Preliminary Results, Criterion Validity Analysis

Expanding Individual Growth & Development Indicators of Language and Early Literacy for

Universal Screening in Multi-Tiered Systems of Support with Three-Year-Olds

Technical Report #12

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

Wackerle-Hollman

IGDILab

Department of Educational Psychology

University of Minnesota

Updated: 10/29/21

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 Renaissance Learning, for commercial development and sale. These individuals and the University of Minnesota have a consulting relationship and/or receive royalties from 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 teachers and students who contributed their time and effort to this project, and the program leaders who made this work possible.

Preliminary Results, Criterion Validity Analysis

Expanding Individual Growth & Development Indicators of Language and Early Literacy for

Universal Screening in Multi-Tiered Systems of Support with Three-Year-Olds

Technical Report #[Index]

Abstract

State and local education agencies are calling for the expansion of evidence-based

assessment to improve the quality of early childhood education classrooms. However, few

conceptually and empirically sound tools exist. This is particularly true for young preschoolers,

or students who are more than one year away from the age cutoff for kindergarten entry. This

document presents preliminary results for validity analyses of IGDI assessments designed for 3-

year-olds (IGDI3 measures) using correlations with criterion measures (TOPEL, PPVT, GRTR,

QUILS). In this study, 136 students completed at least one criterion assessment and IGDI from a

corresponding domain. Sample characteristics are presented as well as results of correlations of

IGDIs with criterion measures.

Preliminary Results, Criterion Validity Analysis

Expanding Individual Growth & Development Indicators of Language and Early Literacy for

Universal Screening in Multi-Tiered Systems of Support with Three-Year-Olds

Technical Report #12

Project Introduction

This document was been developed as part of a four-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

is to extend the practical array of Individual Growth and Development Indicators (IGDIs) for

assessing language and early literacy development to 3-year-old preschool children in ways that

inform and enhance multi-tiered systems of support, 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 children[2] in areas of oral language, phonological awareness, and alphabet knowledge.

We assessed the psychometric characteristics of these measures (including their classification

accuracy for identifying candidates for more intensive intervention) and used these measures to

describe growth across time in PK3 and PK4 in ways that help describe the broader domain of

language and early literacy development, and that note relations between achievement in this

area and characteristics of services children receive in early childhood classrooms.

---

[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.

Universal reading success is broadly accepted as an essential goal for education of elementary students in the United States and around the world (National Institute of Child Health and Human Development, 2000). An individual's ability to access and understand text, to make sense of what is read, and to use this information to communicate with others or perform functional tasks – and to explore, wonder, and create – is a central goal of early elementary schooling, and lays the foundation for achievement in other domains and for overall educational attainment.

For decades, research has demonstrated that specific skill development in the preschool years (generally, the time before a child enrolls in Kindergarten, typically at age 5) is directly and causally related to successful acquisition of early reading skills in elementary school (National Early Literacy Panel, 2008). This growing empirical base of knowledge has led to related increase in the attention to, development of, and funding for an array of early childhood education programs specifically intended to promote language and early literacy development.

Multi-tiered systems of support and other early intervention frameworks have been created within this broader context of early childhood education programs promoting language and early literacy development (Carta & Young, 2019). Multi-tiered systems of support include at least three common components, including a) universal screening to identify children for whom supplemental or more intensive intervention may be warranted; b) an array of evidence-based interventions implemented in the child's natural educational environment (from universal, or Tier 1, efforts to supplemental, or Tier 2, interventions to individualized intensive, or Tier 3, instruction); and c) procedures for gathering and evaluating progress monitoring data both to evaluate whether the indicated level of intervention is promoting expected levels of growth and, when appropriate, to signal the appropriateness of moving an individual to a different level of

intervention service (Carta, 2019; Hojnoski & Polignano, 2019).

Individual Growth and Development Indicators (IGDIs) were developed specifically to address universal screening and progress monitoring requirements for MTSS programs to promote language and early literacy development (McConnell et al., 2015). IGDIs are examples of General Outcome Measures (Fuchs & Deno, 1991) designed and evaluated specifically for use in preschool programs. Like other General Outcome Measures, IGDIs are brief, easy to administer, can be repeated over relatively short intervals to produce estimates of growth, related to long-term meaningful outcomes, and produce results that are easily interpreted and actionable by most classroom teachers.

Over more than 20 years of research and development (c.f., Priest et al., 2001), IGDIs for preschool children have developed and evolved. Earliest versions of these measures were developed using classical test theory models and with a broad array of administration options; assessment was conducted with a randomly selected set of physical cards, each assessment was timed for 1 or 2 minutes, and student performance was reported as the number of correct responses per minute (or 2 minutes). While the assessments were generally well-accepted by teachers and programs and produced positive and intriguing results (McConnell & Missall, 2008; Missall et al., 2006; Missall & McConnell, 2010; Missall et al., 2008), significant shortcomings in psychometric utility and logistical feasibility prompted significant revision.

Currently available IGDIs are no longer timed, but rather have been developed using item response theory for a comparable, but logistically different, approach to placing current child performance on a growth trajectory. As a consequence of both this transition to modern test theory and rapid development of portable digital technology, these measures are also now available in a computer-adaptive format, with administration conducted using two wirelessly

connected iPads. Administration is still brief, child-engaging, and psychometrically rigorous (see McConnell, et al., 2015, for more detail), and has the added advantage of affording better access and more useful information much more quickly to classroom teachers (Will et al., 2019).

While these innovations have expanded utility and improved quality of data provided by IGDIs, to date the measures have been developed for and evaluated only with children at least 4 years of age, or specifically those children enrolled in preschool in the year before they become age-eligible for kindergarten. Given continued advances in research support for the role of language and early literacy development even among younger preschool students, and growing provision of early education services to children from age 3 (i.e., two years prior to being age-eligible for kindergarten), research recently turned to adapting and expanding available IGDIs to efficiently and rigorously assess language and early literacy development for three-year-olds.

To date, there are very few research-based measures, and even fewer practice-ready measures, for assessing language and early literacy and support MTSS for these younger students. To address this need, one extension of an existing measure of oral language and new measures of alphabet knowledge and phonological awareness were developed to complement and align with existing IGDIs. The purpose of this paper is to report results of evaluation of relations both among these three new measures, and between these three measures and other measures of language and early literacy to provide initial evidence of concurrent validity.

Theoretically, IGDIs and other General Outcome Measures are indicators of broad, often multidimensional skill sets. Deno (1997) describes essential features, and primary assets, of this approach focusing both on practical relations (particularly utility for intervention planning and progress monitoring). In this logic, validity claims are focused on Cronbach and Meehl's (1955) notion of *relational fertility*, or the degree to which any measure "support[s] the adequacy and

appropriateness of interpretations and actions on the basis of test scores or other modes of

assessment" (Messick, 1995, p 741), a characteristic sometimes described as a nomothetic net

(Deno, 1997).

Individual Growth and Development Indicators for preschool children, like other GOMs

(e.g., words read correct in grade-level passages), appear to have domain-specific characteristics

(e.g., "oral language" or "phonological awareness") largely because preschool children cannot

yet perform the comprehensive, and targeted skill of reading. Thus, while one goal of assessment

and intervention planning might be to describe a child's acquisition of a multi-dimensional or

complex skill like reading, measures of known components can be evaluated by examining their

relative relation to other measures of known characteristics. Kane (2013) argues that these types

of evaluations can be facilitated by test developers (and others) by an argument-based approach,

where inferences regarding test interpretation or use are articulated and then tested empirically.

Purpose of this Report

This document presents preliminary results for validity analysis of IGDI3 measures by

addressing the following research questions:

a. To what extent do IGDI scale scores relate to standardized measures of oral language and

preschool early literacy development in domain-specific ways that support established

validity claims?

b. To what degree do IGDI scale score correlate within domain to demonstrate evidence to

support convergent validity?

Methods

Participants

Study participants were recruited as part of a larger research and development project that assessed language and early literacy development among PK3 students in classrooms located in the upper Midwest and eastern United States. From the existing pool of 477 participants in the larger study, we sought to sample at least 120 students for the present study.

Due to geographic and logistical limitations of our research team, students were quasi-randomly assigned to one of four criterion measures groups, each consisting of different combinations of criterion measures (see procedures for group descriptions). We selected 26 classroom sections from a total of 29 across three participating school districts, resulting in a pool of 346 students. One classroom was excluded because it had less than six participating students and the remaining two were excluded to reduce the burden on the teacher if they had multiple sections participating in the study. Next, a subsample of students from each section was randomly assigned to one of four groups, depending on class size, using the RAND function in Microsoft Excel (2003). In order to achieve complete data on 120 students, we oversampled, resulting in a total of 136 children completing at least one criterion measure.

Participants ranged in age from 44.2 to 56.2 months (M = 51.5, SD = 3.5) at the time of the administration of spring IGDIs. Demographic variables are summarized in Table 1 for this sample of children, with counts for gender, race, home language, IEP status and free and reduced lunch status.

Measures

IGDI Measures

Three IGDI measures were administered to all participating children. All measures were

administered using two digital devices in a student-assessor dyad and a computer-adaptive

algorithm (CAT) for selection of specific items (Wang et al. 2017). Students responded to 15

items for each measure.

Alphabet Knowledge. We tested four measures of alphabet knowledge. *Letter*

Orientation (LO) is a receptive task. The child was presented an array of three instances of a

single letter; only one of the three instances was oriented appropriately (see item samples,

Appendix). The examiner told the child "These are all letters. Point to one that's the right way."

The child indicated a correct response by pointing or touching.

Letter Find (LF) is also a receptive task. The child was presented an array of one letter

and two non-letter shapes. The examiner stated, "Point to the letter." The child indicated a

correct response by pointing or touching.

Point to Letter (P2L) is also a receptive task. The child was presented with an array of

three upper- and/or lower-case letters. The examiner stated, "Point to the [names one of the

letters]." The child indicated a correct response by pointing or touching.

Letter Naming (LN) is an expressive task. The child was presented with a single upper- or

lower-case letter and asked "what is this this letter called?" The child responded orally, and the

examiner records correct or incorrect response.

Oral Language. The oral language measure combined tasks that required the child to

make either a receptive (Point to Picture) or expressive response (Picture Naming). In the *Point*

to Picture (P2P) task, the child was presented with an array of two or three photos of non-

ambiguous objects known to typically developing 3-year-old preschoolers. The examiner stated,

"Point to the [name of one object]." The child indicated a correct response by pointing or

touching. The expressive counterpart of the P2P task is *Picture Naming* (PN). In this task, the

child was presented with a single image and asked what that image is a picture of.

Phonological Awareness. *Robot Blending* (RB) was developed to assess preschool

children's ability to hear stretched or elongated words, and to receptively identify the

pronounced word. The child first "meets" Igdi the Robot, who "talks funny." Multiple sample

items were used to illustrate this slow, stretched speech. Then for each item, the child was

presented with an array of two or three images, these images were labeled by the examiner, and

then a standardized recording of one of the object name played a segmented robot voice at the

word level (for compound words), syllable level, or phoneme level. Each segment was spoken

somewhat slowly, and a standard .75 sec interval of silence was placed between each segment.

After hearing the word, the child was asked to point to or touch the displayed object that they

thinks was named.

Criterion Measures

Test of Preschool Early Literacy (TOPEL). The TOPEL (Lonigan et al., 2007) is a

norm-referenced, individually administered paper-and-pencil assessment of emergent literacy. It

is designed for children age 36-71 months and measures print knowledge, definitional

vocabulary, and phonological awareness. It has 98 items and takes approximately 25-30 minutes

to administer. Student performance is reported in raw scores, standard scores, sum of standard

scores, and overall percentile rank. Reported test-retest reliability ranged from .48 to .81 (Wilson

& Lonigan, 2009). Reported validity evidence of convergent validity correlations with the

GRTR-R ranged from .65-.73 for TOPEL print knowledge, .43-.48 for TOPEL definitional

vocabulary, .43-.51 for TOPEL phonological awareness, and .71-.72 for TOPEL early literacy

index. Convergent validity correlations with original IGDIs 1.0 measures (McConnell, 2002) for

4-year-olds were also reported by Wilson and Lonigan (2009), but are not reported here since

updated IGDIs measures exist.

Peabody Picture Vocabulary Test, Fourth Edition (PPVT-IV). The PPVT (Dunn & Dunn, 2007) is a norm-referenced, individually administered assessment of vocabulary. It is designed for individuals 30 months of age and older and measures receptive vocabulary and oral language development. In the current study, the paper-and-pencil version of PPVT-IV was used. The administration time varies because individuals continue to respond to items until basal and ceiling rules are met. Student performance is reported as a raw score and standard score. Reported reliability evidence included.87-.93 alternate form reliability, .87-.93 test-retest reliability, .89-.97 internal consistency on age-norm samples, and .87-.97 internal consistency on grade-norm samples (Dunn & Dunn, 2013). Reported validity evidence included test correlations of .67-.75 with the Clinical Evaluation of Language Fundamentals (CELF; Semel, et al., 2003) and .80-.84 with the Expressive Vocabulary Test (EVT-2; Williams, 2007).

Quick Interactive Language Screener (QUILS). QUILS (Golinkoff et al., 2017a) is a screening measure individually administered with a touchscreen tablet device via the QUILS website. Designed for children age 36-71 months, it has 48 items measuring vocabulary, syntax, and processing. It takes approximately 15-20 minutes to administer. Responses are automatically scored. Student performance is reported in raw scores, a single scaled score, standard scores, and percentile rank. Reported convergent validity evidence indicated correlations with the PPVT-4 standard score (Dunn & Dunn, 2007) from .54-.67 across vocabulary, syntax, process, and overall standard score (Golinkoff et al., 2017b). Correlations with the Preschool Language Scales (PLS-5; Zimmerman et al., 2011) Auditory Comprehension standard score ranged from .54-.65 across vocabulary, syntax, process, and overall standard score. Reported test-retest reliability was .71 for vocabulary, .73 for syntax, .69 for process and .83 overall (Golinkoff et al., 2017b).

Internal consistency reliability was .79 for vocabulary and syntax, .87 for process, and .93 overall.

Get Ready to Read, Revised (GRTR-R). The GRTR-R (Lonigan & Wilson, 2008) is a screening measure for early literacy used with children ages 36-60 months. It has 25 items measuring print knowledge, linguistic awareness, and emergent writing. The GRTR is administered individually to each child. Items are presented on cardstock, and responses are recorded by the assessor on a paper score form. It takes approximately 10-15 minutes to administer and results in a single score of the total items correct. The reported overall internal consistency was α =.88 (Lonigan & Wilson, 2008). Broken down by age, internal consistency was .77 for 3-year-olds, .84 for 4-year-olds, and .83 for 5-year-olds. Reported validity evidence compared GRTR-R to the TOPEL (Lonigan et al., 2007) and found correlations with the TOPEL Print Knowledge ($r = .76$, $p < .001$) and overall Early Literacy Index ($r = .72$, $p < .001$). Correlations with the TOPEL Phonological Awareness ($r = .39$, $p < .001$) and Definitional Vocabulary ($r = .44$, $p < .001$) were lower in comparison.

Procedures

As part of the larger investigation, all participating children completed seasonal assessments with all three IGDIs. Tests were administered by trained members of our research team, in or near each student's classroom.

Administration of criterion measures began in January and continued until completion in May. In order to obtain samples for within- and across- domain analysis, students were assigned to one of the following groups: (a) TOPEL and QUILS *(n = 30)*; (b) TOPEL and GRTR-R *(n = 30)*; (c) TOPEL, QUILS, and PPVT *(n = 30)*; and (d) TOPEL, GRTR-R, and PPVT *(n = 30)*. This sampling structure resulted in all students completing the TOPEL (*n* = 120), and 60 students

completing assessments at least one of the PPVT, GRTR-R, and QUILS assessments. However, due to data collection scheduling issues (e.g., snow storms, child attendance) half way through data collection we modified group assignment to ensure we had adequate number of students completing each criterion measure, resulting in 136 students completing at least one criterion measure (see Table 3 for ns for individual criterion measures).

Members of the research team completed training in assessment administration, human subjects protection, and confidentiality protection as well as demonstrated fidelity with test administration requirements prior to data collection. Data collectors attended a 1.5-hour workshop facilitated by graduate research assistants. During the workshop, each criterion measure was described, and data collectors practiced administering the assessment and reviewed administration logistics. Data collectors worked with children during their regular school day and assessed students in the hallway or an empty space near their classrooms. Test materials were set up on tables if available, otherwise assessments were completed on the floor.

Data collectors administered each criterion measure following the standardized instructions specific to that assessment (i.e., administration instructions and basal and ceiling scoring rules). The majority of assessments were completed in two sittings, with some students needing additional sessions to complete their assigned measures. Data collectors were instructed to attempt to complete each measure within a sitting, with the exception of the TOPEL which was administered across two sittings, with the first two sub-tests (Print Knowledge and Definitional Vocabulary) administered in one sitting and Phonological Analysis in the second, along with any other assigned assessment. At the end of each session children received a sticker or small toy.

The analytic data set for this study included student demographic information, results

from one or more criterion measures, and Spring IGDI results for all participants.

Results

Descriptive Statistics

Descriptive statistics for IGDIs completed by participating children are reported in Table 2. While reported in theta values centered at an arbitrary zero, means and standard deviations across all three measures indicate student performance that is slightly above that arbitrary center and with normal distributions.

Descriptive statistics for all four criterion measures are reported in Table 3. Three of these measures (i.e., TOPEL, PPVT-V and QUILS) are norm-referenced by child age; in addition to normative scale scores we retained raw scores to represent variance due to child performance across age-based norm groups for these three measures. Review of standard score results for these three measures demonstrates all means slightly above, but within one-third standard deviation, of the normative mean. SD of scale scores ranged from 11.51 to 16.4, again approximating expected standard score SDs of 15.

Correlations

To answer the first research question, "To what extent do IGDI scale scores relate to standardized measures of oral language and preschool early literacy development in domain-specific ways that support our validity claims?, we computed Pearson's correlations to examine the relation between IGDI and criterion measures. We expected correlations to be significant between the oral language Three-year-old IGDIs and the PPVT, QUILS vocabulary, syntax and composite scores, and TOPEL DV and sum scores; between the phonological awareness IGDIs and the TOPEL PA, sum, and GRTR scores; and between the alphabet knowledge IGDIs and the TOPEL PK, sum, and GRTR scores. For this analysis, missing data were removed using pairwise

deletion. As a result, the samples used to estimate each correlation differed by pair of variables.

Results are presented in Table 4, with expected within domain correlations bolded. Sample sizes

are shown in parentheses.

Finally, to answer the third research question, to what degree do IGDI scale score

correlate within domain to demonstrate evidence to support convergent validity?, we computed

Pearson's correlations among the IGDI measures. Results are presented in Table 5.

Discussion

Conceptually and empirically sound assessments are crucial for multi-tiered systems of

support to be effective in promoting academic success. As states and local education agencies

expand early childhood education programs to serve younger preschool children, we must

increase the availability of tools designed for this age group and evaluate against psychometric

standards. As such, this study provides preliminary evidence that IGDI3 measures moderately to

strongly correlate with domain-matched criterion assessments.

References

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.

Carta, J. J., & Young, R.M. (Eds.). (2019). *Multi-tiered systems of support for young children:*

Driving change in early education. Brookes Publishing.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological*

Bulletin, 52, 281–302.

Deno, S. L. (1997). Whether thou goest. . . Perspectives on progress monitoring. In J. W. Lloyd,

E. J. Kame'enui, & D. Chard (Eds.), *Issues in educating students with disabilities* (pp.

77–99). Erlbaum.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test (PPVT-IV)*. New York:

Pearson.

Dunn, L. M., & Dunn, D. M. (2013). *Peabody Picture Vocabulary Test (PPVT-4)* Technical

Report.

https://www.pearsonassessments.com/content/dam/school/global/clinical/us/assets/ppvt-

4/ppvt-4-technical-report.pdf

Fuchs, L. S., & Deno, S. L. (1991). Paradigmatic distinctions between instructionally relevant

measurement models. *Exceptional Children, 57*, 488-500.

Golinkoff, R. M., de Villiers, J., Hirsh-Pasek, K., Iglesias, A., & Wilson, M. S. (2017a). *Quick*

Interactive Language Screener (QUILS^{TM}): A measure of vocabulary, syntax, and

language acquisition skills in young children. Baltimore, MD: Paul H. Brookes.

Golinkoff, R. M., de Villiers, J., Hirsh-Pasek, K., Iglesias, A., & Wilson, M. S. (2017b). User's

manual for the Quick Interactive Language Screener *(QUILS^TM)*.

Hojnoski, R. L., & Polignano, J. C. (2019). Using data-based decision making to improve

learning outcomes for all children. In J. J. Carta, & R. M. Young (Eds.), *Multi-tiered

systems of support for young children* (pp. 97-111). Brookes Publishing.

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational

Measurement, 50*(1), 1-73.

Lonigan, C. J., Wagner, R. K., & Torgeson, J. K. (2007). *Test of Preschool Early Literacy

(TOPEL)*. Austin, TX: Pro-Ed.

Lonigan, C. J., & Wilson, S. B. (2008). Report on the Revised Get Ready to Read! screening

tool: Psychometrics and normative information (Technical Report). National Center for

Learning Disabilities.

http://www.getreadytoread.org/images/content/downloads/GRTR_screening_tool/grtrnor

mingreportfinal-july-2008.pdf

McConnell, S. R. (2002). Individual Growth and Development Indicators. University of

Minnesota.

McConnell, S. R., & Missall, K. N. (2008). Best practices in monitoring progress for preschool

children. In A. Thomas & G. Grimes (Eds.), *Best practices in school psychology*. 5th

Edition, National Association of School Psychologists, pp. 561-573.

Missall, K. N., Carta, J. J., McConnell, S. R., Walker, D., & Greenwood, C. R. (2008). Using

Individual Growth and Development Indicators to measure early language and literacy.

Infants & Young Children, 21(3), 241-253.

Missall, K. N., & McConnell, S. (2010). Early literacy and language IGDIs for preschool children. In J. J. Carta, C. R. Greenwood, D. Walker & J. Buzhardt (Eds.), *Using IGDIs: Monitoring progress and improving intervention results for infants and young children* (pp. 181–201). Brookes.

McConnell, S. R., Wackerle-Hollman, A. K., Roloff, T. A., & 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

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from person's responses and performances as scientific inquiry into score meaning. *Methodological Issues and Strategies in Clinical Research, 2*, 241-260.

Missall, K. N., McConnell, S. R., & Cadigan, K. (2006). Early literacy development: Skill growth and relations between classroom variables for preschool children. *Journal of Early Intervention, 29*(1), 1-21.

National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel.* National Institute for Literacy.

National Institute of Child Health and Human Development. (2000). Report of the National Reading Panel: Teaching children to read. U.S. Government Printing Office. https://www.nichd.nih.gov/sites/default/files/publications/pubs/nrp/Documents/report.pdf

Priest, J. S., McConnell, S. R., Walker, D., Carta, J. J., Kaminski, R. A., McEvoy, M. A., Good, R., Greenwood, C. R., & Shinn, M. R. (2001). General growth outcomes for young children: Developing a foundation for continuous progress measurement. *Journal of Early Intervention, 24*, 163–180.

Semel, E., Wiig, E. H., & Secord, W. A. (2003). Clinical evaluation of language fundamentals, fourth edition (CELF-4). Toronto, Canada: The Psychological Corporation/A Harcourt Assessment Company.

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.

Whitehurst, G. J., & Lonigan, C. J. (2010). *Get Ready to Read! Revised*. San Antonio, TX: Pearson.

Will, K. K., McConnell, S. R., Elmquist, M., Lease, E. M., & Wackerle-Hollman. (2019). Meeting in the middle: Future directions for researchers to support educators' assessment literacy and data-based decision making. *Frontiers in Education, 4*.

Williams, K. T. (2007). *Expressive Vocabulary Test* (2nd edition; EVT-2). Pearson.

Wilson, S. B., & Lonigan, C. J. (2009). An evaluation of two emergent literacy screening tools for preschool children. *Annals of Dyslexia, 59*(2), 115-131. doi:10.1007/s11881-009-0026-9

Zimmerman, I. L, Steiner, V. G., & Pond, R. E. (2011). *Preschool Language Scales*, Fifth Edition (PLS-5). Pearson Education.

Table 1

Sample Characteristics for Demographic Variables

Demographic	*n* (%)
Gender	
Female	70 (51%)
Male	57 (42%)
Missing	9 (7%)
Race	
White	75 (55%)
Black	20 (15%)
Asian/Pacific Islander	10 (7%)
Latino	4 (3%)
American Indian	1 (1%)
Multiple*	14 (10%)
Missing	12 (9%)
Home Language	
English	111 (82%)
Other*	14 (10%)
Missing	11 (8%)
Free/Reduced Lunch	
Yes	21 (15%)
No	96 (71%)
Missing	19 (14%)
IEP	
None	106 (78%)
Current	21 (15%)
Missing	9 (7%)

Note. Total sample consisted of 136 participants.

*We recognize the limitations in combining multiple race and home language categories, for an additional breakdown of these categories, please email the study authors.

Table 2

Individual Growth and Development Indicator Means, Standard Deviations, Skew, and Kurtosis (Spring Screening)

Individual Growth and Development Indicator	*M*	*SD*	Skew	Kurtosis	*N*
Alphabet Knowledge	0.65	1.24	-0.14	2.41	136
Oral Language	0.62	0.79	-0.42	3.02	136
Phonological Awareness	0.18	1	0.43	2.28	136

Note. Table 2 presents results for 136 of 140 participating students who completed at least one of the four criterion measures. Scores for all three Individual Growth and Development Indicators are reported in Rasch theta values.

Table 3

Descriptive Statistics for Standardized Norm-referenced Tests

Criterion Measure	*M*	*SD*	Skew	Kurtosis	*N*
Test of Preschool Early Literacy					
Print Knowledge					
Raw Score	16.25	10.11	0.24	1.88	111
Standard Score	103.86	16.4	0.35	2.13	111
Definitional Vocabulary					
Raw Score	45.15	10.53	-0.87	3.51	110
Standard Score	100.75	11.51	-0.76	3.36	110
Phonological Analysis					
Raw Score	13.44	5.44	0.09	2.54	107
Standard Score	96.13	14.55	0.11	2.61	107
Sum Score	301.04	31.59	0.21	2.48	107
Peabody Picture Vocabulary Test-IV					
Raw Score	71.57	18.45	-0.25	2.68	54
Standard Score	104.76	14.9	-0.09	2.27	54
Get Ready to Read – Revised					
Raw Score	14.41	4.89	-0.11	2.56	54
Quick Interactive Language Screener					
Vocabulary					
Raw Score	10.02	3.37	-0.33	1.99	51
Standard Score	103.63	13.13	-0.19	2.52	51
Syntax					
Raw Score	9.57	3.37	-0.31	2.29	51
Standard Score	104.65	13.05	-0.28	2.56	51
Process					
Raw Score	8.82	3.7	-0.31	2.07	51
Standard Score	103.41	10.57	0.07	2.54	51
Total Score	28.41	9.28	-0.38	2.13	51

Table 4

Correlations Between IGDI and Criterion Measures

| Criterion Measures | Individual Growth and Development Indicators | | |
	Alphabet Knowledge	Oral Language	Phonological Awareness
TOPEL PK Raw	**.80*** (109)**	.20* (109)	.31*** (109)
TOPEL PK Standard Score	**.74*** (109)**	.17 (109)	.23** (109)
TOPEL DV Raw Score	.24** (108)	**.46*** (108)**	.29** (108)
TOPEL DV Standard Score	.18 (108)	**.44*** (108)**	.23* (108)
TOPEL PA Raw Score	.42*** (105)	.51*** (105)	**.52*** (105)**
TOPEL PA Standard Score	.39*** (105)	.50*** (105)	**.48*** (105)**
TOPEL Sum Score	.63*** (105)	.52*** (105)	.43*** (105)
GRTR Raw Score	.52*** (53)	.15 (53)	.50*** (53)
QUILS Vocab Raw Score	.38** (50)	**.55*** (50)**	.44*** (50)
QUILS Vocab Standard Score	.21 (50)	**.45*** (50)**	.33* (50)
QUILS Syn Raw Score	.35* (50)	**.44*** (50)**	.52*** (50)
QUILS Syn Standard Score	.29* (50)	**.41** (50)**	.41** (50)
QUILS Process Raw Score	.39** (50)	**.41** (50)**	.63*** (50)
QUILS Process Standard Score	.23 (50)	**.33* (50)**	.47*** (50)
QUILS Total Raw Score	.42** (50)	**.52*** (50)**	.60*** (50)
PPVT Raw Score	.09 (53)	**.72*** (53)**	.51*** (53)
PPVT Standard Score	.09 (53)	**.69*** (53)**	.41** (53)

Table 5

Correlations Among IGDI Measures

	Oral Language	Phonological Awareness
Alphabet Knowledge	.41 (410)	.35 (405)
Oral Language	-	.53 (405)