

Research and Development of an Oral Language Measure for 3-Year-Old Children
Expanding Individual Growth & Development Indicators of Language and Early Literacy for
Universal Screening in Multi-Tiered Systems of Support with 3-Year-Olds

Technical Report #9

Kelsey Will, Scott McConnell, Anthony Albano, Marianne Elmquist, Erin Lease, and Alisha

Wackerle-Hollman

IGDILab

Department of Educational Psychology

University of Minnesota

Updated: 10/19/22

IGDILAB



Preliminary Research

Contact Authors for Updates

Authors' Note: This work was supported by Grant Number R305A160034, *Expanding Individual Growth & Development Indicators of Language and Early Literacy for Universal Screening in Multi-Tiered Systems of Support with 3-Year-Olds* from the National Center for Educational Research, Institute of Education Sciences to the University of Minnesota, Scott McConnell, Principal Investigator.

This Technical Report presents preliminary findings or intermediary results of our work. Please contact the authors for a more up-to-date version or for permission before citing or distributing. For more information, email igdilab@umn.edu.

Scott McConnell, Alisha Wackerle-Hollman, and colleagues developed Individual Growth and Development Indicators, or *IGDIs*. Intellectual property from this research has been licensed to Early Learning Labs, Inc., and subsequently to Renaissance Learning for commercial development and sale. Scott, Alisha, and the University of Minnesota have royalty and/or equity interest in Renaissance Learning. These relationships have been reviewed and managed by the University of Minnesota in accordance with its conflict of interest policies.

The authors acknowledge the teachers and children who contributed their time and effort to this project, and the early childhood program leaders who made this work possible. We'd also like to acknowledge earlier project contributors Kristin Schuster, Alyssa Schardt, and Michael Rodriguez. We extend a special thanks to our group of data collectors who drove throughout the metro in the Minnesota sunshine and snow: Cassie Neu, Drake Bauer, Grace Vieth, Rachel Garofalo, Alex Willets, Meredith Reese, Lora Kussman, and Laurie Orstad.

Preliminary Results: Contact Authors for Update

Research and Development of an Oral Language Measure for 3-Year-Old Children

Technical Report #9

Abstract

This report is a detailed technical report of research completed as part of an ongoing program of research on assessment of language and early literacy development among three-year-old children. This study sought to assess whether less-difficult receptive items relate to more difficult expressive items in ways that allow for combination of the item types to assess oral language in 3- and 4-year-old children. In particular, we wanted to determine if new receptive items demonstrate acceptable psychometric features, and whether these new items are associated with existing, and more difficult, expressive items in ways that produce a single scale dimension. We evaluated whether items in the expanded scale sample a broader range of oral language development, with some degree of overlap in abilities sampled by receptive and expressive items, such that procedures like computer-adaptive testing would have an appropriate item bank for assessing 3- and 4-year-old children. Four hundred and forty-nine children participated in one or more seasonal assessments while enrolled in 3-year-old classrooms. A combination of linear and computer-adaptive tests were completed to a) assess item characteristics and locate both receptive- and expressive-response item types on a common Rasch scale, b) evaluate the extent to which this scale could be considered unidimensional for purposes of item banking and score reporting, c) examine relation between person ability and item bank coverage across seasons, and d) descriptively evaluate changes in mean group performance across three successive seasons. Taken together, results presented here indicate that an expanded approach to assessing growth and development in oral language – an approach that adequately samples and describes the performance of lower-performing younger 3-year-olds as

well as higher-performing older 4-year-olds – can be achieved. Additional research will be needed to describe the content, construct, discriminant, and instructional utility of this new measure; initial stages of this research have been completed and will be reported in separate IGDILab technical reports.

Preliminary Results: Contact Authors for Update

Research and Development of an Oral Language Measure for 3-Year-Old Children

Technical Report #9

Project Introduction

This document has been developed as part of a five-year (2016 – 2021) research and development project funded by the Institute of Education Sciences, *Expanding Individual Growth & Development Indicators of Language and Early Literacy for Universal Screening in Multi-Tiered Systems of Support with 3-Year-Olds*. The overarching purpose of this project is to extend the practical array of Individual Growth and Development Indicators (IGDIs) for assessing language and early literacy development to developmental ability and skills typical of 3-year-old preschool children in ways that inform and enhance multi-tiered systems of support (MTSS), and to enrich current knowledge of the developmental course of language and early literacy development in this early preschool age group. Over the course of this project, we developed and evaluated measures appropriate for 3-year-old children¹ in areas of oral language, phonological awareness, and alphabet knowledge. During this process, we evaluated and refined definitions of domains of assessment; developed and pilot tested different formats for gathering child responses; wrote and field-tested items for each domain; refined and organized these items into scales; assessed the psychometric characteristics of these scales (including their classification accuracy for identifying candidates for more intensive intervention); and used them to describe growth across time in 3-year olds and 4-year-olds in ways that help describe the broader domain of language and early literacy development.

¹ While we refer to 3-year-old children throughout this report, IGDIs are *not* age-normed nor intended to support inferences of age-based development. Rather, the focus here is development and evaluation of measures and data utilization tools that support assessment of language and early literacy skills that precede, developmentally, those measured by current IGDIs – and that, in broad terms, are likely to be developed by children more than one and less than two years prior to kindergarten entry – a “grade” we reference as PK3.

Purpose of this Report

This report presents Year 3 research and development of the Age 3 IGDIs Oral Language measure (McConnell et al., 2019), which focused on dimensionality and scaling of Point to Picture (P2P) and Picture Naming (PN) item types, and location of all items on the existing Age 4 Picture Naming scale. This work builds on domain specification and a related literature review presented in Schuster, Cai, et al. (2017), *Defining a Domain of Assessment for Language and Early Literacy in 3-Year-Olds: Oral Language (Technical Report #3)*. Technical report 3 defines oral language as “the ability to use words to communicate ideas and thoughts and to use language as a tool to communicate to others....Within the broader category of oral language, expressive language is the use of words to express meaning, and receptive language is the ability to listen, process, and understand the meaning of spoken words” (Bradfield et al., 2014, p. 234).

Oral Language Development

Oral language development, or a child’s growing receptive and expressive vocabulary and ability to communicate with and understand others, is a hallmark of the preschool years. Children’s acquisition of basic language skills is not only an important developmental task (e.g., Kaiser & Roberts, 2011; Kuhl, 2004), but also is an important, if not necessary, condition for developmental growth in other domains (Casenhiser et al., 2013; Gonzalez et al., 2010; Hoff, 2013; Pellegrini et al., 1998; Pruden et al., 2005; Scarborough, 2001; Walker et al., 1994).

Due to its central role in young children’s current and future development, detection and early intervention for delays in language development are of paramount importance. Assessment and intervention models vary widely from home to community to classroom settings (Noam & Hermann, 2002), with growing attention to classroom-based interventions as the prevalence of preschool and early education services for young children has continued to expand (Haring Biel

et al., 2019; Wasik & Hindman, 2011; Zucker et al., 2013). This expansion in both the prevalence of early education programs and initiatives to promote oral language development has accompanied growth in research, development, and dissemination of multi-tiered systems of support for preschool-aged children (Carta, 2019; McConnell, 2019).

Multi-tiered systems of support (MTSS) require integrated assessment and intervention resources (Carta, 2019). In general, MTSS includes a) access to high-quality and evidence-based interventions that support the development of most children; b) universal screening to identify children for whom that general level of intervention is not sufficient for them to meet current developmental achievement expectations; c) supplemental and more intensive evidence-based interventions that can be provided to these children; and d) systematic monitoring of progress for children receiving supplemental intervention to evaluate the effects of this additional support.

To help meet this greater demand for MTSS in early education programs, researchers have developed and are continuing to expand the number and quality of assessment tools designed specifically for screening and progress monitoring (Carta, 2019). Developers of one such set of tools, Individual Growth and Development Indicators (IGDIs), initially focused on measures appropriate for use primarily in the year prior to kindergarten entry, PK4, for children who speak English and/or Spanish (Bradfield et al., 2014; Durán et al., 2019; McConnell et al., 2015). Research and implementation to date has demonstrated both the psychometric rigor and practical utility of these measures (Kincaid et al., 2020).

While IGDIs research and development to date appears to offer resources for some MTSS implementations, design and evaluation has not kept pace with expansion of broad access to early education, and the commitment to preventive and early intervention like MTSS. As public and private preschool programs with specific curricular objectives expand services to

younger, lower-ability children, measures must be adapted to accommodate the characteristics of this new population. In particular, two priorities have emerged: First, the form and difficulty of measures must be examined and, if necessary, adapted to better match the skills and abilities of younger children. Second, to preserve measurement and intervention continuity that is central to both principles of General Outcome Measurement (Fuchs & Deno, 1991) and MTSS in early education (McConnell, 2019), any new measures should be aligned conceptually and statistically with those used as children continue to develop and move toward kindergarten enrollment.

The purpose of research reported here is to address these two priorities for screening oral language development among younger preschool students. Given both practical experience and prior research suggesting that receptive language tasks are easier than expressive ones like our existing Picture Naming measure for individuals either close to their third birthday or with relatively less-developed oral language skills, we thought it worthwhile to explore receptive item types as we expanded our ability to estimate earlier developmental status in this domain. We also wanted to preserve the noun-focused naming function of current administration, both due to practical experience that Picture Naming was engaging and efficient with young children and to increase efficiency for vertical scaling of child performance across up to two years prior to kindergarten entry.

Early-stage work to expand this measure followed previous research and development efforts (see McConnell et al., 2015), including review and confirmation of the target domain (Schuster, Cai, et al., 2017), brainstorming a variety of formats for sampling this domain with early educators, and testing prototypes of most-promising formats with small samples of young children (Schuster, Schardt, et al., 2017). A receptive task, Point to Picture, emerged from these efforts as a promising format for designing easier items for younger, lower-ability children.

During Point to Picture, a child is presented a horizontal array of three nonambiguous images and asked to point to the one image the examiner prompts. A review of developmental literature suggested that this receptive response format would produce less-difficult items (Schuster, Cai, et al., 2017). The task was easy to administer, requiring little or no instruction and training for children before beginning assessment. The item design process leveraged existing stimuli and technology assets used in the existing Picture Naming task. Finally, logical analysis suggested that Point to Picture items had potential to be combined with expressive Picture Naming items to form a single scale with application across a broader range of ages and the oral language abilities expected for 3- and 4-year-old children.

Given an overall intention to form a single oral language measure appropriate for use across at least two years of preschool development and an expectation to link items within the measure to an existing pool of items originally designed for older children, the work here focused on using a computer-adaptive testing model for item selection and scoring. Computer adaptive testing will allow children across the two years prior to kindergarten enrollment to access developmentally appropriate IGDI items. However, to psychometrically evaluate this model, we required evidence of common construct-relevant variance in the sets of both item types, and calibration scales placing the new Point to Picture items on the existing Rasch Picture Naming scale.

As such, this project aimed to answer two research questions:

1. To what degree can Point to Picture and Picture Naming items be combined into a single scale to assess developmental ability across a wider range of ages?
 - a. To what extent can two item types, Point to Picture and Picture Naming, be represented in a single measurement factor?

2. What are the results of scaling all items of both types on a single existing scale, and to what extent do item locations and child ability scores align?

Finally, in preparation for subsequent and more formal analyses, we evaluated descriptively distributions of item selection and student performance within and across three seasonal assessments.

Method

Participants

Children were recruited from three large urban and “first-ring suburban” districts and a community-based Head Start program in the Midwest, and one charter Head Start program in an urban area on the East Coast. To be included in this sample, children had to be enrolled in a participating early childhood program, be between the ages of 36-59 months at the time of enrollment, have basic English language proficiency as determined by parent report of home language or teacher judgement, and parental consent to participate. Children were excluded from the study if they were not within the appropriate age range, did not demonstrate basic English language proficiency, or if their parent/guardian returned the passive consent form indicating they did not want their child to participate.

A total of 449 children completed at least one oral language assessment during any of three seasonal assessment periods. Children were 36 to 52 months of age ($M = 43.59$, $SD = 3.61$) at the Fall assessment. The sample was 44.1% boys and 49% girls, with 6.9% of the sample missing reported gender information. Approximately 12% of the sample was reported as receiving special education services with an IEP (6% missing data) and approximately 16% of the sample was reported as eligible for free/reduced lunch (38% missing data). Table 1 presents a

summary of these participant characteristics as well as their reported races/ethnicities and primary home language.

Measures

Typically, all IGDI measure are brief (1-5 minutes) and include a three or four item sample structure that aids in determining if children have a baseline skill set necessary to engage with the task. For the purposes of this study, the sample mechanisms were not used to evaluate how children with very low oral language skills engage with the new easier items.

Point to Picture

Point to Picture (P2P) is a receptive oral language task. Assessment is delivered via two iPads, linked by Bluetooth. The examiner's device controls all activities, delivering administration and scoring directions to the examiner as well as navigation resources to advance the assessment on both devices. The child's device is used to display images and other stimuli for each item of assessment, and to record tactile responses from the child for receptive tasks. Each item presents 2-3 images to the child. The examiner verbally prompts the child to "point to the (target item)." The child touches a response on the iPad and it is automatically scored as "correct" or "incorrect," and the examiner confirms the child's selection on the examiner's device. The item is scored as "no response" if the child does not select a response or says, "I don't know."

Prior to this investigation, a total of 66 P2P items were written and field-tested and provisionally calibrated with 3-year-old children (McConnell et al., 2018). All of these items were used in the Fall administration except two items that had been used for demonstration and had not been calibrated. After the Fall administration and item calibration, 12 items were

removed due to poor infit or outfit. Detailed results are provided in Technical Report #6 (McConnell et al., 2018).

Picture Naming

Picture Naming (PN) is an expressive oral language task. Each item presents one image to the child. The examiner asks, “What is this?” or “What’s that?” The examiner scores a child’s verbal response as “correct” for an exact match or appropriate extension of the correct response. A child’s verbal response is scored by the examiner as “incorrect” if the response is not an exact match to the correct response or is an inappropriate extension or generalization of the correct response. The item is scored as “no response” if the child does not respond to the item or says, “I don’t know.” Screenshots of a P2P item and a PN item are in Appendix A.

Procedures

Form Creation

To evaluate possible unidimensionality of P2P and PN items, Fall assessment was conducted using six linear forms. Each form included both P2P and PN items, with common items of each type repeated across all six forms. These forms were constructed to include all 64 existing P2P items and 20 of the lowest-difficulty PN items. Each form had 27-28 items (see Appendix B).

Winter and Spring assessments were completed using an existing computer-adaptive testing (CAT) algorithm (Wang et al., 2017) and calibration of P2P items based on concurrent assessment of P2P and PN items during Fall assessment. This adaptive process, identical to that used for other IGDIs, begins seasonal assessment with an item randomly selected within .25 logits of the presumed seasonal benchmark. Child ability was estimated after each response, and successive items were selected randomly from items with difficulty $\pm .1$ logit from this estimated

ability. Children saw 25 items in the Winter and 25 items in the Spring. After completing 25 items, the child's final ability estimate was converted to a reporting score scale, and reported with a confidence interval of \pm the standard error of measurement (Wang, et al., 2017).

Training and Fidelity

Graduate research assistants trained data collectors on administration of all Age 3 IGDI measures (Oral Language, Alphabet Knowledge, and Robot Blending) during a single hourlong training session. Data collectors were able to practice administering the measures before being asked to demonstrate fidelity on each of the measures. Data collectors only moved on to live administration after achieving at least 90% fidelity on each measure.

Administration and Scoring

We completed all measures with child participants in their early childhood program classrooms or in the hallways outside their classrooms. Lead graduate research assistants and data collectors administered all measures.

During Fall assessment, child participants each completed two linear oral language forms. Participants were randomly assigned to complete either forms 1 and 4, forms 2 and 5, or forms 3 and 6. Each form had 27-28 items and took less than five minutes to administer. Participants completed the first oral language form, followed by the alphabet knowledge IGDI, the second oral language form, and the phonological awareness IGDI. Participants did not need to complete all four measures in a single sitting if they needed a break. During Winter and Spring assessments, children completed a single CAT form for each of three Age 3 IGDI (oral language, alphabet knowledge, phonological awareness).

Results

Across three seasons of assessments, a combination of linear and CAT forms were administered. In Fall, six linear forms, each with 27 or 28 items, were administered; each child participant completed two forms. An average of 143 children (range 101 – 222) completed each form, generating 23,544 item-level responses ($M = 3,924$ per form). In Winter and Spring, the full item banks of 52 P2P items and 248 PN items were available for computer-adaptive administration. Forms were completed by 406 children in winter and 411 children in spring, producing 20,425 item-level responses on 296 unique items.

Scale Characteristics

Child responses from Fall administration of six linear forms were evaluated using traditional classical test theory item- and scale-level metrics. Appendix C presents results by item by form. These detailed results are summarized in Table 2, including means, standard deviations, minimum and maximum estimates across all included items. Results also provide mean scored response (m , also known as p-value or item difficulty), standard deviation (SD), number of responses (n), number of children with missing responses (Missing), the item-total correlation (ITC, a measure of item discrimination), the corrected item total correlation (CITC), and the alpha internal consistency for the full scale if the item were deleted (AID).

Item difficulties ranged from .28 to .99, and item-total correlations .047 to .615. While some individual items in single forms failed to meet one or more *a priori* criteria, in general items performed as required for inclusion in item pools for subsequent CAT testing. As a result, all items were retained for further analysis.

Confirmatory Factor Analysis

To evaluate whether receptive P2P and expressive PN items could reasonably be combined in a single scale, we evaluated unidimensionality across these two item types with three confirmatory factor analysis models for each linear form. The first model included all receptive P2P and expressive PN items, the second only the P2P items, and the third only the PN items. Models were all fit in Mplus with a single factor and categorical outcomes. Models for each linear form were compared using the root mean square error of approximation (RMSEA) and comparative fit index (CFI). For this analysis, primary attention is directed to the all-items model; testing the fit of this model allows evaluation of the degree to which the two types of items constitute a common scale for measurement purposes. For completeness, we also include brief reports of results for the two single-item-type model fits.

Table 3 presents CFA results for all items for each form. RMSEA results met a criterion value of less than .05 for two of six forms, with three other results marginally higher. Similarly, CFI results exceeded a criterion of .90 for two of six forms, with two additional forms marginally below criterion. Given some disagreement about absolute standards for RMSEA and CFI results related to tests of unidimensionality (Kline, 2010), the number of results that approached but did not meet these criteria, and considerations in establishing both validity and utility of oral language measures, we elected to retain the mixed-item-type format for oral language assessment. Note that in form ID 281118005, one item (item ID 210056) was removed from the model because of a linear dependency which negatively impacted model fit.

Calibration

Point to Picture items were initially calibrated as part of our initial item testing research, and initial item fit and location statistics are presented in a report of earlier research (see

Technical report 6.; McConnell et al., 2018). To create a single scale for describing oral language performance across the ability scale now sampled by receptive and expressive items, Rasch modeling was used to calibrate receptive P2P items onto the existing expressive PN scale, where the PN items served as anchors with known parameter values based on prior investigations (Bradfield et al., 2014). This was achieved by fitting a Rasch model to the full OL Fall data set across all six forms, obtaining item difficulties for all items, P2P and PN, and then linking parameters to the existing PN scale using a mean/sigma transformation (Kolen & Brennan, 2014).

Twenty PN items included in the linear forms served as anchors to the existing PN scale. Prior to calibrating new items, we compared Rasch item difficulties for these 20 anchor items when completed by 4-year-olds in prior research and the current sample of 3-year-old children; the intent of this analysis was to evaluate stability of relative item difficulty, assuming a change in absolute difficulties when items were completed by younger students.

Relations in Rasch item difficulty were roughly linear and without any significant outliers (see Figure 1). The expected shift in observed difficulty was noted, with items completed by 3-year-olds tending to be more difficult (theta ranging from -2 to +3) when compared to difficulties on these items when completed by 4-year-olds (thetas ranging from -5 to 0).

The mean/sigma linking process, using Rasch item difficulties from the current sample and Rasch difficulties in the prior data set (see Appendix D), produced linear coefficients ($A = 1.17$, $B = -2.51$) that were then used to transform all item locations to the Age 4 scale. These linked item locations were then used to estimate ability, referred to as theta, for each participant responding to each linear form.

Final item locations, based on this analysis, for P2P and PN items are presented in Appendix D. New item locations were used for all P2P items in all subsequent analyses.

Seasonal Assessments

Fall Administration

Table 4 summarizes results for Fall assessment of all participating children. For ability estimation, only responses from the blocked administration forms were used (281118001, 281118002, 281118003), resulting in one form used for each participant. As reported in Table 4, theta, an arbitrary value typically ranging from -5 to +5 in initial calibrations like this, demonstrated acceptable skew, kurtosis, and range (c.f., McConnell & Wackerle-Hollman, 2016). Children responded to up to 28 items (n items administered), with an average of 27.74 item-level responses per child. On average, children responded correctly to 75.7% of administered items.

To assess item coverage for operational assessment, we also compared distributions of child ability and item difficulty along the theta scale. Figure 2 presents these results in terms of densities for all receptive P2P and expressive PN items. These maps include all P2P and PN items, not only those administered to participants in this study. Theta values for receptive and expressive items overlap substantially. As a result, an expanded range of ability was captured by the combined P2P and PN item pool.

Winter Administration

Using all receptive and expressive oral language items, item locations developed in the prior calibration, and item selection using an existing CAT engine to produce 25-item forms, Winter assessment was completed for all participants. Any child's assessment might include

items from P2P, PN, or both, depending on performance of the test taker and selection of items by the CAT algorithm.

Table 5 summarizes child performance for Winter assessment. Mean theta for persons increased over Fall assessment (Fall = -0.45, Winter = +0.16), and distributions of child scores continued to represent a roughly normal curve. The number of items was fixed at 25, and did not vary over children. Across all participants, 42.6% of administered items were answered correctly.

Figure 3 presents density distributions of items and children across the theta scale, with shaded curves for P2P items, PN items, and Winter child ability estimates. Compared to Fall distributions, the curve of child ability estimates has shifted right but still is well-represented by items available for selection in the computer-adaptive algorithm.

Spring Administration

Spring assessment replicated Winter procedures, with all children assessed using a computer-adaptive form that could draw content from the full bank of P2P and PN items. Table 6 summarizes results for Spring assessment. Mean theta was slightly higher in spring than winter (winter = +0.16, Spring = +0.48). The number of items was again fixed at 25, and did not vary over children. Participants responded correctly, on average, to 44.6% of administered items.

Figure 4 presents density distributions of items and children across the theta scale, with shaded curves for P2P and PN items, as well as Spring child ability estimates. Again, child ability estimates are well-represented by items available for computer-adaptive selection.

Seasonal Results

To descriptively evaluate the combined scale's ability to model children's growth across seasons, we examined density distributions of child ability estimates for fall, winter, and spring

respectively, along with a density plot for item locations for all P2P and PN items on the unified oral language scale (see Figure 5). While seasonal distributions of child ability overlap, mean values and ranges of distributions appear to increase over successive seasons. Replicating season-specific findings, all seasonal child ability distributions fall within the range of available items for computer-adaptive assessment.

Discussion

This study addressed two primary questions: To what degree do less-difficult receptive items relate to more difficult expressive items in ways that allow for combination of the item types to assess oral language in 3- and 4-year-old children? And To what degree can we develop unified and concurrently calibrated pool of receptive P2P and expressive PN items?

Regarding the first research question, we wanted to determine if new receptive language items demonstrate acceptable psychometric features, and whether these new items are associated with existing, and more difficult, expressive language items in ways that produce a single scale dimension. As a corollary to this primary question, we evaluated whether items in the expanded scale sample a broader range of oral language development, with some degree of overlap in abilities sampled by receptive and expressive items, such that procedures like CAT would have an appropriate item bank for assessing 3- and 4-year-old children.

The evidence lends preliminary support to combining receptive and expressive item types into a single scaled assessment of oral language development for preschool children. In isolation, most receptive items fall within acceptable ranges of difficulty, item-total correlation, and contribution to scale stability. Confirmatory factor analyses evaluating six linear forms that combined PN and P2P items all demonstrated fair to reasonable unidimensional fit, providing at least tentative support for creating a single measure with both item types. Finally, review of item

densities by Rasch scale location consistently demonstrated both overlap and distinctive ability coverage for receptive and expressive item types, thus expanding the range of developmental ability sampled by this new measure.

Given both expected demands of additional reliability, validity, and utility testing needed to support use of this measure in applied settings and emerging demands of preschool programming across 3- and 4-year-old children, measure developers concluded that sufficient evidence was available to continue developing and testing use of two item types in operational assessment.

Regarding the second research question, we evaluated a unified and concurrently calibrated pool of receptive P2P and expressive PN items and found the item calibration, continuity and factor structure of the scale to be robust. Such a combined item pool, when deployed with an existing computer-adaptive testing protocol, is expected to provide a single measure of oral language development for use with children across a longer (typically two year) period of preschool development.

Taken together, results presented here indicate that an expanded approach to assessing growth and development in oral language – an approach that adequately samples and describes the performance of lower-performing younger 3-year-olds as well as higher-performing older 4-year-olds – can be achieved. Additional research will be needed to describe the content, construct, discriminant, and instructional utility of this new measure; initial stages of this research have been completed and will be reported in separate IGDILab technical reports. Additional research will identify criterion-based benchmarks for evaluating individual child performance and recommending need for supplemental or different intervention. Further research is also needed that takes advantage of this measure’s capacity to describe growth over

two full academic years, and that identifies classroom and other characteristics moderating this growth. Finally, both additional research and the development and evaluation of professional development and ongoing coaching and technical assistance mechanisms to support teachers' use of this and similar measures in preschool programs will be needed to bring to scale the potential of these measures, and their full contribution to MTSS and improved outcomes for young children.

Preliminary Results: Contact Authors for Update

References

- Bradfield, T. A., Besner, A. C., Wackerle-Hollman, A. K., Albano, A. D., Rodriguez, M. C., & McConnell, S. R. (2014). Redefining Individual Growth and Development Indicators: Oral language. *Assessment for Effective Intervention, 39*(4), 233-244.
- Carta, J. J. (2019). Introduction to multi-tiered systems of support in early education. In J. J. Carta, & R. M. Young (Eds.), *Multi-tiered systems of support for young children* (pp. 1-14). Brookes Publishing.
- Casenhiser, D. M., Shanker, S. G., & Stieben, J. (2013). Learning through interaction in children with autism: Preliminary data from a social-communication-based intervention. *Autism, 17*(2), 220-241.
- Durán L. Wackerle-Hollman, A. K., Kohlmeier, T. L., Brunner, S. K., Palma, J., & Callard, C.H. (2019). Individual Growth and Development Indicators- Español: Innovation in the development of Spanish oral language general outcome measures. *Early Childhood Research Quarterly, 48*, 155-172.
- Fuchs, L. S., & Deno, S. L. (1991). Paradigmatic distinctions between instructionally relevant measurement models. *Exceptional Children, 57*, 488-500.
- Gonzalez, J. E., Pollard-Durodola, S., Simmons, D. C., Taylor, A. B., Davis, M. J., Kim, M., & Simmons, L. (2010). Developing low-income preschoolers' social studies and science vocabulary knowledge through content-focused shared book reading. *Journal of Research on Educational Effectiveness, 4*(1), 25-52.
- Haring Biel, C., Buzhardt, J., Brown, J. A., Romano, M. K., Lorio, C. M., Windsor, K. S., Kaczmarek, L. A., Gwin, R., Sandall, S. S., & Goldstein, H. (2019). Language interventions taught to caregivers in homes and classrooms: A review of intervention

- and implementation fidelity. *Early Childhood Research Quarterly*, 50(1), 140-156.
<https://doi.org/10.1016/j.ecresq.2018.12.002>
- Hoff, E. (2013). Interpreting the early language trajectories of children from low-SES and language minority homes: Implications for closing achievement gaps. *Developmental Psychology*, 49(1), 4–14. <https://doi.org/10.1037/a0027238>
- Kaiser, A. P., & Roberts, M. Y. (2011). Advances in early communication and language intervention. *Journal of Early Intervention*, 33(4), 298–309.
- Kincaid, A. P., McConnell, S. R., & Wackerle-Hollman, A. K. (2020). Assessing early literacy growth in preschoolers using Individual Growth and Development Indicators. *Assessment for Effective Intervention*, 45(3), 173-183.
doi.org/10.1177/1534508418799173
- Kline, R. B. (2010). *Principles and practice of structural equation modeling* (3rd ed.). Guilford Press.
- Kolen, M. J., & Brennan, R. L. (2014). *Test equating, scaling, and linking*. New York, NY: Springer.
- Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, 5, 831-843.
- McConnell, S. R. (2019). The path forward for multi-tiered systems of support in early education. In J. J. Carta, & R. M. Young (Eds.), *Multi-tiered systems of support for young children* (pp. 253-268). Brookes Publishing.
- McConnell, S. R., Wackerle-Hollman, A. K., Roloff, T. A. B., & Rodriguez, M. (2015). Designing a measurement framework for Response to Intervention in early childhood

programs. *Journal of Early Intervention*, 36(4), 263–280.

<https://doi.org/10.1177/1053815115578559>

McConnell, S., Lease, E., Wackerle-Hollman, A., Albano, A., Schardt, A., Schuster, K., & Cai, L. (2018). Testing initial items for alphabet knowledge, phonological awareness, and oral language for 3-year-old children: Report of Phase 2 results. Technical Report #6, IGDILab, University of Minnesota.

McConnell, S. R., & Wackerle-Hollman, A. K., (2016). Can we measure the transition to reading? General outcome measures and early literacy development from preschool to early elementary grades. *AERA Open* 2(3), 1-15.

McConnell, S. R., Wackerle-Hollman, A. K., Albano, A., Lease, E., Elmquist, M., and Will, K. (2019). Individual Growth & Development Indicators (IGDIs), Language and Early Literacy: Age 3. Unpublished measurement system, IGDILab, University of Minnesota.

Noam, G. G., & Hermann, C. A. (2002). Where education and mental health meet: Developmental prevention and early intervention in schools. *Development and Psychopathology*, 14(4), 861-875.

Pellegrini, A. D., Galda, L., Bartini, M., & Charak, D. (1998). Oral language and literacy learning in context: The role of social relationships. *Merrill-Palmer Quarterly*, 44(1), 38–54.

Pruden, S. M., HirshPasek, K., & Golinkoff, R. M. (2005). The social dimension in language development: A rich history and a new frontier. In P. Marshall & N. Fox (Eds.), *The development of social engagement: Neurobiological perspectives*. Oxford University Press.

- Scarborough, H. S. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory and practice. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research* (Vol. 1, pp. 97–110). Guilford.
- Schuster, K., Cai, L., McConnell, S., & Wackerle-Hollman, A. (2017). Defining a domain of assessment for language and early literacy in 3-year-olds: oral language. Technical Report #3, IGDILab, University of Minnesota.
- Schuster, K., Schardt, A., Brandes, D., Cai, L., Albano, A., Wackerle-Hollman, A., & McConnell, S. (2017). Expanding Individual Growth and Development Indicators of language and early literacy for universal screening in multi-tiered systems of support with three-year-olds. Technical Report #4, IGDILab, University of Minnesota.
- Walker, D., Greenwood, C., Hart, B., & Carta, J. (1994). Prediction of school outcomes based on early language production and socioeconomic factors. *Child Development*, *65*(2), 606–621.
- Wang, Q., Rodriguez, M., Wackerle-Hollman, A., & McConnell, S. (2017). IGDI-APEL Computer Adaptive Test Program Manual. Unpublished technical report, IGDILab, University of Minnesota, Minneapolis MN.
- Wasik, B. A., & Hindman, A. H. (2011). Improving vocabulary and pre-literacy skills of at-risk preschoolers through teacher professional development. *Journal of Educational Psychology*, *103*(2), 455–469. <https://doi.org/10.1037/a0023067>
- Zucker, T. A., Solari, E. J., Landry, S. H., & Swank, P. R. (2013). Effects of a Brief Tiered Language Intervention for Prekindergartners at Risk. *Early Education & Development*, *24*(3), 366–392. <https://doi.org/10.1080/10409289.2012.664763>

Table 1*Participant Characteristics*

Variable	<i>n</i> (%) unless specified otherwise	Missing data <i>n</i> (%)
Child age (in months)		1 (0.22)
Mean (<i>SD</i>)	43.59 (3.61)	
Range	36.53-52.24	
Gender		31 (6.90)
Boy	198 (44.10)	
Girl	220 (49.00)	
Race/Ethnicity		34 (7.57)
American Indian	3 (0.67)	
Asian/Pacific Islander	35 (7.80)	
Asian/Pacific Islander & White	8 (1.78)	
Asian/Pacific Islander, White & American Indian	1 (0.22)	
Black	165 (36.75)	
Black & American Indian	2 (0.45)	
Black, Asian/Pacific Islander & White	1 (0.22)	
Black & Latino	1 (0.22)	
Black & White	4 (0.89)	
Latino	8 (1.78)	
White	178 (39.64)	
White & Latino	8 (1.78)	
White & American Indian	1 (0.22)	
Home language		35 (7.80)
Bosnian	1 (0.22)	
Cantonese	1 (0.22)	
Enbosh	3 (0.67)	
English	378 (84.19)	
English & Oromo	1 (0.22)	
Hausa	1 (0.22)	
Hindi	1 (0.22)	
Hmong	4 (0.89)	
Japanese	1 (0.22)	
Mandarin	1 (0.22)	
Marathi	1 (0.22)	
Nepali	1 (0.22)	
Oromo	1 (0.22)	
Russian	1 (0.22)	
Somali	5 (1.11)	
Spanish	5 (1.11)	

Telugu	1 (0.22)	
Tibetan	1 (0.22)	
Vietnamese	3 (0.67)	
Wollof	2 (0.45)	
Yoruba	1 (0.22)	
IEP status		28 (6.24)
Yes	55 (12.25)	
No	366 (81.51)	
Free/reduced-price lunch eligibility		172 (38.31)
Yes	71 (15.81)	
No	206 (45.88)	

Note. Participant characteristics include 449 children who completed at least one oral language assessment during any season.

Preliminary Results: Contact Authors for Update

Table 2

Summary of Item-Level Characteristics

Form	<i>p</i>	<i>SD</i>	<i>N</i> responses	<i>N</i> missing	ITC	CITC	AID
8001							
<i>M</i>	0.76	0.37	212.75	9.25	0.35	0.26	0.75
<i>SD</i>	0.18	0.13	12.09	12.09	0.10	0.09	0.01
Min	0.34	0.12	180	0	0.11	0.08	0.73
Max	0.99	0.50	222	42	0.55	0.45	0.75
8002							
<i>M</i>	0.77	0.36	99.93	2.07	0.41	0.33	0.79
<i>SD</i>	0.19	0.12	2.97	2.97	0.09	0.11	0.01
Min	0.44	0.14	92	0	0.19	0.06	0.77
Max	0.98	0.50	102	10	0.59	0.51	0.80
8003							
<i>M</i>	0.81	0.31	109.56	2.44	0.34	0.26	0.73
<i>SD</i>	0.20	0.15	3.32	3.32	0.15	0.14	0.01
Min	0.38	0.10	98	0	0.05	0.02	0.71
Max	0.99	0.50	112	14	0.59	0.49	0.74
8004							
<i>M</i>	0.74	0.39	207.26	6.74	0.39	0.310	0.79
<i>SD</i>	0.18	0.11	9.23	9.23	0.11	0.109	0.01
Min	0.28	0.18	181	0	0.18	0.116	0.78
Max	0.97	0.50	214	33	0.55	0.485	0.80
8005							
<i>M</i>	0.82	0.34	98.78	2.22	0.41	0.33	0.77
<i>SD</i>	0.16	0.18	3.27	3.27	0.11	0.12	0.01
Min	0.37	0.10	87	0	0.16	0.04	0.76
Max	0.99	0.50	101	14	0.62	0.58	0.79
8006							
<i>M</i>	0.79	0.35	106.30	2.70	0.30	0.20	0.63
<i>SD</i>	0.17	0.12	3.67	3.67	0.13	0.13	0.01
Min	0.39	0.10	92	0	0.07	0.09	0.60
Max	0.99	0.50	109	17	0.55	0.46	0.67

Note. *p* = Item difficulty, or mean scored response; *SD* = standard deviation; *N* responses = number of children responding; *N* missing = Number of children not completing an assigned form; ITC = Item-Total Correlation; CITC = Corrected Item-Total Correlation; AID = alpha internal consistency for the full scale if the item were deleted; *M* = mean; Min = minimum; Max = maximum.

Table 3

Confirmatory Factor Analysis, Linear Forms, Both Item Types

Form and Model	ChiSq	RMSEA	CFI	WRMR
Form 1 all items	sig	0.06	0.83	1.12
Form 2 all items	sig	0.04	0.97	0.87
Form 3 all items	sig	0.11	0.88	1.31
Form 4 all items	non sig	0.02	0.98	0.87
Form 5 all items	sig	0.07	0.90	1.01
Form 6 all items	sig	0.08	0.78	1.13

Confirmatory Factor Analysis, Linear Forms, Point to Picture Items Only

Form and Model	ChiSq	RMSEA	CFI	WRMR
Form 1, P2P items	sig	0.063	0.826	1.131
Form 2, P2P items	non sig	<.001	1.000	0.713
Form 4, P2P items	non sig	.019	0.978	0.848
Form 5, P2P items	sig	0.066	0.929	0.923
Form 6, P2P items	non sig	0.034	0.937	0.825
Form 3, P2P items	Sig	.097	0.960	1.062

Confirmatory Factor Analysis, Linear Forms, Picture Naming Items Only

Form and Model	ChiSq	RMSEA	CFI	WRMR
Form 1, PN items	sig	0.046	0.949	0.848
Form 2, PN items	sig	.067	0.889	0.826
Form 4, PN items	non sig	<.001	1.000	0.590
Form 5, PN items	non sig	<.001	1.000	0.587
Form 6, PN items	sig	0.067	0.897	0.864
Form 3, PN items	sig	0.116	0.800	1.187

Note. ChiSq = chi-square test; Sig = significant at $p \leq 0.05$; RMSEA = root mean square error or approximation; CFI = comparative fix index; WRMR = weighted root mean square residual.

Table 4

Descriptive Statistics for Fall OL Ability Estimation

	<i>M</i>	Median	<i>SD</i>	Skew.	Kurtosis	Min.	Max.	<i>n</i>	Missing
Ability estimate (theta)	-0.45	-0.32	0.90	-0.26	2.83	-3.41	1.50	432	0
<i>n</i> items administered	27.74	28	0.44	-1.11	2.23	27	28	432	0
<i>n</i> items correct	20.99	22	3.93	-0.70	3.23	6	28	432	0
Proportion correct	0.76	0.79	0.14	-0.71	3.20	0.21	1.00	432	0

Note. Total number correct (*n items correct*) treats missing items as zeros. Proportion correct was calculated as total items correct over number of items administered. Skew = skewness; Min = minimum; Max = maximum.

Preliminary Results: Contact Authors for Update

Table 5

Descriptive Statistics for Winter OL Ability Estimation

	<i>M</i>	Median	<i>SD</i>	Skew.	Kurtosis	Min.	Max.	<i>n</i>	Missing
Ability estimate (theta)	0.16	0.17	1.01	-0.74	4.91	-4.00	2.82	406	0
<i>n</i> items administered	25	25	0	NaN	NaN	25	25	406	0
<i>n</i> items correct	10.65	11.00	2.24	-0.7	5.27	0	16	406	0
Proportion correct	0.43	0.44	0.09	-0.701	5.27	0	0.64	406	0

Note. Total number correct (*n* items correct) treats missing items as zeros. Proportion correct was calculated as total items correct over number of items administered. Skew = skewness; Min = minimum; Max = maximum.

Preliminary Results: Contact Authors for Update

Table 6

Descriptive Statistics for Spring OL Ability Estimation

	<i>M</i>	Median	<i>SD</i>	Skew.	Kurtosis	Min.	Max.	<i>n</i>	Missing
Ability estimate (theta)	0.48	0.53	1.03	-0.90	5.13	-4.00	3.41	411	0
<i>n</i> items administered	25	25	0	NaN	NaN	25	25	411	0
<i>n</i> items correct	11.14	11	2.26	-0.64	5.48	0	19	411	0
Proportion correct	0.45	0.44	0.09	-0.64	5.48	0	0.76	411	0

Note. Total number correct (*n* items correct) treats missing items as zeros. Proportion correct was calculated as total items correct over number of items administered. Skew = skewness; Min = minimum; Max = maximum.

Preliminary Results: Contact Authors for Update

Figure 1

Relations in Age 3 and Age 4 Picture Naming Item Difficulties

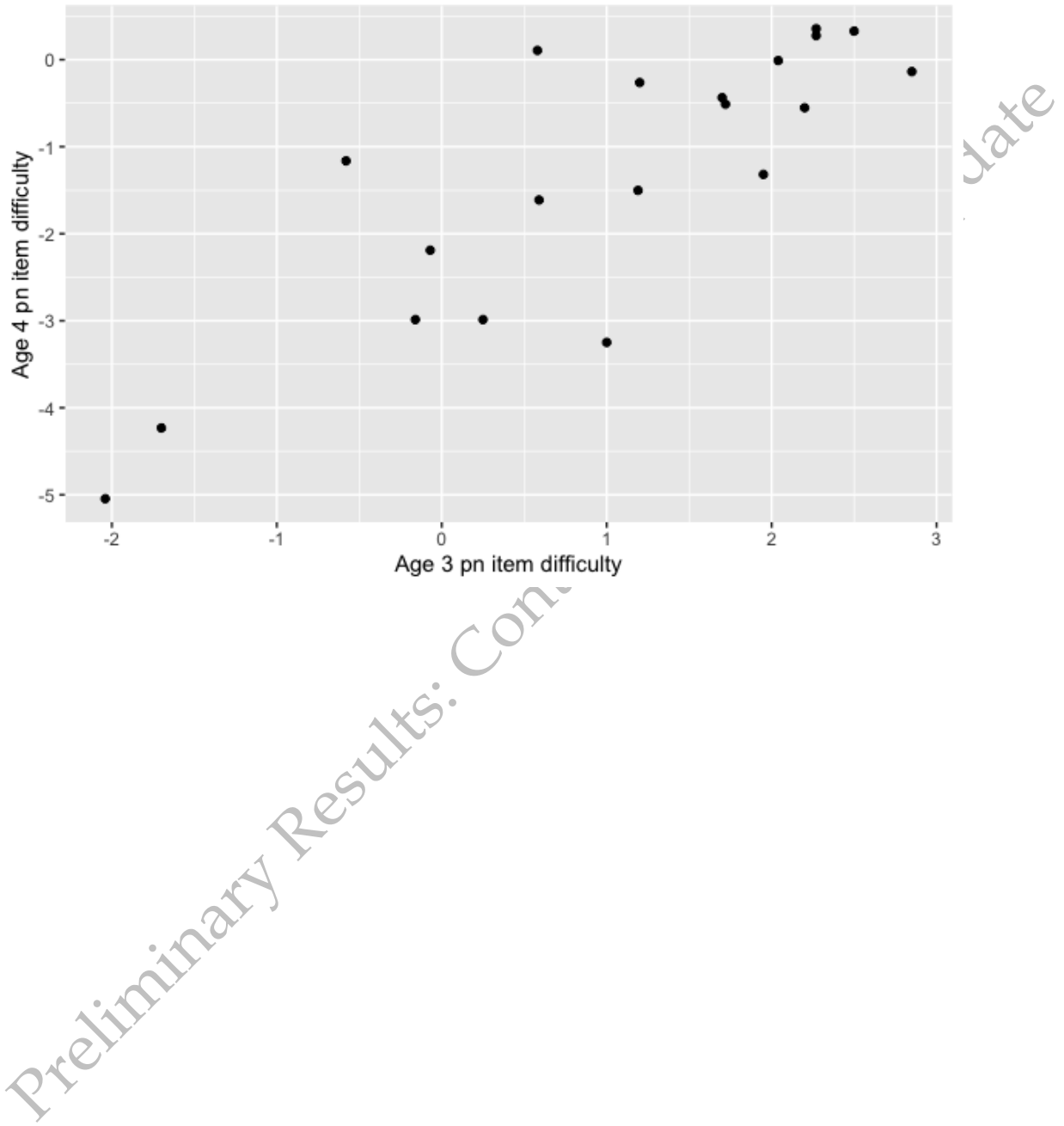


Figure 2

Distributions of Ability and Item Difficulty Along the Theta Scale for Fall

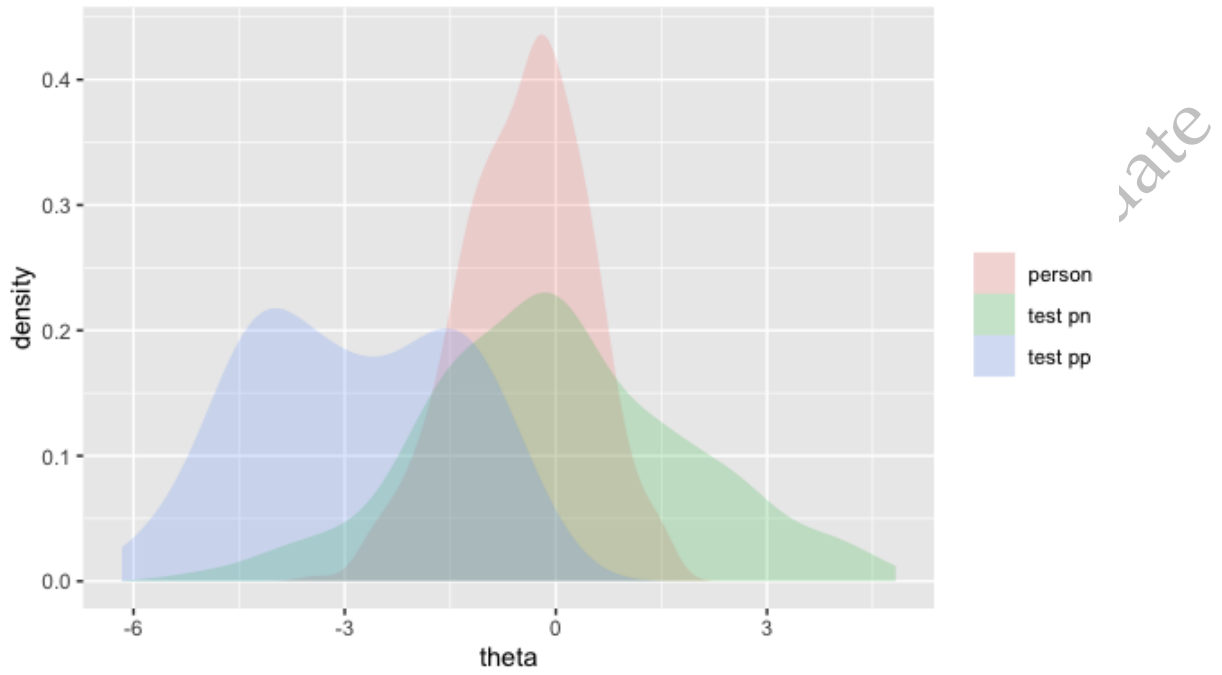


Figure 3

Distributions of Ability and Item Difficulty Along the Theta Scale for Winter

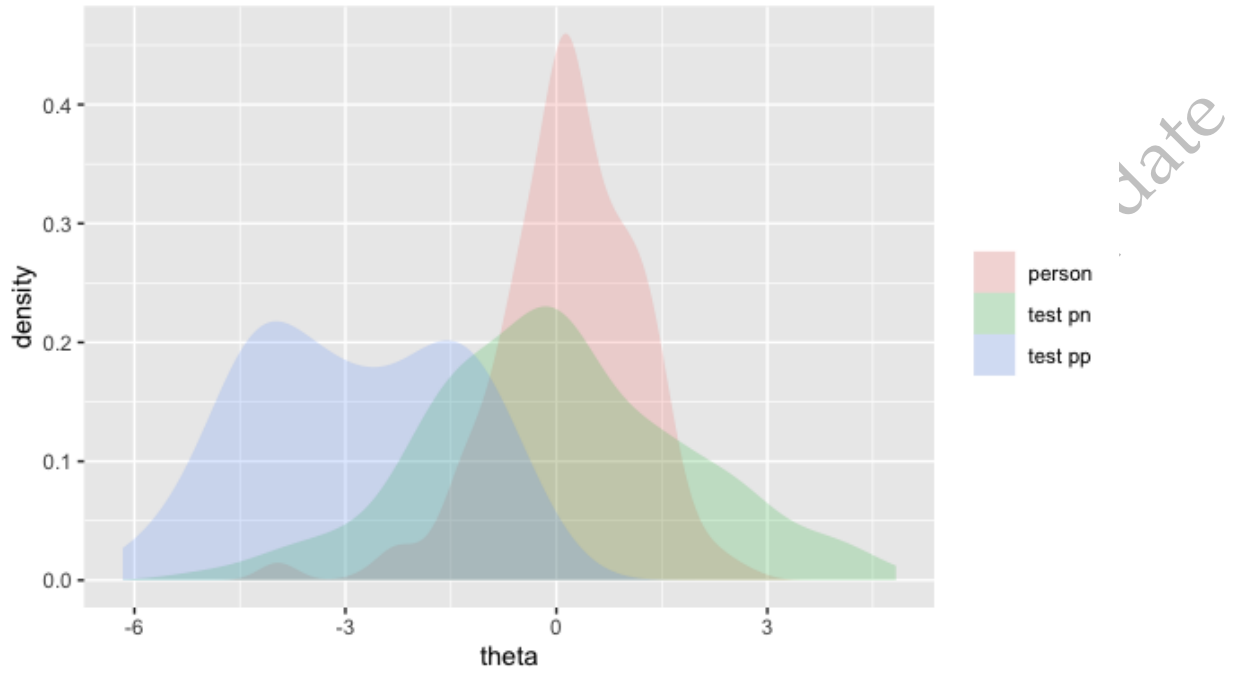
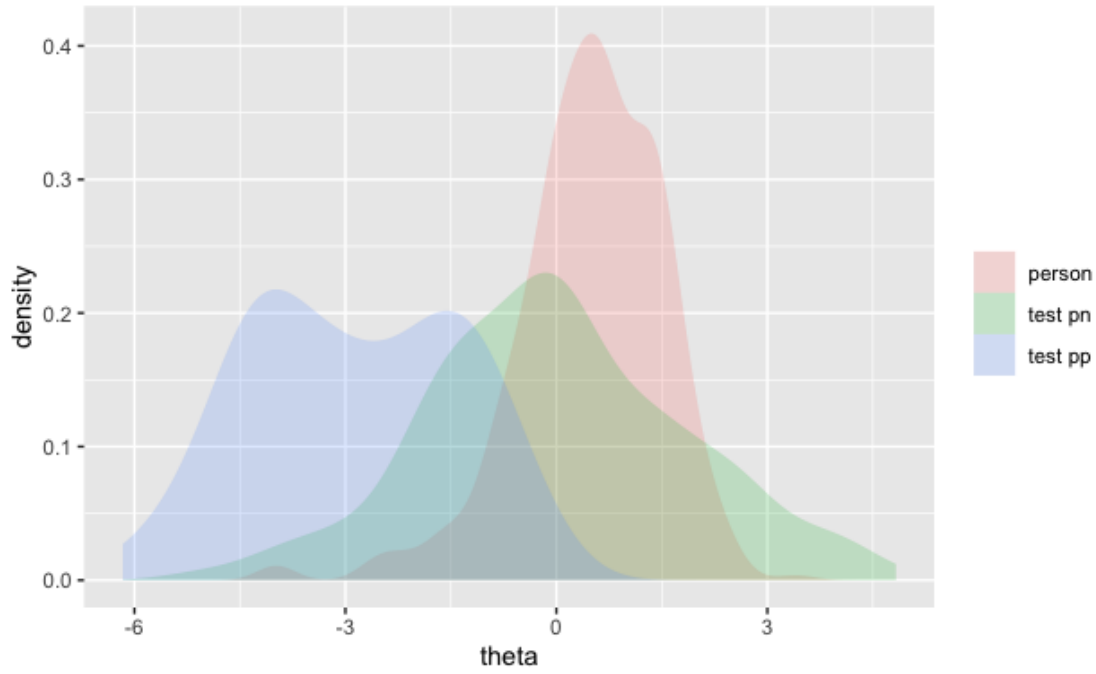


Figure 4

Distributions of Ability and Item Difficulty Along the Theta Scale for Spring

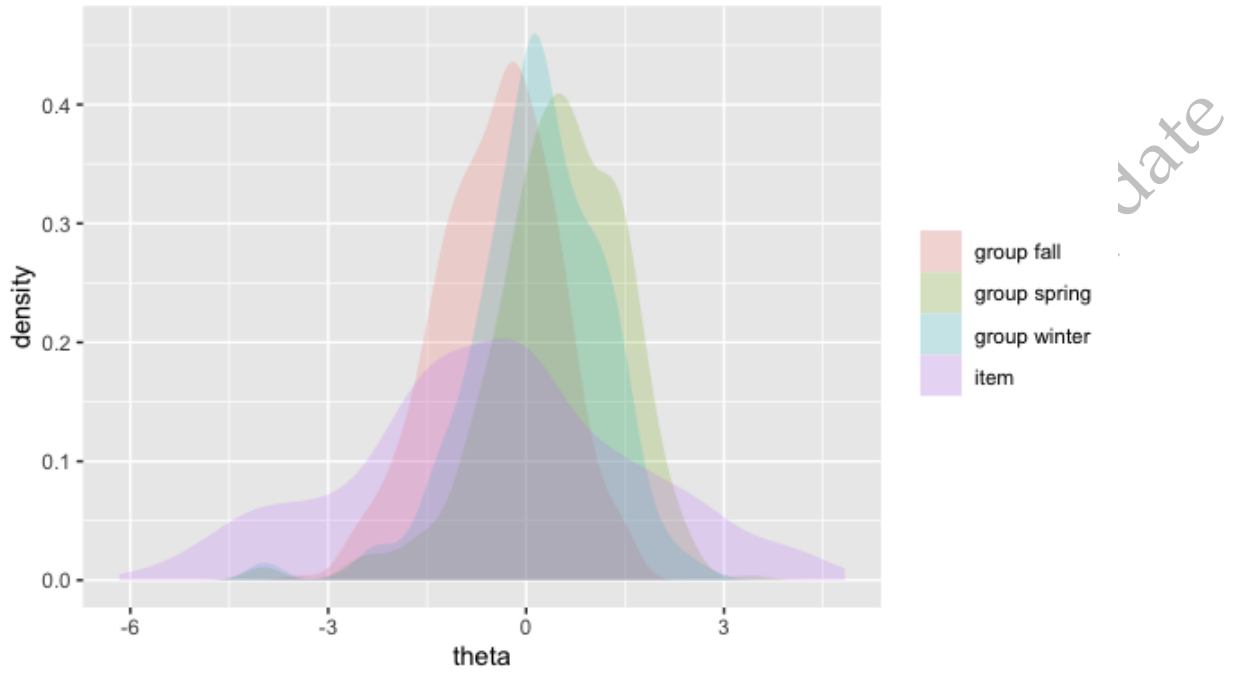


Preliminary Results: Con

date

Figure 5

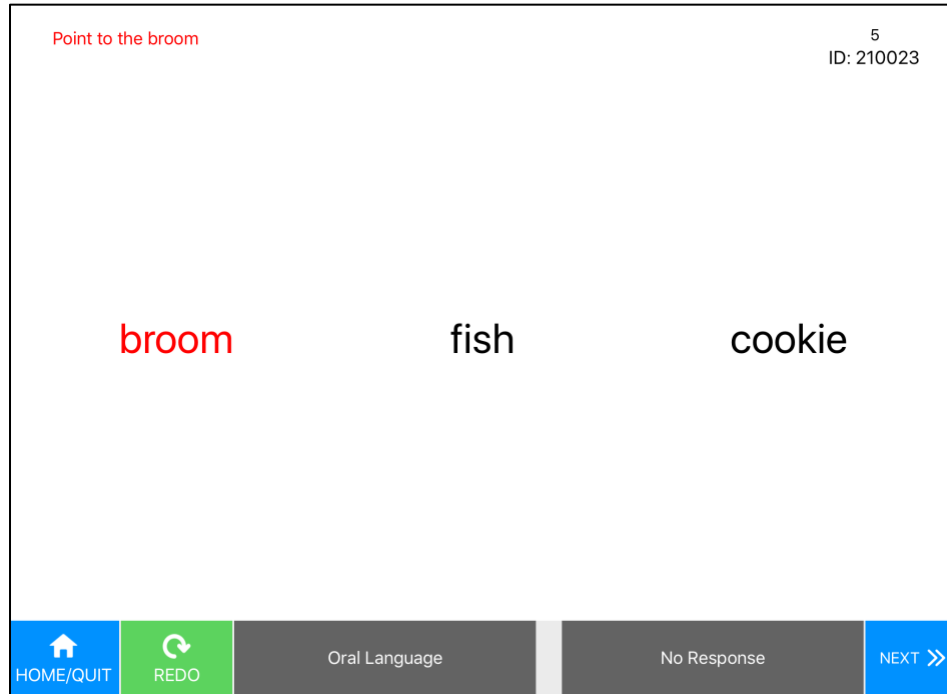
Density Distributions of Child Ability Estimates by Season and Density Plot for Item Locations for All P2P And PN Items on the Unified Oral Language Scale



Preliminary Results: Con

Appendix A Item Examples

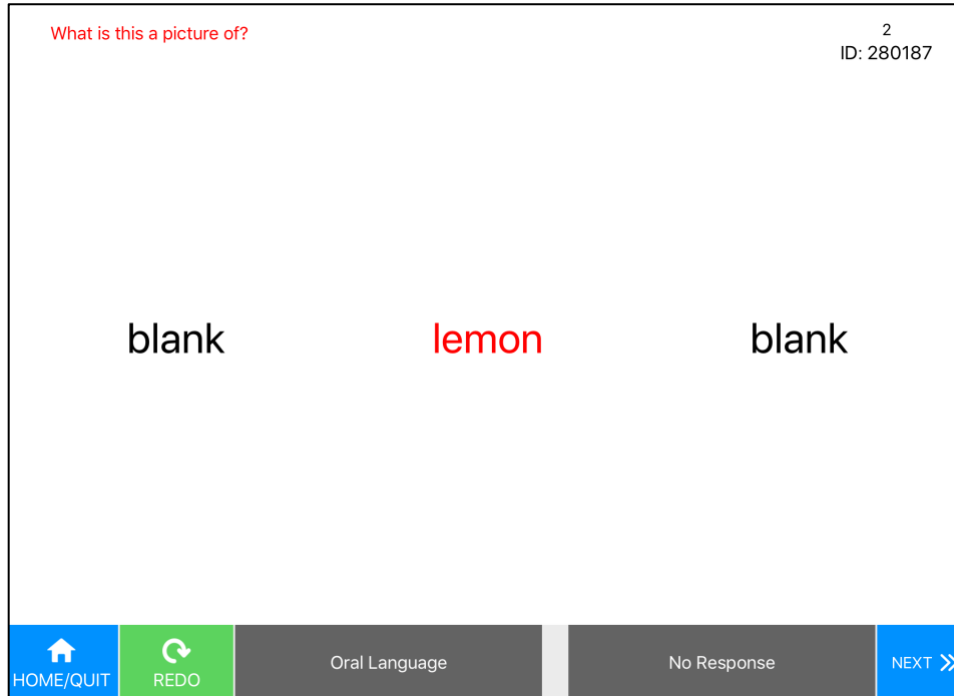
Examiner iPad- Point to Picture



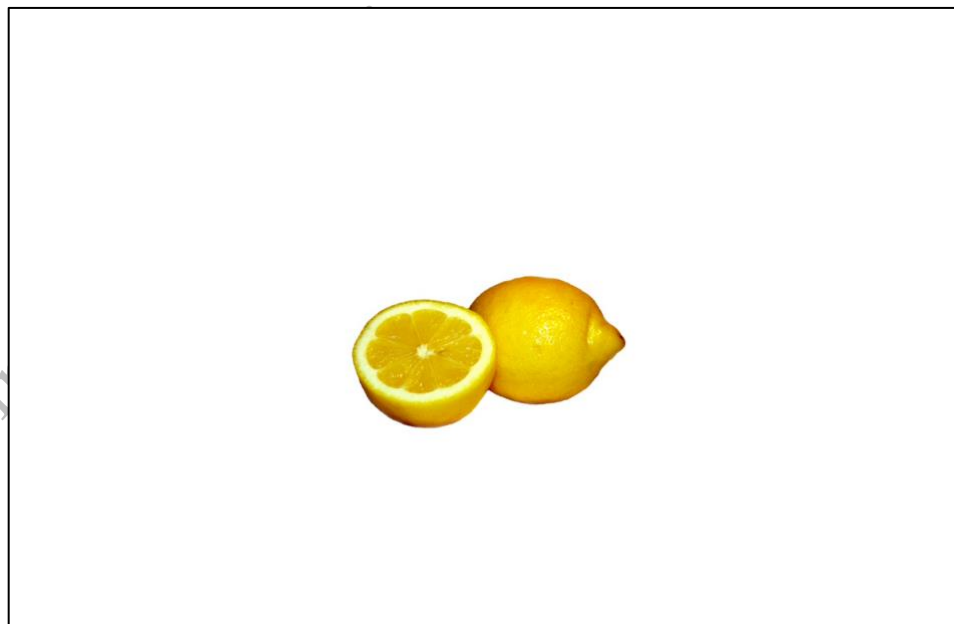
Child iPad- Point to Picture



Examiner iPad- Picture Naming



Child iPad- Picture Naming



Appendix B
Forms for Fall Administration

Form IDs:	281118001	281118002	281118003	281118004	281118005	281118006
Format:	LINEAR	LINEAR	LINEAR	LINEAR	LINEAR	LINEAR
Item IDs:	210008	210008	210008	210056	210056	210056
	210030	210030	210030	280096	280096	280096
	210016	210016	210016	210048	210048	210048
	210058	210058	210058	280105	280105	280105
	210063	210063	210063	210032	210032	210032
	210049	210049	210049	280150	280150	280150
	210052	210052	210052	210046	210059	210022
	210039	210039	210039	210015	210015	210015
	210038	210038	210038	280225	280225	280225
	210006	210018	210018	210041	210020	210003
	210018	210011	210026	210029	210029	210029
	210040	210065	210044	280014	280014	280014
	210051	210060	210019	210066	210023	210025
	210054	210033	210013	210036	210036	210036
	210012	210053	210050	280296	280296	280296
	210031	210024	210007	210045	210009	210004
	210057	210062	210064	210061	210061	210061
	210014	210047	280245	280092	280092	280092
	280245	280245	280091	210021	210005	210028
	280091	280091	280217	210055	210055	210055
	280217	280217	280142	280134	280134	280134
	280142	280142	280164	210034	210017	210043
	280164	280164	280083	210010	210010	210010
	280083	280083	280257	280248	280248	280248
	280257	280257	280299	210035	210027	210037
	280299	280299	280333	210042	210042	210042
	280333	280333	280200	280109	280109	280109
	280200	280200				

Appendix C
Scale Characteristics Item Level Tables

This appendix presents six tables, each with item-level difficulty, standard deviation, number of unique respondents, number of missing respondents, item-total correlation, corrected item-total correlation, and *alpha* internal consistency if that item is removed for 27 or 28 items presented in linear forms during Fall administration.

Table C1*Item Analysis for Fall Oral Language Form 1*

Item	<i>M</i>	<i>SD</i>	<i>N</i>	Missing	ITC	CITC	AID
210006	0.760	0.428	221	1	0.497	0.411	0.738
210008	0.703	0.458	219	3	0.306	0.198	0.752
210012	0.743	0.438	218	4	0.398	0.301	0.746
210014	0.959	0.198	222	0	0.295	0.250	0.748
210016	0.946	0.227	221	1	0.251	0.198	0.749
210018	0.950	0.218	221	1	0.315	0.266	0.750
210030	0.937	0.244	222	0	0.271	0.215	0.751
210031	0.955	0.208	222	0	0.273	0.225	0.749
210038	0.923	0.267	222	0	0.231	0.167	0.752
210039	0.986	0.116	221	1	0.105	0.077	0.754
210040	0.630	0.484	219	3	0.445	0.342	0.742
210049	0.959	0.198	222	0	0.306	0.261	0.749
210051	0.739	0.440	222	0	0.301	0.198	0.751
210052	0.910	0.287	222	0	0.252	0.185	0.752
210054	0.768	0.423	220	2	0.473	0.386	0.746
210057	0.582	0.494	220	2	0.401	0.292	0.749
210058	0.964	0.187	222	0	0.332	0.290	0.748
210063	0.793	0.406	222	0	0.478	0.396	0.738
280083	0.443	0.498	194	28	0.405	0.292	0.743
280091	0.852	0.356	203	19	0.367	0.286	0.744
280142	0.718	0.451	202	20	0.340	0.232	0.746
280164	0.731	0.444	201	21	0.287	0.179	0.751
280200	0.573	0.496	199	23	0.477	0.372	0.739
280217	0.844	0.364	205	17	0.233	0.142	0.752
280245	0.542	0.499	203	19	0.270	0.149	0.753
280257	0.466	0.500	193	29	0.343	0.226	0.746

280299	0.344	0.477	180	42	0.502	0.407	0.738
280333	0.563	0.497	199	23	0.547	0.450	0.732

Preliminary Results: Contact Authors for Update

Table C2*Item Analysis for Fall Oral Language Form 2*

Item	<i>M</i>	<i>SD</i>	<i>N</i>	Missing	ITC	CITC	AID
210008	0.673	0.471	101	1	0.409	0.305	0.786
210011	0.873	0.335	102	0	0.568	0.510	0.776
210016	0.941	0.236	102	0	0.442	0.395	0.784
210018	0.950	0.218	101	1	0.519	0.479	0.783
210024	0.970	0.171	101	1	0.343	0.306	0.788
210030	0.950	0.218	101	1	0.312	0.260	0.788
210033	0.891	0.313	101	1	0.379	0.312	0.786
210038	0.922	0.270	102	0	0.496	0.444	0.783
210039	0.980	0.139	102	0	0.450	0.423	0.787
210047	0.644	0.481	101	1	0.377	0.272	0.788
210049	0.941	0.236	102	0	0.553	0.512	0.781
210052	0.931	0.254	102	0	0.469	0.420	0.783
210053	0.765	0.426	102	0	0.450	0.362	0.782
210058	0.951	0.217	102	0	0.349	0.301	0.787
210060	0.902	0.299	102	0	0.402	0.339	0.785
210062	0.618	0.488	102	0	0.383	0.276	0.790
210063	0.861	0.347	101	1	0.403	0.329	0.785
210065	0.550	0.500	100	2	0.190	0.062	0.802
280083	0.630	0.485	92	10	0.465	0.368	0.782
280091	0.863	0.346	95	7	0.320	0.244	0.788
280142	0.818	0.388	99	3	0.293	0.204	0.790
280164	0.614	0.489	101	1	0.428	0.325	0.785
280200	0.469	0.502	98	4	0.325	0.213	0.791
280217	0.871	0.337	101	1	0.439	0.370	0.782
280245	0.446	0.500	101	1	0.304	0.190	0.793
280257	0.441	0.499	93	9	0.408	0.300	0.786
280299	0.436	0.499	94	8	0.440	0.336	0.782
280333	0.588	0.495	97	5	0.589	0.504	0.773

Table C3*Item Analysis for Fall Oral Language Form 3*

Item	<i>M</i>	<i>SD</i>	<i>N</i>	Missing	ITC	CITC	AID
210007	0.955	0.207	112	0	0.409	0.354	0.729
210008	0.794	0.406	107	5	0.351	0.232	0.735
210013	0.982	0.133	112	0	0.251	0.212	0.734
210016	0.964	0.187	111	1	0.155	0.098	0.737
210018	0.964	0.187	111	1	0.403	0.354	0.728
210019	0.982	0.133	112	0	0.168	0.129	0.737
210026	0.973	0.162	112	0	0.114	0.065	0.739
210030	0.991	0.095	111	1	0.047	0.018	0.739
210038	0.955	0.207	112	0	0.144	0.082	0.737
210039	0.982	0.134	111	1	0.230	0.191	0.734
210044	0.865	0.343	111	1	0.289	0.189	0.733
210049	0.946	0.226	112	0	0.358	0.296	0.728
210050	0.964	0.187	111	1	0.447	0.400	0.726
210052	0.920	0.273	112	0	0.213	0.132	0.737
210058	0.991	0.095	111	1	0.310	0.284	0.734
210063	0.786	0.412	112	0	0.526	0.426	0.718
210064	0.732	0.445	112	0	0.521	0.412	0.716
280083	0.750	0.435	108	4	0.589	0.490	0.710
280091	0.846	0.363	104	8	0.425	0.323	0.726
280142	0.809	0.395	110	2	0.511	0.412	0.718
280164	0.583	0.495	108	4	0.303	0.155	0.739
280200	0.382	0.488	110	2	0.353	0.214	0.736
280217	0.890	0.314	109	3	0.173	0.079	0.741
280245	0.524	0.502	105	7	0.392	0.255	0.733
280257	0.429	0.497	105	7	0.395	0.258	0.734
280299	0.469	0.502	98	14	0.589	0.478	0.710
280333	0.532	0.501	109	3	0.561	0.443	0.712

Table C4*Item Analysis for Fall Oral Language Form 4*

Item	<i>M</i>	<i>SD</i>	<i>N</i>	Missing	ITC	CITC	AID
210010	0.648	0.479	213	1	0.386	0.287	0.791
210015	0.850	0.358	213	1	0.340	0.263	0.789
210021	0.757	0.430	210	4	0.423	0.337	0.784
210029	0.925	0.264	213	1	0.207	0.147	0.794
210032	0.915	0.279	212	2	0.301	0.240	0.791
210034	0.512	0.501	213	1	0.227	0.114	0.799
210035	0.765	0.425	213	1	0.274	0.180	0.794
210036	0.743	0.438	214	0	0.457	0.372	0.784
210041	0.939	0.239	214	0	0.335	0.285	0.789
210042	0.901	0.299	212	2	0.320	0.256	0.789
210045	0.640	0.481	214	0	0.505	0.416	0.781
210046	0.659	0.475	211	3	0.441	0.347	0.785
210048	0.850	0.358	213	1	0.385	0.311	0.786
210055	0.664	0.474	211	3	0.370	0.271	0.789
210056	0.967	0.179	213	1	0.220	0.180	0.791
210061	0.771	0.421	214	0	0.514	0.438	0.782
210066	0.662	0.474	213	1	0.488	0.398	0.782
280014	0.967	0.179	212	2	0.365	0.329	0.788
280092	0.409	0.493	181	33	0.498	0.407	0.780
280096	0.917	0.277	204	10	0.178	0.115	0.794
280105	0.962	0.192	210	4	0.322	0.282	0.789
280109	0.707	0.456	198	16	0.537	0.455	0.777
280134	0.651	0.478	195	19	0.497	0.406	0.780
280150	0.681	0.467	191	23	0.543	0.459	0.778
280225	0.819	0.386	204	10	0.552	0.485	0.777
280248	0.284	0.452	190	24	0.265	0.166	0.793
280296	0.426	0.496	195	19	0.488	0.394	0.781

Table C5*Item Analysis for Fall Oral Language Form 5*

Item	<i>M</i>	<i>SD</i>	<i>N</i>	Missing	ITC	CITC	AID
210005	0.950	0.219	100	1	0.410	0.361	0.772
210009	0.930	0.256	100	1	0.428	0.367	0.769
210010	0.752	0.434	101	0	0.294	0.185	0.778
210015	0.869	0.339	99	2	0.372	0.288	0.772
210017	0.990	0.100	101	0	0.333	0.309	0.776
210020	0.901	0.300	101	0	0.404	0.335	0.770
210023	0.870	0.338	100	1	0.499	0.428	0.764
210027	0.960	0.198	99	2	0.507	0.467	0.769
210029	0.890	0.314	100	1	0.455	0.382	0.770
210032	0.870	0.338	100	1	0.496	0.425	0.766
210036	0.723	0.450	101	0	0.475	0.376	0.765
210042	0.950	0.219	100	1	0.347	0.295	0.772
210048	0.832	0.376	101	0	0.374	0.284	0.772
210055	0.713	0.455	101	0	0.156	0.037	0.788
210056	0.950	0.218	101	0	0.615	0.577	0.763
210059	0.717	0.453	99	2	0.421	0.310	0.772
210061	0.842	0.367	101	0	0.514	0.438	0.764
280014	0.969	0.173	98	3	0.512	0.476	0.770
280092	0.494	0.503	87	14	0.294	0.165	0.782
280096	0.919	0.274	99	2	0.303	0.236	0.775
280105	0.980	0.141	100	1	0.285	0.251	0.776
280109	0.857	0.352	98	3	0.513	0.442	0.762
280134	0.687	0.466	99	2	0.508	0.409	0.762
280150	0.719	0.452	96	5	0.381	0.275	0.771
280225	0.806	0.397	98	3	0.515	0.433	0.764
280248	0.367	0.485	90	11	0.243	0.123	0.784
280296	0.505	0.503	97	4	0.404	0.279	0.774

Table C6*Item Analysis for Fall Oral Language Form 6*

Item	<i>M</i>	<i>SD</i>	<i>N</i>	Missing	ITC	CITC	AID
210003	0.954	0.210	109	0	0.123	0.057	0.641
210004	0.972	0.166	107	2	0.328	0.280	0.630
210010	0.661	0.476	109	0	0.239	0.091	0.639
210015	0.861	0.347	108	1	0.252	0.145	0.639
210022	0.862	0.346	109	0	0.316	0.214	0.628
210025	0.954	0.210	109	0	0.358	0.299	0.627
210028	0.626	0.486	107	2	0.273	0.122	0.642
210029	0.899	0.303	109	0	0.334	0.246	0.623
210032	0.898	0.304	108	1	0.263	0.171	0.632
210036	0.692	0.464	107	2	0.409	0.275	0.624
210037	0.917	0.277	109	0	0.103	0.016	0.644
210042	0.945	0.229	109	0	0.112	0.040	0.643
210043	0.519	0.502	108	1	0.364	0.216	0.631
210048	0.935	0.248	107	2	0.328	0.255	0.629
210055	0.645	0.481	107	2	0.198	0.048	0.652
210056	0.897	0.305	107	2	0.423	0.340	0.620
210061	0.879	0.328	107	2	0.430	0.339	0.621
280014	0.953	0.212	107	2	0.188	0.122	0.636
280092	0.565	0.498	92	17	0.518	0.391	0.602
280096	0.870	0.337	108	1	0.158	0.053	0.646
280105	0.991	0.097	106	3	0.155	0.125	0.639
280109	0.830	0.377	106	3	0.547	0.455	0.607
280134	0.683	0.468	101	8	0.419	0.286	0.615
280150	0.740	0.441	104	5	0.431	0.310	0.614
280225	0.821	0.385	106	3	0.368	0.257	0.630
280248	0.390	0.490	100	9	0.070	-0.087	0.670
280296	0.462	0.501	104	5	0.400	0.254	0.624

Appendix D

As noted in the text, Point to Picture (P2P) items, initially calibrated as a single item type, were recalibrated to place item locations on the scale originally set by Picture Naming items. To create a single scale for describing oral language performance across the ability scale now sampled by receptive and expressive items, Rasch modeling was used to calibrate receptive P2P items onto the existing expressive PN scale, where the PN items served as anchors with known parameter values based on prior investigations (Bradfield et al., 2014). This was achieved by fitting a Rasch model to the full OL Fall data set across all six forms, obtaining item difficulties for all items, P2P and PN, and then linking parameters to the existing PN scale using a mean/sigma transformation (Kolen & Brennan, 2014). The mean/sigma linking process, using difficulties from the current sample and difficulties in the prior data set, produced linear coefficients ($A = 1.17$, $B = -2.51$) that were then used to transform all item locations to the Age 4 scale. These linked item locations were then used to estimate ability, referred to as theta, for each participant responding to each linear form.

The table below contains the final calibration results for oral language, with original item locations labeled `b_original` and item locations linked to the picture naming scale labeled `b_linked`.

itemid	b_original	b_linked	task
210003	-1.28	-4.0052272	pp
210004	-1.81	-4.6241861	pp
210005	-1.37	-4.1103334	pp
210006	0.57	-1.8447101	pp
210007	-1.39	-4.1336904	pp
210008	0.85	-1.5177129	pp
210009	-0.87	-3.5264099	pp
210010	1.13	-1.1907157	pp
210011	-0.23	-2.7789877	pp
210012	0.65	-1.7512823	pp
210013	-2.37	-5.2781805	pp
210014	-1.83	-4.6475431	pp
210015	-0.09	-2.6154892	pp
210016	-1.33	-4.0636195	pp
210017	-3.13	-6.1657443	pp
210018	-1.48	-4.2387966	pp
210019	-2.37	-5.2781805	pp
210020	-0.52	-3.1176634	pp
210021	0.55	-1.8680670	pp
210022	-0.01	-2.5220614	pp
210023	-0.16	-2.6972384	P2P

itemid	b_original	b_linked	task
210024	-2.00	-4.8460771	P2P
210025	-1.28	-4.0052272	P2P
210026	-1.94	-4.7760063	P2P
210027	-1.61	-4.3906167	P2P
210028	1.50	-0.7586123	P2P
210029	-0.64	-3.2578050	P2P
210030	-1.41	-4.1570473	P2P
210031	-1.70	-4.4957230	P2P
210032	-0.56	-3.1643773	P2P
210033	-0.45	-3.0359141	P2P
210034	1.89	-0.3031519	P2P
210035	0.42	-2.0198871	P2P
210036	0.84	-1.5293914	P2P
210037	-0.62	-3.2344481	P2P
210038	-1.03	-3.7132654	P2P
210039	-2.61	-5.5584638	P2P
210040	1.30	-0.9921817	P2P
210041	-1.20	-3.9117994	P2P
210042	-0.90	-3.5614453	P2P
210043	2.02	-0.1513318	P2P
210044	-0.07	-2.5921322	P2P
210045	1.21	-1.0972879	P2P
210046	1.15	-1.1673588	P2P
210047	1.33	-0.9571463	P2P
210048	-0.22	-2.7673093	P2P
210049	-1.44	-4.1920827	P2P
210050	-1.63	-4.4139737	P2P
210051	0.61	-1.7979962	P2P
210052	-0.81	-3.4563391	P2P
210053	0.66	-1.7396038	P2P
210054	0.42	-2.0198871	P2P
210055	1.14	-1.1790372	P2P
210056	-1.29	-4.0169057	P2P
210057	1.47	-0.7936477	P2P
210058	-1.88	-4.7059354	P2P
210059	1.07	-1.2607865	P2P

itemid	b_original	b_linked	task
210060	-0.71	-3.3395543	P2P
210061	0.24	-2.2300996	P2P
210062	1.51	-0.7469338	P2P
210063	0.27	-2.1950642	P2P
210064	0.84	-1.5293914	P2P
210065	1.82	-0.3849012	P2P
210066	1.04	-1.2958220	P2P
280003	NA	-1.3130000	PN
280005	NA	-0.1030000	PN
280011	NA	-3.1050000	PN
280014	-1.70	-4.2320000	PN
280015	NA	-1.7020000	PN
280037	NA	-0.2550000	PN
280038	NA	0.3710000	PN
280041	NA	-1.5340000	PN
280048	NA	2.3400000	PN
280050	NA	0.8630000	PN
280065	NA	0.2760000	PN
280068	NA	-1.7970000	PN
280071	NA	-0.6950000	PN
280075	NA	-1.1940000	PN
280076	NA	-1.1280000	PN
280078	NA	1.2490000	PN
280080	NA	-1.5760000	PN
280081	NA	1.3290000	PN
280083	1.70	-0.4360000	PN
280084	NA	-1.1330000	PN
280085	NA	0.0020000	PN
280086	NA	-1.0710000	PN
280087	NA	0.0000000	PN
280088	NA	-0.3920000	PN
280090	NA	-5.0540000	PN
280091	-0.07	-2.1900000	PN
280092	2.20	-0.5530000	PN
280093	NA	-0.0580000	PN
280094	NA	-0.9570000	PN

itemid	b_original	b_linked	task
280095	NA	-0.9680000	PN
280096	-0.58	-1.1630000	PN
280097	NA	-1.1950000	PN
280098	NA	4.2910000	PN
280099	NA	-1.9080000	PN
280100	NA	-1.6380000	PN
280101	NA	-1.9600000	PN
280102	NA	-2.6130000	PN
280103	NA	-1.7190000	PN
280104	NA	-0.3330000	PN
280105	-2.04	-5.0450000	PN
280107	NA	0.3870000	PN
280108	NA	-0.5900000	PN
280109	0.58	0.1060000	PN
280110	NA	-1.9450000	PN
280111	NA	-2.4420000	PN
280112	NA	-2.1120000	PN
280113	NA	-1.0260000	PN
280114	NA	-1.3610000	PN
280115	NA	0.4390000	PN
280116	NA	2.6790000	PN
280117	NA	-0.5830000	PN
280118	NA	-1.2550000	PN
280119	NA	-1.2900000	PN
280120	NA	-1.8220000	PN
280121	NA	-2.5790000	PN
280122	NA	0.8510000	PN
280123	NA	-0.1710000	PN
280124	NA	3.2660000	PN
280126	NA	-2.3990000	PN
280127	NA	-1.6120000	PN
280128	NA	0.2690000	PN
280129	NA	0.0000000	PN
280130	NA	-0.2120000	PN
280132	NA	-3.3420000	PN
280133	NA	-0.7770000	PN

itemid	b_original	b_linked	task
280134	1.20	-0.2630000	PN
280135	NA	0.8540000	PN
280136	NA	-0.0730000	PN
280137	NA	-0.9440000	PN
280138	NA	0.7510000	PN
280139	NA	0.0000000	PN
280142	0.59	-1.6120000	PN
280143	NA	-1.9530000	PN
280144	NA	0.3650000	PN
280145	NA	-1.1010000	PN
280146	NA	0.5640000	PN
280147	NA	1.0890000	PN
280148	NA	1.5100000	PN
280149	NA	-3.8380000	PN
280150	1.00	-3.2490000	PN
280151	NA	4.3040000	PN
280152	NA	-3.8380000	PN
280153	NA	-3.7810000	PN
280154	NA	-0.3370000	PN
280155	NA	2.5460000	PN
280156	NA	1.8260000	PN
280157	NA	-2.2540000	PN
280158	NA	-1.3680000	PN
280159	NA	-4.2890000	PN
280160	NA	4.8280000	PN
280161	NA	1.6900000	PN
280162	NA	3.5350000	PN
280163	NA	1.3240000	PN
280164	1.19	-1.5020000	PN
280165	NA	-1.0510000	PN
280166	NA	2.2510000	PN
280167	NA	-0.3550000	PN
280168	NA	1.2510000	PN
280169	NA	-2.0030000	PN
280170	NA	-2.3250000	PN
280171	NA	1.7610000	PN

itemid	b_original	b_linked	task
280172	NA	0.9130000	PN
280173	NA	0.0000000	PN
280174	NA	1.3420000	PN
280175	NA	0.3700000	PN
280176	NA	1.4470000	PN
280177	NA	-0.7390000	PN
280178	NA	1.3220000	PN
280179	NA	-1.8880000	PN
280180	NA	2.6820000	PN
280181	NA	0.1940000	PN
280182	NA	-0.4410000	PN
280184	NA	-1.4030000	PN
280185	NA	0.0000000	PN
280186	NA	0.0000000	PN
280187	NA	1.5840000	PN
280188	NA	-1.7970000	PN
280189	NA	1.2680000	PN
280190	NA	2.1250000	PN
280191	NA	-0.5820000	PN
280192	NA	0.0000000	PN
280193	NA	-1.3450000	PN
280194	NA	0.0000000	PN
280195	NA	-2.5030000	PN
280196	NA	-0.6360000	PN
280197	NA	-1.9980000	PN
280198	NA	-0.6870000	PN
280199	NA	-3.2310000	PN
280200	2.04	-0.0100000	PN
280201	NA	-0.5600000	PN
280202	NA	-3.9610000	PN
280203	NA	-3.2750000	PN
280204	NA	-1.8740000	PN
280205	NA	0.1810000	PN
280206	NA	-0.3280000	PN
280207	NA	2.6700000	PN
280208	NA	0.3600000	PN

itemid	b_original	b_linked	task
280209	NA	3.9400000	PN
280210	NA	0.2450000	PN
280211	NA	-2.1280000	PN
280212	NA	-0.1240000	PN
280213	NA	-0.5160000	PN
280214	NA	1.3190000	PN
280215	NA	1.3720000	PN
280216	NA	-1.8780000	PN
280217	-0.16	-2.9870000	PN
280218	NA	2.5820000	PN
280219	NA	-2.4370000	PN
280220	NA	-0.3260000	PN
280221	NA	1.6220000	PN
280222	NA	0.7300000	PN
280223	NA	-3.9170000	PN
280224	NA	2.5130000	PN
280225	0.25	-2.9870000	PN
280226	NA	-1.9890000	PN
280227	NA	0.4460000	PN
280228	NA	1.1230000	PN
280229	NA	1.2240000	PN
280230	NA	-0.0270000	PN
280231	NA	-0.2740000	PN
280232	NA	-0.7920000	PN
280233	NA	0.5460000	PN
280234	NA	-0.9080000	PN
280235	NA	0.9440000	PN
280236	NA	0.0000000	PN
280237	NA	0.0000000	PN
280238	NA	-0.5280000	PN
280239	NA	-0.4050000	PN
280240	NA	0.1130000	PN
280241	NA	0.9620000	PN
280242	NA	0.1780000	PN
280243	NA	1.8030000	PN
280244	NA	-0.8640000	PN

itemid	b_original	b_linked	task
280245	1.95	-1.3190000	PN
280246	NA	-1.3900000	PN
280247	NA	2.0010000	PN
280248	2.85	-0.1370000	PN
280249	NA	1.6550000	PN
280250	NA	-0.0540000	PN
280251	NA	0.1990000	PN
280252	NA	-2.1970000	PN
280253	NA	0.0000000	PN
280391	NA	0.6590000	PN
280262	NA	-0.6430000	PN
280269	NA	-3.3570000	PN
280275	NA	-0.1410000	PN
280294	NA	3.5870000	PN
280364	NA	3.2030000	PN
280325	NA	-1.2220000	PN
280313	NA	1.5030000	PN
280359	NA	-0.8680000	PN
280264	NA	-1.0360000	PN
280315	NA	0.6130000	PN
280324	NA	-1.3240000	PN
280409	NA	1.3250000	PN
280362	NA	2.6900000	PN
280319	NA	2.6900000	PN
280274	NA	0.4010000	PN
280407	NA	2.6010000	PN
280386	NA	-0.8690000	PN
280297	NA	-0.3410000	PN
280347	NA	1.5160000	PN
280320	NA	1.9640000	PN
280335	NA	1.7790000	PN
280358	NA	0.0490000	PN
280429	NA	0.4860000	PN
280408	NA	-1.8370000	PN
280312	NA	2.9260000	PN
280357	NA	0.4440000	PN

itemid	b_original	b_linked	task
280384	NA	1.9170000	PN
280428	NA	-0.6780000	PN
280341	NA	2.2640000	PN
280374	NA	-1.7820000	PN
280326	NA	-1.6700000	PN
280317	NA	0.5180000	PN
280344	NA	4.3290000	PN
280305	NA	0.3440000	PN
280383	NA	-0.5110000	PN
280333	1.72	-0.5110000	PN
280340	NA	3.9400000	PN
280348	NA	-1.4650000	PN
280406	NA	2.9520000	PN
280339	NA	2.2460000	PN
280337	NA	-0.7590000	PN
280293	NA	2.1520000	PN
280352	NA	-0.1290000	PN
280338	NA	1.7100000	PN
280306	NA	0.5270000	PN
280367	NA	2.4980000	PN
280365	NA	0.5270000	PN
280350	NA	3.9380000	PN
280257	2.27	0.3550000	PN
280256	NA	-1.4040000	PN
280377	NA	-1.5090000	PN
280385	NA	-0.2170000	PN
280295	NA	0.0000000	PN
280321	NA	0.3550000	PN
280398	NA	2.4980000	PN
280287	NA	2.1480000	PN
280289	NA	0.8960000	PN
280405	NA	4.0660000	PN
280265	NA	-1.4040000	PN
280303	NA	1.6480000	PN
280263	NA	-2.9250000	PN
280411	NA	2.9670000	PN

itemid	b_original	b_linked	task
280271	NA	-0.1790000	PN
280277	NA	-0.1180000	PN
280278	NA	0.0010000	PN
280351	NA	2.4250000	PN
280366	NA	0.5290000	PN
280426	NA	-1.0150000	PN
280266	NA	-1.5930000	PN
280353	NA	-1.4580000	PN
280296	2.27	0.2760000	PN
280397	NA	-0.1790000	PN
280299	2.50	0.3280000	PN
280375	NA	2.8490000	PN
280310	NA	1.6170000	PN
280380	NA	-1.1630000	PN
280260	NA	-1.5920000	PN
280307	NA	0.1410000	PN
280283	NA	2.7320000	PN
280399	NA	3.9180000	PN
280394	NA	1.0570000	PN
280323	NA	-1.0900000	PN
280330	NA	0.8120000	PN
280422	NA	1.1610000	PN
280387	NA	-1.1630000	PN
280284	NA	-0.2470000	PN
280311	NA	2.5840000	PN
280354	NA	1.5460000	PN
280316	NA	0.8120000	PN
280376	NA	2.1090000	PN
280345	NA	3.8040000	PN
280298	NA	-0.2920000	PN
280425	NA	1.8720000	PN
280318	NA	-1.0390000	PN
280308	NA	2.5560000	PN
280381	NA	-0.6570000	PN
280300	NA	-0.2920000	PN
280280	NA	-1.5310000	PN

<u>itemid</u>	<u>b_original</u>	<u>b_linked</u>	<u>task</u>
280389	NA	1.2350000	PN
280268	NA	-1.5310000	PN
280261	NA	0.8010000	PN
280363	NA	1.2690000	PN
280288	NA	3.3470000	PN
280314	NA	3.7000000	PN

Preliminary Results: Contact Authors for Update

Appendix E
Data Files Used in Analysis

Demographic data

- “idgdi3_demographics_Y4.csv”

Demographic analyses

- “Child-demographics-for-IGDI-3-Y3-analyses.html”

IGDI data export

- “ExportFile07012018-07012019.csv”

Analysis report file

- “igdi3-year3-report-201104.html”

Preliminary Results: Contact Authors for Update