing Participants. The 63-item Snapshot version was sent to families of typically developing children predominantly under 3 years of age who had participated in the *NLS*. Recruitment for this original study involved sending 6,000 recruitment postcards to parents in the Denver-metro area. Interested parents completed a phone interview providing information on household demographics and child language skills. Selected families (n = 329) were from monolingual English-speaking households, reported no language or other developmental delays, and were targeted on the US census on mother's attained education as a proxy for socioeconomic status (SES; see Gilkerson and Richards (2008)). Recruitment for the current study occurred approximately 4 months after their original study had ended, so parents were relatively easy to reach, and the response rate was high. A total of 269 of these families completed the Snapshot and other criterion measures which were sent in the mail. Because a majority of the children whose families had participated in the *NLS* had aged beyond 6 months, an additional sample of 39 families of children ages 2-6 months were recruited using similar procedures to yield a final sample of 308 families (henceforth the Scoring sample) for the current study. Eighty percent

(248) were no older than 36 months, and 92% were 42 months or younger. Data for older children were incorporated into the initial scoring algorithm, but for evaluation analyses, the typically developing sample was restricted to ages 36 months or younger.

Pilot Testing Procedures. Parents completed different criterion measures depending on their child's age. Parents of infants (< 6 months) were mailed the Snapshot and then completed the REEL-3 with a certified SLP during a scheduled phone interview. Parents of older children were sent the Snapshot, MBCDI (8-30 months only) and Child Development Inventory to complete and return in a postage-paid envelope. Participants were compensated $100 for completing the questionnaires and were not provided with assessment results.

Results

Item Analysis. Eleven of the 63 items with poor discriminability (e.g., of chronological age) were eliminated. The resulting 52-item set and its sequencing were selected primarily by applying an Item Response Theory (IRT) framework, specifically Roberts et al.'s Generalized Graded Unfolding Model (GGUM), and examining the item characteristic curve (ICC) of each item, generated using the GGUM2004 software (Roberts et al., 2006). ICCs graphically illustrate the transition over ages from "Not Yet" to "Yes" responses; for example, an ideally discriminating ICC has a relatively steep transition slope and sharp age threshold. However, not all items demonstrated ideal characteristics. Items intended to discriminate at the earliest stages of language development (e.g., "When you talk to your child, does he/she look in the direction of your voice?") were endorsed by most or all parents but were nevertheless retained on theoretical grounds. Items intended to demark later stages did demonstrate age-related discrimination as expected, and item difficulty was consistent with final ordering. To reduce administration time for younger children, we adopted a rule for parents to stop after five 'Not Yet' responses in a

row, since in the majority of such cases no subsequent items were endorsed. Thus, final item sequencing was based on response characteristic metrics informed by clinical expertise. This final version of the Snapshot (including the first ten items) is provided in Appendix A.

Snapshot Scoring. The Snapshot raw score is the total count of "Yes" responses before the five-in-a-row "Not Yet" threshold is reached. To provide a more familiar clinical metric, we transformed this raw count into a Development Age (DA). Median distribution values are commonly used to define development age, but given the limited sample sizes for some age months we chose not to utilize this method. Instead, we calculated the best fit to our data via least squares regression modeling. Ultimately, a nonlinear power function predicting chronological age from total "Yes" count (see Figure 1) provided a good fit to these data (R^2 = 0.90). Snapshot DA correlated well with child chronological age, $r(306) = .92$, $p < .01$, and did not differ significantly from it. To facilitate comparison with age-standardized measures and provide percentile rankings, we generated two additional variant scores. First is the Development Quotient (DQ), calculated in the standard way as 100*(Development Age/Chronological Age). Second, we computed age-standardized Snapshot scores (SS) from the raw score fit at each age month assuming a normal (Gaussian) distribution. From these, percentile ranking values were obtained.

INSERT FIGURE 1 ABOUT HERE

Phase II: Psychometric Evaluation

Participants

Psychometric properties of the Snapshot were assessed first for subsets of the Scoring sample and then compared to additionally recruited samples of children diagnosed with language delays (LD/ASD). Monthly sensitivity to change over time and test-retest reliability were

assessed on families of Scoring sample children 36 months or younger balanced across gender and SES who had completed the Snapshot monthly for up to 26 months as part of the *NLS* (*n* = 59, age *M* = 26.6, *SD* = 4.3 months). These families (henceforth the Monthly sample) had been chosen from the original pool of *NLS* participants to match the U.S. census on mother's attained education and to reflect a broad range of language skills.

To compare the extent to which the Snapshot could discriminate TD children from children with known communication problems, we selected a subset of 225 Scoring sample participants with children up to 36 months of age (henceforth the TD sample). To these families we added two additional samples from the *NLS*, the first 49 children ages 10-44 months diagnosed with language delays unrelated to ASD (sample), and the second 77 children ages 16-48 months with language delays related to ASD (ASD sample). Children who were older than 36 months were included for these delay samples since they could be expected to demonstrate language skills within the Snapshot's development age range of 36 months or younger. Children in these two groups did in fact evidence expressive language delays; the Child Development Inventory expressive language Development Quotient (DQ) means were within the clinical range for language delay, 76.7 (*SD* = 21.2) for the group and 64.9 (*SD* = 19.3) for the ASD group. Extensive diagnostic detail on these delay samples can be found in Oller et al. (2010). Sample demographics are shown in Table 1.

INSERT TABLE 1 ABOUT HERE

Procedures

Procedures for the TD sample were as previously described. The Monthly sample completed the Snapshot at home monthly, the *MBCDI* and *Child Development Inventory*

alternate months, and a certified SLP evaluated children on the PLS-4 and REEL-3 at 4-6 month intervals. The and ASD samples completed a single Snapshot from home.

Results

Snapshot Reliability. Snapshot internal reliability was computed via Cronbach's alpha on the 248 families of children 36 months of age or younger from the Scoring sample and yielded a value in the "Excellent" range ($\alpha = 0.98$). Test-retest reliability was examined in the Monthly sample; specifically, we compared within-family month-to-month changes in raw score total (i.e., Month 1 vs. Month 2, Month 2 vs. Month 3, etc.) over 12 months. Correlations between adjacent monthly scores ranged from $r = .93$-$.98$ (Mean $r = .95$, all $p < .001$). Average within-family monthly change in raw score (i.e., month-to-month change averaged within and across families) was 0.94 items ($SD = 0.56$). That is, the average Snapshot raw score for a typically developing child increased with age at a rate of just under one item per month. The standard error of measurement of average monthly raw score change was $SEM = .125$, corresponding to a 95% confidence interval of [0.70 – 1.19] items. We further examined reliability for these families with respect to monthly change in Development Age (DA). Figure 2 displays average DAs obtained from ages 11-35 months plotted against average DAs scored each subsequent month. For example, when a child obtained DA = 24 months, the average DA received one month later was DA = 24.95 months. Mean within-child monthly DA change was 0.89 months ($SD = 0.47$), which was significantly greater than zero, $t(57) = 14.33$, $p < .001$.

INSERT FIGURE 2 ABOUT HERE

Snapshot Criterion Validity: Standardized Assessments. We examined criterion validity first comparing Snapshot DA to that of SLP-administered and parent-report language-related criterion measures. For the MBCDI we correlated Snapshot raw score to raw Vocabulary score,

controlling for child age. We included only Snapshot assessments completed within 6 weeks of criterion measures for children up to 36 months of age, and thus sample sizes varied. Sample and assessment information and Pearson correlations are provided in Table 2.

INSERT TABLE 2 ABOUT HERE

Snapshot Divergent Validity: Language Delay Identification. We next evaluated criterion validity with respect to delay detection comparing Snapshot DQ scores from 126 children up to 48 months of age with language delays both related to ASD and not (ASD and samples) with those without delays from 225 TD children (to 36 months of age). Delay detection sensitivity and specificity are a function of threshold cutoff value selection; here we evaluated at the value at which sensitivity equaled specificity, i.e., the equal error rate (EER) point. Delay detection performance was very good at EER = 87%, corresponding to DQ = 77. Similarly, the area under the Receiver Operating Characteristic (ROC) curve was high at 0.94; that is, a randomly selected child with a Delay diagnosis would have a lower Snapshot DQ than 94% of TD children. Table 3 provides performance results at the EER point, and Figure 3 displays the sensitivity and specificity relationship for the full ROC curve.

INSERT TABLE 3 ABOUT HERE

INSERT FIGURE 3 ABOUT HERE

Discussion

The purpose of this effort was to develop and initially validate the Developmental Snapshot, a parent friendly, brief, and repeatable measure of children's expressive and receptive communication skills. Details of item development, pilot-testing and refinement, and psychometric properties were reported. Results indicated the Snapshot was reliable internally

and over repeated administrations. The measure was also sensitive to growth over time, strongly correlated with criterion measures, and discriminated significantly between the skills of typically developing children compared to two sub-samples of children with communication challenges. Given these attributes, the Snapshot represents a technological, methodological and conceptual advance in early progress monitoring of children learning to talk and communicate.

Monthly test-retest reliability was well established over a full year of measurement, addressing the first research question. It is important to consider that progression would be expected without the impact of intervention, as typically developing children continually add to their vocabulary and acquire new language skills over time as a natural part of development (Fromkin et al., 2013). It remains to be seen whether more accelerated advancement would be observed for typically developing children who are experiencing positive changes in language input and interaction or for children with language delays participating in intervention.

Our second research question concerned criterion validity of the Snapshot compared with standardized language assessments. Significant correlations observed with other parent report questionnaires were unsurprising given structural and methodological similarities, though the Snapshot is significantly shorter than the Child Development Inventory and MBCDI. More surprising perhaps are strong correlations found with SLP-administered language and cognitive assessments (e.g., PLS-4 and CAT/CLAMS) which include items that require direct observation by a licensed professional. Such correlations indicate overlap between information derivable from the Snapshot and clinician-administered evaluation tools, but the Snapshot requires no clinical expertise to administer or score. Results could be attributed to a number of characteristics of the tool, including pointed questions on specifically observable behaviors followed by clarifying examples meant to reduce subjectivity in responses. However, it should

be noted that the validation participant subsample over-represented higher educated mothers; 79% held a bachelor's degree or had completed some college classes. Future work is needed with more representative samples, including caregivers with lower literacy levels.

Our third research question addressed the extent to which the Snapshot could identify children with known language delays. Although it was not primarily designed to identify language delays, results suggest the potential for its use to flag at-risk children. The empirically-derived equal error rate cutoff of DQ = 77% is consistent with moderate delay. Thus, use of the Snapshot as a monthly measurement tool could also help alert parents and professionals more quickly to the need for a comprehensive evaluation. Note that estimates from these data of the overall performance of the Snapshot when implemented in this manner are inflated by the fact that the proportion of delay diagnosis cases (36%) is considerably larger than would be expected according to more usual estimates of language delay base rates in the general population. Nelson et al. (2006) reported the prevalence in preschool age children of speech and language delays combined to range from 5-8%, though Horwitz et al. (2003) found somewhat higher rates of 8-20% in their sample of more than 1000 children up to age 36 months. Estimates of the precision (positive predictive value) of the Snapshot when used to identify children with language delays can nevertheless be calculated based on the observed sensitivity and specificity for different prevalence rates. For Delay base rates of 5-8%, precision ranges from 25.6% to 36.3%. Precision increases with the base rate; for example, precision is 42.1% at a prevalence of 10%, 53.6% at 15% prevalence, and 62.1% at a high estimate of 20% prevalence. In practice, the EER point of equal sensitivity and specificity may not be the preferred threshold, nor may the general population prevalence be of most interest, and thus precision values would vary accordingly.

Given its demonstrated reliability and validity, it is important to note how the Snapshot may fit practically into a broader assessment portfolio. First, the results presented here suggest it could be used as a level 1 screen for language delays, potentially administered at pediatric check-in visits or preschools to flag children who could benefit from early intervention. Once intervention begins, it can be used both as a progress monitoring tool and/or an effectiveness measure for the intervention itself. As a progress monitoring tool, it can inform professionals about changes in the child that may or may not result from the intervention, allowing them to make necessary adjustments or move the family to a new intervention. As an overall pre- to post-intervention effectiveness measure, Snapshot scores can be used to determine whether an intervention impacted a child's development beyond what would be expected.

The Snapshot represents an advance in language assessment in several ways. First, it can be completed efficiently by a caregiver observing the child in his/her natural environment, rather than relying solely on the judgment of professionals in a clinical environment. Second, it can be completed quickly (10-15 minutes) and automatically scored, minimizing the need for additional resources. Third, the Snapshot is sensitive to language behavior changes over relatively short periods of time; assessment over 12 months showed consistent small increases in monthly raw scores, noteworthy given the lesser resolution of other professionally administered assessments.

Results suggest the Snapshot can be used as a monthly feedback tool for parents participating in interventions designed to enhance the early language environment. More specifically, because scoring supports frequent administration, parents can obtain rapid feedback on changes in their child's development which could reflect parental efforts to implement intervention strategies. Thus, in this context it can provide encouragement to support parents' own behavioral changes as they notice their child acquiring new language skills on a monthly

basis. This approach represents a distinct shift from the current standard in language testing, which typically requires 6 months between test administrations and precludes frequent progress monitoring. The ability to sample skills monthly allows professionals to make informed, data-driven decisions about an intervention. Further, it underscores language milestone behavior, so parents become better informed observers of their child's progress. Perhaps most significantly, the Snapshot represents a conceptual framework that focuses the parental mindset on the importance of continual observation in the natural home environment while they are with their child, highlighting behaviors to watch for and expect in the early stages of development. In this way, it encourages a shift toward empowering parents with knowledge of milestones, nurturing the expectation that they can be first to notice developmental progress occurring at home and during daily routines, as opposed to relying solely on the input of professionals who see the child intermittently and often in unfamiliar environments.

Given that language assessments are traditionally completed at 6 or 12 month intervals, the feasibility of monthly administrations may seem dubious from a clinical perspective. However, the Snapshot is particularly well-suited as an add-on to interventions already implementing regular visits or routine check-in meetings with parents, since it was designed to be completed fairly quickly (10-15 minutes). As well, since it can be completed by the parent without professional supervision, it can be sent in the mail when in-person meetings are impractical. The feasibility of monthly administration has been demonstrated in a recent yearlong LENA Research Foundation pilot intervention including bi-weekly and monthly parent group meetings, with 74% of participants completing it at 4-week intervals. Note too that successful utilization of frequent monitoring tools in early intervention has also been demonstrated with the Early Communication Indicator (ECI), a brief, professionally

administered observational tool completed at monthly and bi-weekly intervals to gauge response to intervention in terms of child communication skills (Greenwood et al., 2010).

Limitations

Despite its psychometric properties and potential benefits for early progress monitoring, there are several limitations to use of the Snapshot to note. First, it is focused almost exclusively on receptive and expressive language development and does not specifically address other aspects of development such as gross/fine motor, social skills or pragmatic language. Although a more comprehensive assessment would include these items, at the early ages there is generally a strong correlation between language acquisition and other areas of development, so if other issues are present they could be reflected in aspects of language (e.g., Iverson, 2010; Schuster, 2000; Dale et al., 1989; Horwitz et al., 2003). Relatedly, the Snapshot offers no developmental subscales and is limited in the number of items developmentally appropriate between 2-36 months. Also, although reliability and validity are strong, the Snapshot was age-referenced on a relatively small, local sample and to date has only been validated in English. The parent report modality of the Snapshot may also be considered a limitation of sorts. Although relying solely on parent report to assess early language skills and other behavioral domains has been called into question due to the potential for response bias (Sameroff et al., 1982; Seifer et al., 1994) a number of parent questionnaires have demonstrated strong reliability compared to professional evaluation (Doig et al., 1999; Dale, 1991; Rescorla and Alley, 2001). Correlations observed between professional judgment and parent report of language skills in the current study suggest minimal parental bias, at least in this sample. Two further caveats should be noted. In addition to being based on relatively sparse data at some ages, interpretation of age-standardized values may be subject to a ceiling effect. Although Snapshot items were designed to cover children up

to 36 months of age reflecting a wide range of development levels, younger but more advanced children conceivably could achieve the maximum raw score. Consequently, this cap may not indicate these children's true upper range of language development. Consideration is warranted also when the Snapshot is used with delayed children older than the reference set. Although these children may be within the scope of the Snapshot developmentally, the appropriate age reference group may not be clear.

Future Research and Practice

Ongoing and planned research projects intend to expand use of the Snapshot to other languages. Thus far, the Snapshot has been translated into Korean, Mandarin and Spanish. In each case there have been items that could not be directly translated which were replaced with developmentally appropriate, language-specific items. The Asian-language versions are currently being tested in Seoul and Shanghai and will soon be standardized to larger populations in these languages. We further expect to expand the current English-speaking sample to include more representative geographical, racial and cultural cross-sections of children.

Although the Snapshot demonstrated respectable discrimination in identifying children with language delays, the samples were small, and more research is warranted. Indeed, the ease of use and frequency of administration this tool enables could contribute substantially to research on early screening. A review of 24 studies focusing on screening in preschool children concluded that more research is needed to determine the optimal age, administration interval and modality for early screening (Nelson et al., 2006). Use of the Snapshot as a monthly progress monitoring tool in preschools or during regular pediatric visits could help provide more precise information on when language delays are likely to surface for certain subsets of children, as well as the ages at which identification and treatment of delays are most effective.

Additional next steps in research with the Snapshot include examination of its sensitivity to parent-implemented intervention effects as intended. Given reported findings, practitioners working with parents to bridge the word gap in disadvantaged families would likely benefit from inclusion of the Snapshot into their intervention practices. Contributions of the Snapshot toward the study of parent implemented intervention effects and sustainability await additional research.

Conclusions

By providing a means to quickly and efficiently evaluate infant/toddler language development on a monthly basis, the Snapshot supports the developing trend in collaborating with parents in the implementation of parent-focused interventions. The Snapshot can be used as a valid and reliable measurement tool to focus parental attention on dynamic behavioral changes in linguistic development. This approach is essential to intervention and progress monitoring because, unlike intermittent professional evaluations that are static in time, parents are continually present and can be sensitive to more fluid behavioral advancements. By encouraging parental observation of child behavior over time, and reliably reporting behavioral changes at monthly intervals, the Snapshot can facilitate professional understanding of the topography of child language behavior in a way that has not been previously accessible.

Acknowledgements

We deeply appreciate the time and effort made by all our participant families. We gratefully acknowledge speech language pathologists Jennifer Cruz and Michelle Bargo who contributed valuable input during the development phase, and Rebecca Mills, Tammi Trojan and April Teehan who worked tirelessly to collect and manage participant data. We thank Charles Baer and Steven Johnson for programming the computer software, Dongxin Xu for helpful input

during development, and Tricia Striano for editorial input. Finally, we are grateful to Terrance Paul and the LENA Research Foundation for development, project feedback and support.

Funding

This projected was funded by the LENA Research Foundation.

Conflict of Interest Statement

The work described here is part of the development of the Developmental Snapshot. This intellectual property may be licensed or sold by the non-profit 501(C)3 LENA Research Foundation, but none of the authors will receive financial gain for products related to the research described in this paper. Dr. Greenwood and Dr. Montgomery are unpaid volunteer members of the LENA Research Foundation scientific advisory board, and Dr. Gilkerson and Mr. Richards are staff employees of the LENA Research Foundation.

References

Accardo PJ and Capute AJ. (2005) *The Capute Scales: Cognitive Adaptive Test/Clinical Linguistic and Auditory Milestone Scale (CAT/CLAMS),* Baltimore, MD: Paul H. Brookes Publishing Co.

Bzoch KR, League R and Brown VL. (2003) *Receptive-expressive emergent language test,* Austin, TX: PRO-ED.

Crow S and O'Leary A. (2015) Word health: Addressing the word gap as a public health crisis.

Dale PS. (1991) The validity of a parent report measure of vocabulary and syntax at 24 months. *Journal of Speech and Hearing Research* 34: 565-571.

Dale PS, Bates E, Reznick JS, et al. (1989) The validity of a parent report instrument of child language at twenty months. *Journal of Child Language* 16: 239-249.

Doig KB, Macias MM, Saylor CF, et al. (1999) The child development inventory: A developmental outcome measure for follow-up of the high risk infant. *The Journal of Pediatrics* 135: 358-362.

Fenson L, Marchman VA, Thal DJ, et al. (2007) *MacArthur-Bates communicative development inventories,* Baltimore, MD: Paul H. Brookes Publishing Co.

Fromkin V, Rodman R and Hyams N. (2013) An introduction to language (10th ed.). Boston, MA: Cengage Learning.

Gilkerson J and Richards JA. (2008) The LENA natural language study. *Technical Report No. LTR-02-2.* Boulder, CO: LENA Research Foundation.

Greenwood CR, Carta JJ, Walker D, et al. (2006) Preliminary investigations of the application of the Early Communication Indicator (ECI) for infants and toddlers. *Journal of Early Intervention* 28: 178-196.

Greenwood CR, Walker D and Buzhardt J. (2010) The early communication indicator for infants and toddlers: Early Head Start growth norms from two states. *Journal of Early Intervention* 32: 310-334.

Halfon N, Olson L, Inkelas M, et al. (2002) Summary statistics from the National Survey of Early Childhood Health, 2000. National Center for Health Statistics. *Vital Health Stat* 15: 1-34.

Hart B and Risley TR. (1992) American parenting of language-learning children: Persisting differences in family-child interactions observed in natural home environments. *Developmental Psychology* 28: 1096-1105.

Hart B and Risley TR. (1995) *Meaningful differences in the everyday experience of young American children,* Baltimore, MD: Paul H. Brookes Publishing Co.

Hart B and Risley TR. (1999) *The social world of children learning to talk,* Baltimore, MD: Paul H. Brookes Publishing Co.

Hoff E. (2003) The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development* 74: 1368-1378.

Horwitz SM, Irwin JR, Briggs-Gowan MJ, et al. (2003) Language delay in a community cohort of young children. *Journal of the American Academy of Child & Adolescent Psychiatry* 42: 932-940.

Huttenlocher J, Haight W, Bryk A, et al. (1991) Early vocabulary growth: Relation to language input and gender. *Developmental Psychology* 27: 236-248.

Ireton H. (1992) *Child development inventory,* Minneapolis, MN: Behavior Science Systems, Inc.

Iverson JM. (2010) Developing language in a developing body: The relationship between motor development and language development. *Journal of Child Language* 37: 229-261.

Kaiser AP, Hemmeter ML, Ostrosky MM, et al. (1996) The effects of teaching parents to use responsive interaction strategies. *Topics in Early Childhood Special Education* 16: 375.

Landry SH, Smith KE and Swank PR. (2006) Responsive parenting: Establishing early foundations for social, communication, and independent problem-solving skills. *Developmental Psychology* 42: 627-642.

Landry SH, Smith KE, Swank PR, et al. (2008) A responsive parenting intervention: The optimal timing across early childhood for impacting maternal behaviors and child outcomes. *Developmental Psychology* 44: 1335-1353.

Nelson HD, Nygren P, Walker M, et al. (2006) Screening for speech and language delay in preschool children: Systematic evidence review for the US Preventive Services Task Force. *Pediatrics* 117: e298-e319.

Oller DK, Niyogi P, Gray S, et al. (2010) Automated vocal analysis of naturalistic recordings from children with autism, language delay, and typical development. *Proceedings of the National Academy of Sciences* 107: 13354-13359.

Providence Talks. (2015) Providence Talks: Pilot findings & next steps.

Rescorla L and Alley A. (2001) Validation of the Language Development Survey (LDS): A parent report tool for identifying language delay in toddlers. *Journal of Speech, Language, and Hearing Research* 44: 434.

Roberts JS, Fang H, Cui W, et al. (2006) GGUM2004: A Windows-based program to estimate parameters of the generalized graded unfolding model. *Applied Psychological Measurement* 30: 64-65.

Roberts MY and Kaiser AP. (2011) The effectiveness of parent-implemented language interventions: A meta-analysis. *American Journal of Speech-Language Pathology* 20: 180-199.

Rowe ML. (2008) Child-directed speech: Relation to socioeconomic status, knowledge of child development and child vocabulary skill. *Journal of Child Language* 35: 185-205.

Sameroff AJ, Seifer R and Elias PK. (1982) Sociocultural variability in infant temperament ratings. *Child Development* 53: 164-173.

Schuster MA. (2000) Developmental screening. In: McGlynn EA (ed) *Quality of Care for Children and Adolescents: A Review of Selected Clinical Conditions and Quality Indicators.* Santa Monica, CA: RAND, 157-168.

Seifer R, Sameroff AJ, Barrett LC, et al. (1994) Infant temperament measured by multiple observations and mother report. *Child Development* 65: 1478-1490.

Shankar M. (2014) Empowering our children by bridging the word gap. *The White House blog*.

Suskind DL, Leffel KR, Graf E, et al. (2015) A parent-directed language intervention for children of low socioeconomic status: A randomized controlled pilot study. *Journal of Child Language*.

The Family Conservancy. (2015) *Talk, read, play (with your child every day)*. Available at: https://thefamilyconservancy.org/talkreadplay.

Warren SF and Brady NC. (2007) The role of maternal responsivity in the development of children with intellectual disabilities. *Mental Retardation and Developmental Disabilities Research Reviews* 13: 330-338.

Zimmerman IL, Steiner VG and Pond RE. (2002) *Preschool language scale,* San Antonio, TX: The Psychological Corporation.

Table 1. Participant Demographics across Samples

	Scoring Sample		Monthly Sample		TD Sample		LD Sample		ASD Sample	
	N	%	N	%	N	%	N	%	N	%
Total	308	100%	59	100%	225	100%	49	100%	77	100%
Female	152	49%	33	56%	117	52%	16	33%	13	17%
Age Group										
2-12	82	27%	8	14%	61	27%	3	6%	0	0%
13-24	83	27%	36	61%	64	28%	11	22%	7	9%
25-36	83	27%	15	25%	100	45%	29	59%	34	44%
37-48	56	18%	0	0%	0	0%	6	12%	36	47%
49+	4	1%	0	0%	0	0%	0	0%	0	0%
Mother's Educ										
Some HS	35	11%	7	12%	33	15%	3	6%	2	3%
HS or Equiv	100	32%	14	24%	63	28%	9	18%	14	18%
Some College	86	28%	23	39%	69	31%	21	43%	13	17%
BA+	87	28%	15	25%	60	26%	16	33%	48	62%

Note. Mother's attained education groups are: Some High School (HS; no degree); HS or Equivalent (includes GED and Trade School diploma); Some College (college courses but no bachelor's level degree); and Bachelor's Degree or higher (BA+).

Table 2. Relationship of Snapshot to Criterion Measures for Children up to 36 Months

Criterion Assessment	Completed By	Age Range (Months)	N	Pearson Correlation
PLS-4 Receptive	SLP	5-36	62	.93
PLS-4 Expressive	SLP	5-36	62	.93
REEL-3 Receptive	SLP	5-35	63	.94
REEL-3 Expressive	SLP	5-35	63	.94
Child Development Inventory Receptive	Parent	16-36	143	.84
Child Development Inventory Expressive	Parent	16-36	142	.81
CLAMS	SLP	5-35	61	.97
CAT	SLP	5-35	61	.94
MBCDI Vocabulary	Parent	8-31	155	.67

Note. Pearson correlations here reflect associations on development age between Snapshot and criterion assessments completed within six weeks of Snapshot administration, except MBCDI Vocabulary which associates Snapshot raw score with Vocabulary raw score, controlling for child age. All correlations $p < .001$.

Table 3. Snapshot Performance for Detecting Delay at the Equal Error Rate Point

	Delay Diagnosis	Typically Developing	Total
DQ < 77	110	30	140
DQ >= 77	16	195	211
Total	126	225	351

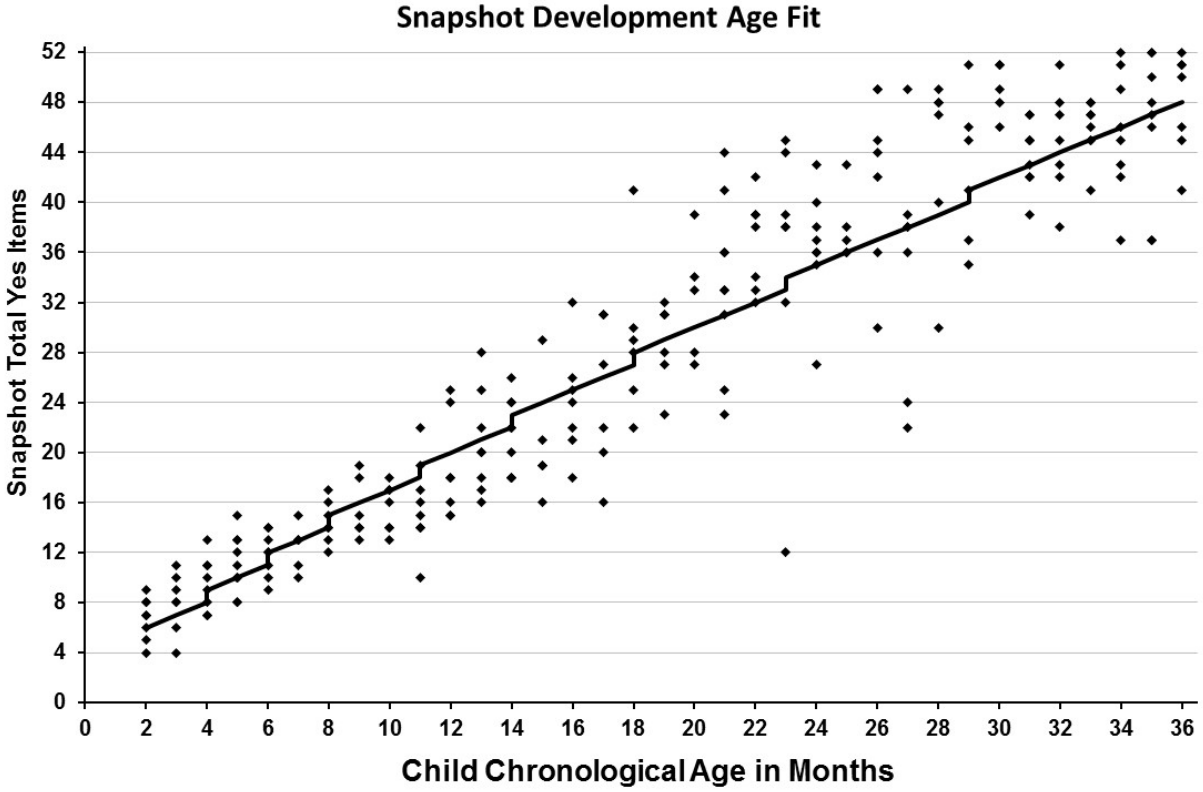
Note. Development Quotient (DQ) is computed as 100*Development Age/ Chronological Age.

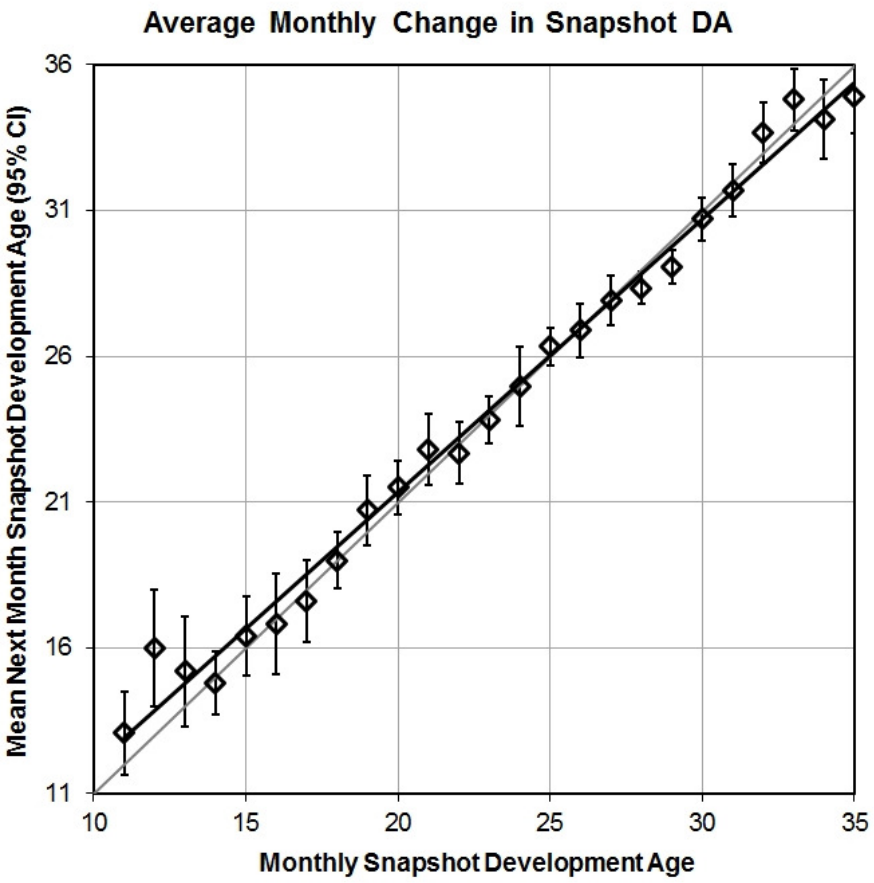
Figure Captions

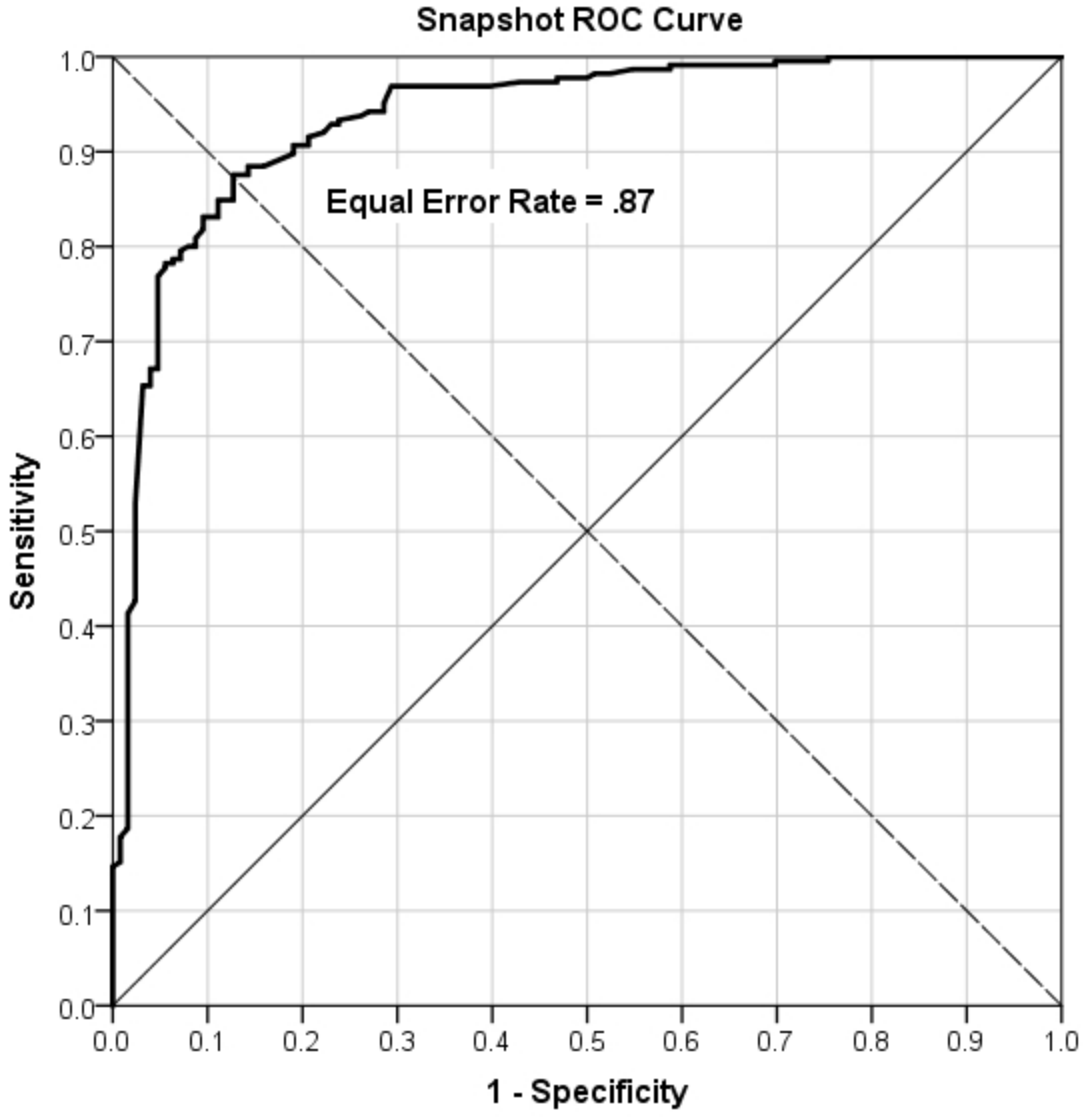
Figure 1. Snapshot total scores across child chronological age with nonlinear fit line.

Figure 2. Average monthly change in Snapshot Development Age.

Figure 3. Snapshot ROC curve for identification of language delay samples.


Snapshot Development Age Fit
Snapshot Total Yes Items
0
4
8
12
16
20
24
28
32
36
40
44
48
52
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
Child Chronological Age in Months


Average Monthly Change in Snapshot DA
Mean Next Month Snapshot Development Age (95% CI)
36
31
26
21
16
11
10
15
20
25
30
35
Monthly Snapshot Development Age


Snapshot ROC Curve
Equal Error Rate = .87
Sensitivity
1 - Specificity
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0