RENAISSANCE®

myIGDIs Early Literacy Technical Manual

myIGDIs by Renaissance

Renaissance
PO Box 8036
Wisconsin Rapids, WI 54495-8036
Telephone: (800) 338-4204
(715) 424-3636

Outside the US: 1.715.424.3636
Fax: (715) 424-4242
Email (general questions): answers@renaissance.com
Email (technical questions): support@renaissance.com
Email (international support): worldsupport@renaissance.com
Website: www.renaissance.com

Copyright Notice

6/2020 myIGDIs

Contents

Introduction

What Is myIGDIs?

Individual Growth & Development Indicators (IGDIs) are preschool assessments for monitoring the growth and development of children on the pathway to kindergarten. Scientifically validated for identifying children who are experiencing difficulties acquiring fundamental skills necessary for academic success, IGDIs can also be used to measure developmental gains and inform instructional needs of individual children.

Our early childhood assessments are specifically designed to give preschool educators the decision-making tools they need in order to monitor growth throughout a child's early years. For over a decade, IGDIs early childhood assessments have enabled early childhood and pre-K instructors to identify children at-risk for developmental delays early and monitor development gains often to help children become school-ready.

myIGDIs includes these measures:

- Picture Naming
- Rhyming
- Sound Identification
- Which One Doesn't Belong
- Alliteration

Background

In 1998, "Individual Growth and Development Indicators of Early Literacy" were developed to monitor early literacy development of preschool-aged children. More recently, these tools have been extensively redesigned, developed, and evaluated to address some of the technical issues of the first edition. The outcome: *myIGDIs Early Literacy+*.

myIGDIs Early Literacy+ includes measures of early literacy development that have been designed under the auspices of the University of Minnesota and Center for Response to Intervention in Early Childhood to support the identification of students requiring additional levels of intervention in the key early literacy domains of oral language, phonological awareness, alphabet knowledge, and comprehension. Most significantly, the new measures were constructed with Item Response Theory and are closely aligned with important instructional decisions that need to be made.

As part of a larger model of RTI in early childhood programs, myIGDIs Early Literacy+ was developed to inform decisions about whether children are demonstrating adequate levels of performance given the general level of instruction ("Tier 1"), or if their performance indicates a need for more intense levels of instruction ("Tier 2" or "Tier 3").

Note: These assessment measures are the second edition of the "Individual Growth and Development Indicators" (IGDIs). The original IGDIs, sometimes referred to as "Get It, Got It, Go!," were first developed in 1998. The measures described here were completed in 2012 and are intended for identification/screening purposes.

Special Recognition of the myIGDIs Early Literacy+ Authors

The Individual Growth & Development Indicators of Early Literacy were originally developed by Drs. Scott McConnell, Tracy Bradfield, Alisha Wackerle-Hollman and Michael Rodriguez of the Center for Response to Intervention in Early Childhood at the University of Minnesota under grant funds from the Institute of Education Sciences. For more about the authors, see Appendix A, About the myIGDIs Authors.

Administration

For administration details, refer to the *myIGDI's by Renaissance Test Administration Manual* at https://www.myigdis.com/wp-content/uploads/2019/09/myIGDIs_Literacy_Screening_Administration_Manual.pdf.

Development

Response to Intervention (RTI) and other multi-tiered systems of support have several common features, regardless of their populations, domain(s) of concern, implementation settings, and purposes (Frieden, 2010; U.S. Public Health Service, 1994). Principal among these common characteristics is a measurement framework that supports the identification of individuals who require additional intervention and that assists in monitoring effects of intervention variation for selected individuals. Like other multi-tiered systems, RTI models most often include *universal screening* of all enrolled students to identify appropriate candidates for more intensive or differentiated intervention (i.e., Tier 2 or 3 services), and *progress monitoring* of these individuals' achievement as tiered intervention is provided (Chard et al., 2008; Kratochwill, Clements, & Kalymon, 2007; McConnell & Greenwood, 2013).

Frequently, General Outcome Measures (GOMs) provide the measurement infrastructure for RTI models in educational settings (Fuchs & Fuchs, 2007, 2010). GOMs are brief, easy-to-collect, and psychometrically rigorous indices of academic or related achievement that describe both current levels of performance and rates of progress over time (Fuchs & Deno, 1991). In RTI, GOMs have emerged as a mature measurement approach, providing both individual and group-level data, and informing teachers and others about child achievement at a single point in time (e.g., universal screening) as well as assessing changes in child achievement over time during varying conditions of intervention (e.g., progress monitoring; Fuchs & Fuchs, 2007).

This GOM approach has been extended to early childhood programs in the form of Individual Growth and Development Indicators (IGDIs). Initially, IGDIs were created to measure developmental achievement in multiple domains as part of a larger effort to establish program performance measures in early childhood special education (Priest et al., 2001). Noteworthy for the discussion here, research has developed IGDIs of early literacy and language for infants and toddlers (Early Communication Indicators, or ECIs; Greenwood, Walker, & Buzhardt, 2010) through early elementary grades (Dynamic Indicators of Basic Early Literacy Skills, or DIBELS; Good, Gruba, & Kaminski, 2002).

The original preschool IGDIs resulting from this work—measures of oral language development (*Picture Naming*) and phonological awareness (*Rhyming and Alliteration*)—were implemented in Early Reading First and other language and literacy intervention efforts (McConnell, Priest, Davis, & McEvoy, 2002). Although these preschool early literacy IGDIs were broadly disseminated and used, limitations of the "first-generation" IGDIs—particularly for functions of assessment relevant to RTI—became apparent (McConnell &

Missall, 2008). In particular, first-generation early literacy IGDIs proved to be too difficult for many low-performing or younger students, yielding frequent zero scores when conducting assessments (Roseth, Missall, & McConnell, 2012); for RTI models, where the emphasis is on identifying and distinguishing levels of performance among lower-performing children, this was particularly problematic. In addition, due to the somewhat unstructured approach to selecting items for any single administration, difficulty of the item sets varied across multiple administrations, producing both relatively large standard errors of measurement and undesirable variation in children's assessed performance as a function of the IGDI (McConnell & Missall, 2008). In RTI models, where progress monitoring is centrally important, such variation greatly limits the utility of any measure.

To address these shortcomings, and to more closely focus on the domains of language and early literacy skill, investigators associated with the Center for Response to Intervention in Early Childhood (CRtIEC) initiated a program of research and development to revise or replace original early literacy and language IGDIs. This research and development process was designed specifically to complement and align with other intervention and implementation work planned for the Center (see Greenwood et al., 2015). In particular, measurement research proceeded in five phases, depicted in Figure 1.

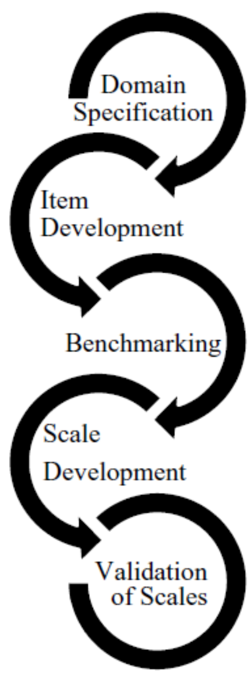

Figure 1: Graphic representation of individual growth and development indicator revision research and development process.

First, investigators associated with both assessment and intervention development reviewed both primary research and reviews of language and early literacy development (e.g., National Early Literacy Panel, 2009; Snow, Burns, & Griffin, 1998) to craft working definitions of four domains of interest (i.e., oral language, phonological awareness, alphabet knowledge, and comprehension); these definitions served as a foundation for measurement and intervention initial design.

Second, after pilot efforts to test multiple formats for acquiring and scoring child performance, the research team developed sets of items and carefully reviewed the characteristics of these items within an Item Response Theory (IRT) framework (Rodriguez, 2010; Wilson, 2005). In particular, University of Minnesota researchers developed items that conceptually represented varying ability within domains of interest, and then tested empirically the extent to which these items contributed to unidimensional descriptions of child performance. This work led to more than 150 items for each domain of interest, with item-level metrics of fit for each.

Third, IGDI developers calibrated and determined Rasch locations for each item and, provisionally, set seasonal benchmarks for identifying children performing at expected levels of achievement within each domain versus those who warranted additional intervention to achieve expected levels of proficiency (Bradfield, Besner, et al., 2014; Wackerle-Hollman, Schmitt, Bradfield, Rodriguez, & McConnell, 2013). These benchmarks were based on criterion evaluations derived from teacher judgments and student performance on standardized tests.

Fourth, these benchmarks were used to identify 15 items for seasonal screening scales. Items for these scales were selected to increase measurement focus near the benchmark to increase precision of classification decisions. Practically, this required selecting items for each scale that had locations closely surrounding the benchmark, both higher and lower on the ability scale. A similar logic was used to construct progress-monitoring scales; here, however, items were selected based on locations at or below the seasonal benchmark.

Fifth, the researchers evaluated classification accuracy of seasonal screening scales, both by replicating assessment and teacher judgment with new samples of children and teachers and in comparison with standardized criterion measures.

The initial intent of this research and development effort was to produce, in a relatively short time, fully developed and empirically supported screening and progress-monitoring measures (IGDIs) for early language and literacy. As this process unfolded, University of Minnesota researchers discovered unexpected challenges and modest successes against their original objectives. The next sections describe key assumptions that have governed this work, the results of research to date, described through the five phases, and insights into the challenges and opportunities that University of Minnesota researchers' efforts have revealed.

IGDIs in Early Childhood RTI: Key Assumptions

Application of new IGDIs (IGDIs 2.0) as the measurement infrastructure for an early childhood model of RTI has been in many ways a straightforward adaptation of GOM applications in RTI for elementary- and secondary-grade children. Scales have been developed for fall, winter, and spring universal screening; benchmarks have been identified for these screening assessments to identify children who might be candidates for more intensive intervention; and progress monitoring measures have been developed to assess children's progress (and identify changes in need for intervention) as tiered services are provided (Greenwood, Bradfield, et al., 2011; McConnell, Bradfield, & Wackerle-Hollman, 2014). At the same time, differences in the history and core characteristics of early childhood education, and an evolving policy and practice environment, have all required adaptation or change from a "traditional" RTI model seen in later grades.

At its core, logic for design and evaluation of IGDIs 2.0 is similar to that found in other RTI assessment systems, and is grounded in existing theory and empirical research (cf. Fuchs & Fuchs, 2007; National Early Literacy Panel, 2009). University of Minnesota researchers identified key aspects of language and early literacy skill *a priori*, based both on scholarship and planned focus of intervention. The researchers developed and tested items that measured these aspects developmentally and conducted research to locate individual items along an achievement continuum. As with other GOMs, University of Minnesota researchers made every effort to maintain focus on measurement growth toward a broad outcome—being a successful reader in early elementary school. Although the measures may seem to represent narrow skill sets of subdomains, the intent of University of Minnesota researchers in measure development and application is to create broad indicators of overall and ongoing development.

These foundational elements lead to three key assumptions driving this approach to early childhood RTI assessment. First, users of these measures assume that children's language and literacy skills develop from experience—both informal interaction with spoken and written text and formal intentional intervention. As a result, University of Minnesota researchers assume that purposeful intervention, appropriately directed, can prevent or remediate deficits or delays in the development of these skills.

Second, this use assumes the skills and competencies develop over time and, as a result, need for intervention can be operationalized as a difference between a child's observed performance at a single point in time and level of skill needed at that time to achieve important long-term goals. In essence, grade- or age-based proficiency standards (e.g., "reading proficiently by third grade") can be considered as terminal points on a developmental trajectory, and status on that trajectory at any earlier point can be used to infer the likelihood of later success.

Third, one can assume that successful intervention should in turn lead to accelerated development or growth on the assessed trajectory. This assumption sets the ground for using IGDIs and other GOMs to assess the efficacy of practice, and to cue changes in intervention services to produce better, more desired long-term outcomes.

While empirical evaluations of this measurement model are ongoing, several foundational elements are already in place that likely will affect IGDIs 2.0's long-term utility in early childhood RTI. These features include the validity framework on which IGDIs 2.0 and any associated interventions are based, the care and attention to assessment development, and the features and quality of decisions made regarding individuals' need for supplemental or tiered intervention.

Using a Validity Framework to Define and Guide Measure Development

Early in this research and development cycle, the University of Minnesota researchers adopted foundational concepts and analytic methods of IRT to guide item and scale development (Wilson, 2005). In earlier works (e.g., McConnell & Missall, 2008; Priest et al., 2001), IGDI research and development was governed by a classical test theory model, in which items sampled a broad domain and were selected randomly for assessment within that domain. The change to IRT allowed greater precision in constructing items that measured domains of interest, locating those items ordinally, and in selecting sets of items as scales to meet the particular functions of seasonal universal screening and progress monitoring within RTI.

This change in methodology also supported a much stronger focus on validity dimensions of the overall Center for Response to Intervention in Early Childhood (CRtIEC) model (Rodriguez, 2010). Validity is a somewhat ill-defined and broadly used concept (Newton & Shaw, 2013) that is rarely applied to relations between assessment or measurement and intervention. However, a strong relation is essential in RTI between what is measured, the decisions that measurement supports, and the scope, focus, and content of subsequent intervention. For development purposes, IGDIs reflect Kane's (1992, 2013) approach, in which validity is seen as the soundness of all interpretations and uses of a test as presented in various claims or arguments.

The logical foundation for this relation is likely clear; if similar or closely related dimensions of child performance are not the basis for both assessment decision-making and intervention, it is unlikely that a universal screening and progress monitoring system will help identify children who will indeed benefit from resulting intervention. Also, if alignment of measurement and intervention is very high, children who receive intervention will perform at higher levels when assessed. Although the logic for this relation may be clear, to date there is relatively little history of formal investigation of validity relations between measurement and intervention in early childhood and in the context of RTI.

Creating Domain Specifications and Item Content

Domain Specification

Center for Response to Intervention in Early Childhood investigators worked to define common assessment and intervention foci early in the Center's effort. At the outset, teams of researchers who later would be responsible for, respectively, intervention or assessment procedures worked cooperatively to identify domains of assessment (i.e., oral language, comprehension, phonological analysis, and alphabet knowledge) and operational definitions for each. These definitions guided the initial development of the assessment as well as Tier 2 and Tier 3 interventions. As research continued, early-stage measures and interventions were used together in early efficacy trials, and similarities or differences in what was assessed and what was taught were reviewed continuously.

Item Development

To design IGDIs 2.0, the authors worked from the common validity framework and domain specifications, as well as the research literature on skills and competencies in each domain of interest (oral language, phonological awareness, alphabet knowledge, and comprehension). In total, they designed five IGDIs that reached the end phases of the research and development process, including one measure of oral language (*Picture Naming*), two measures of phonological awareness and analysis (*Rhyming* and *First Sounds*), and one measure each of alphabet knowledge (*Sound Identification*) and comprehension (*Which One Doesn't Belong?*) (Bradfield, Besner, et al., 2014; Wackerle-Hollman, Rodriguez, Bradfield, Rodriguez, & McConnell, 2015; Wackerle-Hollman et al., 2013). The developers' goal was to create a set of measurement tools with rigorous evidence of score reliability and validity for multiple purposes, including screening and progress monitoring. All IGDIs 2.0 were presented in flash-card format; each card featured isolated photographs of everyday objects presented on a white background. Examples of all five IGDIs are presented in Figure 2.

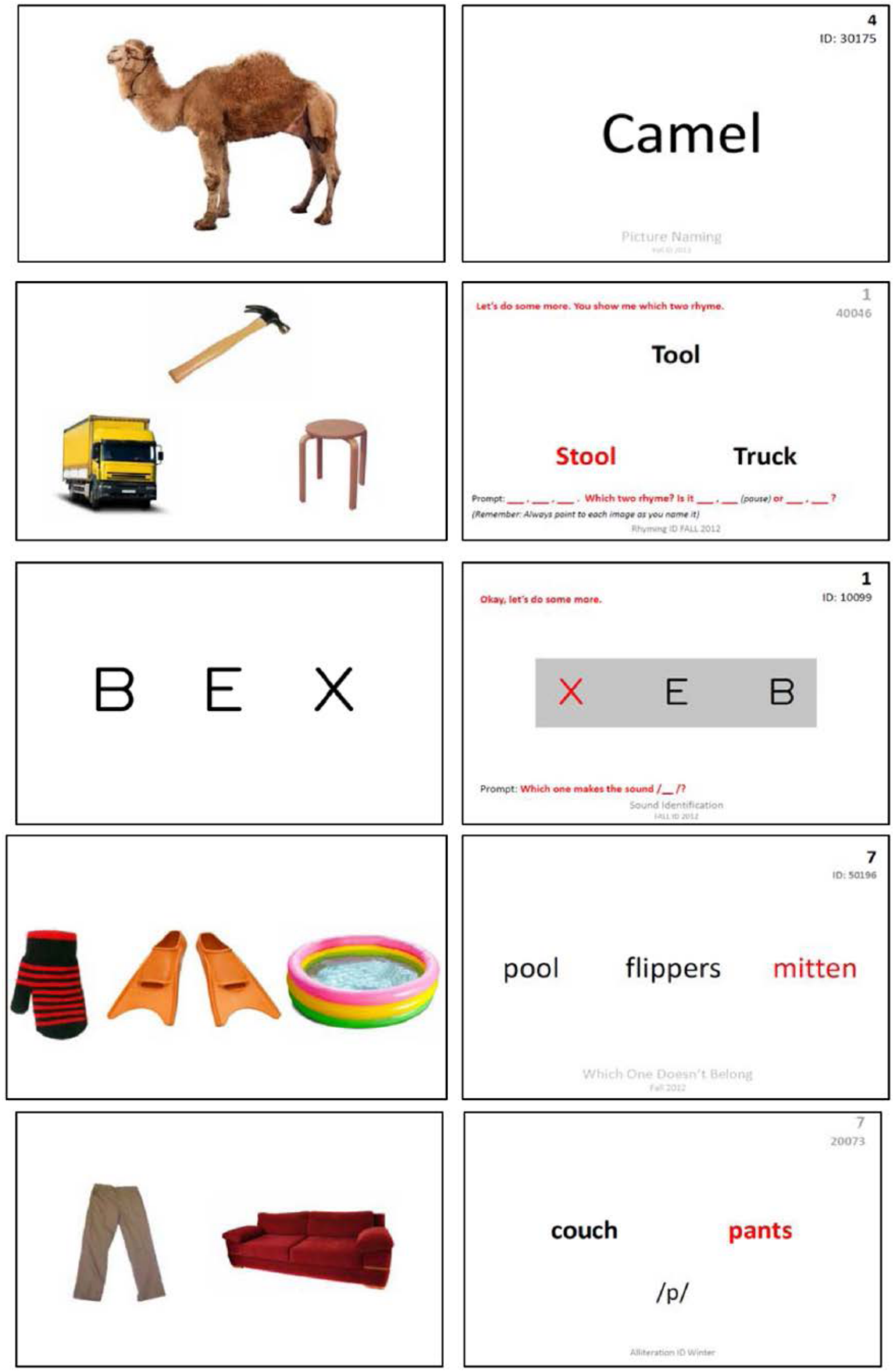

Figure 2: **IGDIs 2.0 item examples: Picture Naming, Rhyming, Sound Identification, Which One Doesn't Belong, and First Sounds. Note: All images © University of Minnesota.**

Items were developed iteratively for all five measures. Item design is a complex process that revolves around increasing the information function of each item by removing as many construct-irrelevant features as possible (Wilson, 2005). When error is reduced or removed, items become more robust indicators of true student ability.

The IGDI item development process worked to build a strong item pool by clearly articulating the content to be assessed, isolating that content in ways that allow students to meaningfully interact with it, and carefully controlling the remaining variables present in each item to manipulate the level of difficulty. Four design goals for ensuring robust item design and limiting error are described here: (a) creating construct-relevant items, (b) ensuring measurement invariance across subgroups, (c) ensuring psychometric quality for each item, and (d) sampling characteristics and robust scale development.

Creating Construct-Relevant Items

Use of IRT for new IGDIs in turn required developing items that, in every instance, were strong and specific samples of a particular domain being assessed. This focus on maximizing information and minimizing error is a common feature of any rigorous measurement development (Cronbach, 1990), affecting both the psychometric quality and overall efficiency of any test. To the extent that test developers can gather more data of direct interest, or "signal," and reduce effort devoted to measuring error, or "noise," the value and quality of gathered information increase.

In this work, child responses across an array of items from any one measure needed to fully represent a child's current skill within a particular domain, and any information presented in an item that did not contribute to, or detracted from, this representation, was a threat to the validity argument. Construct-irrelevant variance exists when a sample or item "contains excess reliable variance that is irrelevant to the interpreted construct" (Messick, 1989, p. 34). To identify sources of possible error in the items, the development team focused particularly on reducing construct-irrelevant variance (Downing, 2002).

To control for the content of assessment and to facilitate easy administration and scoring, IGDIs often use standard images presented on cards. These images present the single greatest potential source of construct-irrelevant variance. In some instances, features of an image may confuse or distract from the specific skill being measured (for instance, an image of a cow in a fenced pasture might be labeled "cow" or "farm" when only the former is the desired response), or may provide cues that guide correct (or incorrect) responding regardless of the child's skill (for instance, when one of three response options is a photo and the correct response and the other two distractors are line drawings). In cases like these, individual test takers' background knowledge, interest, and opinions about any given image—and not their ability to respond to the intended content of the item—may affect their performance.

Several different tactics decrease construct-irrelevant features in new IGDIs. First, whenever possible, each image is presented as isolated on a white background. This prevents additional content contributing to variance in performance. Second, whenever possible, photographs of actual objects are used. By reducing the symbolic representation present during a student's interaction with the items, University of Minnesota researchers reason that students' varying experience with those symbols will be eliminated (Horst, Samuelson, Kucker, & McMurray, 2011).

Third, for receptive tasks where a target is compared with two or three possible responses, IGDIs controlled difficulty of items by varying characteristics of the distractors, including number or similarity of distractors. Fewer choices (i.e., the correct response and one distractor) were generally less difficult than items with more choices (i.e., correct response and two distractors). Similarly, difficulty was varied by altering the similarity or difference of distractors and correct responses; for instance, a *Which One Doesn't Belong?* item with images of "horse, cow, and paper" was less difficult than images of "horse, cow, and fish," with "paper" sharing no functional, feature, or class similarities with the other images, and "fish" potentially included in the same class.

The research team evaluated control for construct-irrelevant features and effects on item difficulty empirically. In these studies (Bradfield, Besner, et al., 2014; Wackerle-Hollman et al., 2015; Wackerle-Hollman et al., 2013), student response patterns were carefully studied to identify consistent errors that interfered with assessment performance across participants. For example, when presented with seemingly novel images in receptive tasks, students appeared to choose a response that was most common or had high interest value (Horst et al., 2011). One specific image of a campfire was often selected by children, whether or not it was the correct answer and regardless of the child's ability. As a result, "campfire" was removed from the distractor image pool to reduce construct-irrelevant features in the items.

Psychometric Quality at the Item Level

Population-Level Analyses

This approach to measure development assumes that each IGDI is internally consistent, or that items used in that measure are related to one another and, thus, to the same developmental domain or construct. To test this assumption, authors examined item-level statistics produced in a Rasch model, a more prescriptive form of IRT (Wilson, 2005). The Rasch model produces a series of item-level statistics that allow for critical evaluation of each item's contribution to the item pool. Using Rasch, they reviewed four metrics for each item: *measure*, or item location, the empirical difficulty value of each item (i.e., ability level required to have a 50% chance of correctly responding to the item); *mean-square infit and outfit statistics*, which provide information about each item's fit with the measurement model near the item location (*infit*) and at the extremes (*outfit*). In addition, although not a component of the Rasch model, *point-biserial correlations* (PBSC), or the degree to which a specific item correlates with the total score of a scale and contributes to a reliable total score, were evaluated, as Rasch models assume relative uniformity in item discrimination, a function of item-total correlations. In development of the IGDIs, retained items met explicit criteria: infit and outfit statistics less than 2, PBSC greater than 0.2, and item means between 0.2 and 0.8.

Item difficulty statistics were examined to ensure that University of Minnesota researchers had relatively large pools of items with difficulty values below mean ability (i.e., easier items located lower on the Rasch ability scale) to allow scales to assess and distinguish the performance of candidates for Tier 2 or Tier 3 interventions. Finally, item discrimination, or the degree to which an item is able to discern between students with high levels of ability and students with low levels of ability, was evaluated through item-total correlations and item-level distractor frequency tables.

After initial selection of assessment content and preparation of individual items, UM researchers assessed item functioning across these four metrics in relatively large samples of preschool children (cf. Bradfield, Besner, et al., 2014; Wackerle-Hollman et al., 2013). Items that met quality criteria were retained, and items that failed to meet one or more quality standards were examined and, if possible, revised. These revised items were then evaluated with a second sample; only those items that now met quality standards were added to the item pool for each measure.

Measurement Invariance Across Subgroups

IGDIs 2.0 are designed to be used in programs with diverse student populations. In particular, UM researchers assume these programs include many children who are English Language Learners (ELLs) or receive special education services. Given the likely presence of both ELL students and those with disabilities participating in IGDI assessment, it is imperative to minimize or control bias in these measures.

IRT provides strong tools for assessing bias of differential item functioning, or DIF, across known subgroups. In simple terms, DIF identifies individual items that sample a skill of interest differently in two groups; while the two groups may vary in their likely response to these items (that is, one group can be more likely to respond incorrectly), the relation of that item to individuals' overall ability should not vary between the groups. IGDI developers assessed DIF to examine and identify the extent to which any item performed differently, suggesting bias, for ELL or special education students (Walker, 2011). To defend the ability to interpret scores in consistent ways across subgroups, items that functioned consistently, with item location, conditioned on overall ability, not varying by subgroup, were retained.

A small percentage (<10%) of items were identified that demonstrated DIF. Because these DIF results represented so few items, they were removed from the pool. The remaining item pool contained items that functioned equally well across the population of interest (4- to 5-year-old students who speak English or another language at home and who have or do not have an identified disability).

Collectively, these analyses identify items that were conceptually and empirically strong contributors to assessment in each of the four domains of interest, and also provide information for further development or refinement of items in areas of ability not yet well sampled by available items. As a result of the validity analysis, the DIF analysis, and the Rasch analysis and an iterative item development and testing process, the IGDI developers were able to select and retain only those items that met all criteria for stability and consistency with the intended areas of assessment.

Setting Standards for Evaluating Performance: Benchmarking

Once individual items had been constructed, thoroughly tested, and revised in isolation, development turned to selecting and combining items into scales used to inform assessment decisions in universal screening and progress monitoring. This work proceeded in two steps: calibrating and locating items on a common ability scale, and then identifying levels of child performance associated with critical assessment decisions.

Item Location and Calibration

Rasch models, where data fit the model well, have the property of *sample-invariance*, producing consistency in relative item difficulty and overall scale characteristics across different samples of respondents, and, similarly, person-ability estimates independent of the items chosen for a given administration. To support quality results, items must be calibrated and located on an arbitrary ability scale with a large and heterogeneous sample (Embretson & Reise, 2000). Calibration is the process of using a measurement model to estimate the parameters describing the measurement characteristics of a set of items, and with the Rasch model, estimating item location or item difficulty.

In developing new IGDIs, item calibration was routinely completed with samples of more than 1,000 preschool children from classroom programs across the United States, including public and private preschool programs, Head Start centers, and fee-based child care programs. The resulting samples varied in ethnicity, special education status, socioeconomic status, and geography; all participants participated in English-language instructional activities, although some spoke languages other than English at home. Children responded to a set of “common” items as well as a subset of remaining items for each IGDI, with more than 100 children responding to each item in the pool.

These child responses were then used to calibrate each item and, in turn, to calculate a Rasch score for each child. Rasch calibration locates items and children on the same scale, allowing researchers to evaluate relative difficulty for individual items as a function of the assessed performance level of individuals or groups. In total, an initial item pool of more than 160 items was produced for each new IGDI.

Preliminary Benchmarking

One of the major functions of these new measures is to identify individual children who may benefit from more intensive intervention—the *screening* function of RTI assessment systems. Note the assumption that this need for differentiated intervention can be defined as a difference between a child's skill at any point in time and the level of skill needed to achieve later reading competence. To this end, IGDIs have identified *benchmarks*, or levels of performance for each season of the academic year that represent this expected level of development.

Two specific challenges exist in developing universal screening benchmarks in early childhood programs. First, the time lag from preschool assessment to later reading proficiency is long; empirical benchmarking is affected by features such as the passage of time (from four-year-old preschool to third-grade reading), confounding intervening variables (instruction between PK4 and third grade), and definitional challenges. Because of these and other challenges, benchmarking is an iterative process of setting and refining benchmarks for preschool performance with the intent of identifying a credible criterion for decision making and refining benchmarks over time.

Second, screening in most multi-tiered system of supports (MTSS) systems, including RTI, assumes universal access to a population of interest (e.g., all third-grade students). However, current publicly funded early childhood programs in the United States are not typically universal or population-based; rather, many programs have been specifically created for and only include individuals from selected groups (e.g., children living in poverty, children with disabilities). Furthermore, these selection factors frequently relate closely to empirically identified risk factors for delays in language and early literacy development (Snow et al., 1998). As a result, it was not reasonable to assume a "typical population" distribution of performance, but rather, a distribution of risk and performance that leads to relatively higher proportions of individuals who might require more intensive intervention.

Because of this, preliminary IGDI benchmarks were established using criterion-referenced approaches. Specifically, initial benchmarks used both teacher judgments and standardized achievement tests of individual child performance to identify levels of performance associated with need for additional intervention (Bradfield, Besner, et al., 2014; Wackerle-Hollman et al., 2013). A three-step standard-setting process was used. First, "tier-level descriptors" were developed for each domain and measure. These tier-level descriptors provided operational definitions for each domain assessed, and described three levels of child performance: children who were demonstrating expected levels of performance and making adequate progress toward end-of-preschool expectations given the typical level of classroom intervention, children who were demonstrating less than expected levels of performance and were unlikely to achieve end-of-preschool expectations without some modest increase in intervention intensity, and children who were demonstrating less than expected levels of performance and were unlikely to achieve end-of-preschool expectations without some significant increase in intervention intensity.

Next, teachers whose students had recently completed IGDI 2.0 assessments were asked to consider these three descriptors of student performance and likely need for support, and to identify the appropriate level for each individual student. These teacher judgments were then used as independent variables for contrasting groups analyses (Cizek & Bunch, 2007), or empirical analyses to identify IGDI Rasch score benchmarks that best differentiated groups of students identified by their teachers. Regression analyses and receiver operator curves were reviewed, and specific benchmarks identifying the Rasch location that best distinguished between Tier 1 and Tier 2/3 classification.

Early in this set of analyses, the researchers determined that teachers demonstrated low reliability in making fine-point distinctions between children identified as candidates for Tier 2 or Tier 3 intervention. For this reason, a single benchmark separating Tier 1 and Tier 2/3 candidates was identified for each measure (Bradfield, Besner, et al., 2014; Wackerle-Hollman et al., 2015; Wackerle-Hollman et al., 2013).

Developing Scales for Seasonal Screening and Progress Monitoring

Card Administrations

Preliminary benchmarks contributed directly to selection of items for seasonal universal screening scales for the card-based version of myIGDIs. With the Rasch benchmark locations established, universal screening scales were constructed for each of the five measures, selecting 15 items located within 0.50 logits of the Rasch benchmark location. This tight clustering of items increases each scale's efficiency for describing a child's performance as above or below a particular benchmark (thus increasing information for this particular function of assessment), but at the cost of providing reliable measurement of performance across the full ability scale (Embretson & Reise, 2000).

For card-based progress monitoring, items were selected with Tier 2/3 candidates in mind. As such, items within the progress monitoring sets were selected from below the criterion benchmark between Tier 1 and Tier 2/3 specified in the screening scale. Furthermore, because of the focus on evaluating growth over time, 30 items were selected for each scale to allow the student more opportunities to interact with testing materials and improve IGDI 2.0 tasks' sensitivity to growth. Progress monitoring sets were constructed for two seasons, fall and winter, with two empirically parallel forms for each season.

Although screening and progress monitoring represent two primary IGDI 2.0 scales, they are not the only scales that are available from existing item pools. Using information from Rasch analyses and different selection criteria, scales could be constructed to assess growth in higher levels of performance (including progress monitoring in Tier 1 or for gifted and talented students), more general scales of performance that span the ability range represented in each scale (for descriptive research or program evaluation), or to identify more efficiently current level of performance for individual children (using, for instance, computer-adaptive testing procedures). These principles, along with development and evaluation of computer-adaptive testing formats for item selection, contributed directly to development of the iPad-based format for language and early literacy assessment.

iPad Administrations

Research and development of this new, card-based version of myIGDIs was accompanied by rapid development of personal computing—specifically, the iPad developed by Apple, Inc. Opportunity to merge the expanded and IRT-based item pool with administration and scoring on digital devices produced a new form of myIGDIs.

iPad administration is completed with two devices, connected via Bluetooth. The examiner's device provides a variety of dashboards (class roster, assessment results to date), as well as capacity to conduct an individual child assessment. After the two devices are connected via Bluetooth, the examiner's device provides both general administration directions and text to read to the child and gesture controls to initiate assessment, present images being displayed on the child's device, and score child responses. The child's device presents images under control of the examiner's actions, with some prompts (e.g., expanding images to highlight individual photos) to assist in administration.

Most significantly, adoption of digital administration and scoring allowed for conversion of myIGDIs on iPads to use computer-adaptive testing methods for item selection. Seasonal screening assessments are 15 items in length (like card-based assessments) and progress-monitoring assessments have 25 items. For each assessment, the first item presented to the child is close to either a recent myIGDIs ability estimate for that child or the seasonal "cut score" for that measure. Successive items in that administration are selected based on child performance; if the child responds to an item correctly, a more difficult item is administered next, and if the child responds incorrectly, an easier item is administered next. In this way, iPad-administered myIGDIs provide information on child performance relative to seasonal benchmarks like the card-based version, but also provide more precise estimates of child ability on every assessment occasion. This is particularly important for progress monitoring, where child-specific assessment is key.

Scale Validation and Features and Quality of Decision Making

After a rigorous process for scale construction, yielding the IGDI 2.0 screening and progress monitoring scales, attention turned to collecting evidence to support interpretation and use arguments. This included work to develop, evaluate, and revise a decision-making framework (DMF) that places IGDIs 2.0 within a broader context for selecting and evaluating interventions for RTI in early childhood language and early literacy.

Screening/Identification

As noted earlier, universal screening scales using the card administration format were constructed specifically to maximize information about individual children's performance at and near a seasonal benchmark or standard for judging need for additional intervention. The universal screening IGDIs have now been applied in a number of field studies, in which researchers have evaluated tier candidacy decisions made with IGDIs 2.0 only by comparing these results with one or more standardized tests or other measures of language and early literacy. Across domains, congeneric or internal reliability estimates range from 0.74 to 0.90 for individual scales, and concurrent construct validity coefficients range from 0.49 to 0.71. Classification accuracy, a difficult concept to assess in isolation (McConnell et al., 2014), has also been established. In an effort to reduce false-negative decisions (decisions that deny access to more intensive intervention to those who would likely benefit from it), sensitivity for all benchmarks has been held to a minimum of 0.70 (Jenkins, Hudson, & Johnson, 2007). With this constraint, specificity levels varied, with a mean of 0.56 across measures (Bradfield, Besner, et al., 2014).

Contextual Characteristics and Technical Challenges for Measurement in EC RTI

Work on IGDIs 2.0 to date has been intentional and iterative, working to both refine and improve assessments while monitoring the development and evaluation of related interventions (see Kelley & Goldstein, 2015; Kaminski, Powell-Smith, Hommel, McMahon, & Aguayo, 2015). At the same time, contextual features of early childhood education and technical challenges in the development of assessment and intervention in early childhood RTI affect the utility, and thus the likely implementation and impact, of measurement systems and larger RTI models.

Contextual Characteristics

Although a host of contextual characteristics might be identified, two have been particularly noteworthy. First, and perhaps most important, the utility and efficacy of any RTI or other MTSS measurement system are directly affected by the populations served in existing preschool programs that may implement these procedures. Although universal access to pre-kindergarten programs is being provided in increasing numbers of states and local communities, it is still generally true that many preschool programs—particularly those funded by federal, state, or local governments—serve a higher proportion of "high-risk" children. The children served in these programs are, by design, more likely to demonstrate lower levels of development in language and early literacy skills, and thus more likely to require some form of relatively intensive intervention to achieve expected levels of performance. That is, children at risk may require more intensive intervention than the Tier 1 interventions provided in typical preschool settings that do not target children at risk. As a result, base rates of children identified in traditional RTI-like tiers of intervention may be significantly higher; depending on criteria considered, as many as 80% of children enrolled in some programs may be performing at levels that would indicate Tier 2 or greater service in traditional RTI models (Carta et al., 2015).

Although it is still very likely that differentiation of services would be of benefit in these selected-population programs, strategies for implementing differentiated service models are not yet well developed. What would constitute "Tier 1" services in these settings? What outcome performance standards are reasonable for children served in these settings? How can more intensive intervention, a level that may be required for many children enrolled in these programs, be delivered efficiently and with fidelity? These are questions that must be addressed before the utility and accuracy of any measurement model can be appropriately evaluated.

In addition, increasing expectations for performance standards at kindergarten entry, and the availability of specific instructional services to help children achieve these higher standards, will affect the criteria that any measurement system must meet. How these standards change, how they are described, and their relation to the constructs and validity standards reported here (or for any other measurement system) are still very much in process. As these changes become clear, some adjustment, refinement, or even substantial revision of measurement procedures may be needed.

Finally, direct assessment and use of child-level performance data are still relatively new practices in many early childhood programs (Carta et al., in press; Greenwood, Carta, & McConnell, 2011). Adoption of "on-demand" seasonal assessments, and high-fidelity use of the data resulting from these assessments, may represent implementation challenges in some preschool settings.

References

Bradfield, T. A., Besner, A., Wackerle-Hollman, A. K., Albano, A., Rodriguez, M. C., & McConnell, S. R. (2014). Redefining individual growth and development indicators: Oral language. *Assessment for Effective Intervention*, *39*, 233–244. doi:10.1177/1534508413496837

Bradfield, T. A., Vue, K., Rodriguez, M. C., & McConnell, S. R. (2014, February). *Confident decision-making: Validation of an early childhood response to intervention decision-making framework.* Paper presented at the Conference for Research Innovation in Early Intervention, San Diego, CA.

Buzhardt, J., Greenwood, C. R., Walker, D., Anderson, R., Howard, W. J., & Carta, J. J. (2011). Effects of web-based support on Early Head Start home visitors' use of evidence-based intervention decision making and growth in children's expressive communication. *NHSA Dialog: A Research-to-Practice Journal for the Early Childhood Field*, *13*, 121–146.

Carta, J. J., Greenwood, C. R., Atwater, J., McConnell, S. R., Goldstein, H., & Kaminksi, R. A. (2015). Identifying preschool children for higher tiers of language and early literacy instruction within a response to intervention framework. *Journal of Early Intervention*, *36*(4), 281–291. doi:10.1177/1053815115579937

Carta, J. J., Greenwood, C. R., Goldstein, H., McConnell, S. R., Kaminski, R., Bradfield, T. A., . . . Atwater, J. (in press). Advances in multi-tiered systems of support for prekindergarten children: Lessons learned from 5 years of research and development from the Center for Response to Intervention in Early Childhood. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *The handbook of response to intervention: The science and practice of multi-tiered systems of support* (2nd ed.). New York, NY: Springer.

Chard, D. J., Harn, B. A., Sugai, G., Horner, R. H., Simmons, D. C., & Kame'enui, E. J. (2008). Core features of multi-tiered systems of reading and behavioral support. In C. R. Greenwood, T. Kratchowill, & M. Clements (Eds.), *Elementary school-wide prevention models: Real models and lessons learned in real schools* (pp. 31–60). New York, NY: Guilford.

Cizek, G. J., & Bunch, M. B. (Eds.). (2007). *Standard setting: A guide to establishing and evaluating performance standards on tests.* Thousand Oaks, CA: SAGE.

Cronbach, L. J. (1990). *Essentials of psychological testing* (5th ed.). New York, NY: Harper & Row.

Downing, S. M. (2002). Construct-irrelevant variance and flawed test questions: Do multiple-choice item writing principles make any difference? *Academic Medicine*, 77, S103–S104.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists.* New York, NY: Psychology Press.

Frieden, T. R. (2010). A framework for public health action: The health impact pyramid. *American Journal of Public Health, 100*, 590-595. doi:10.2105/AJPH.2009.185652

Fuchs, L. S., & Deno, S. L. (1991). Paradigmatic distinctions between instructionally relevant measurement models. *Exceptional Children, 57*, 488–500.

Fuchs, L. S., & Fuchs, D. (2007). The role of assessment in the three-tier approach to reading instruction. In D. Haager, J. Klingner, & S. Vaughn (Eds.), *Evidence-based reading practices for response to intervention* (pp. 29–44). Baltimore, MD: Brookes.

Fuchs, L. S., & Fuchs, D. (2010). *Using CBM to determine response to intervention.* Washington, DC: National Center on Response to Intervention.

Good, R. H., Gruba, J., & Kaminski, R. A. (2002). Best practices in using dynamic indicators of basic early literacy skills (Dibels) in an outcomes-driven model. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology* (4th ed., Vol. 1, pp. 699–720). Washington, DC: National Association of School Psychologists.

Greenwood, C. R., Bradfield, T., Kaminski, R., Linas, M., Carta, J. J., & Nylander, D. (2011). The response to intervention (RTI) approach in early childhood. *Focus on Exceptional Children, 43*(9), 1–22.

Greenwood, C. R., Carta, J. J., Goldstein, H., Kaminski, R. A., McConnell, S. R., & Atwater, J. (2015). The Center on Response to Intervention in Early Childhood: Developing evidence-based tools for a multi-tier approach to preschool language and early literacy instruction. *Journal of Early Intervention, 36*(4), 246-262. doi:10.1177/1053815115581209

Greenwood, C. R., Carta, J. J., & McConnell, S. R. (2011). Advances in measurement for universal screening and individual progress monitoring of young children. *Journal of Early Intervention, 33*, 254–267. doi:10.1177/1053815111428467

Greenwood, C. R., Walker, D., & Buzhardt, J. (2010). The Early Communication Indicator (ECI) for infants and toddlers: Early Head Start growth norms from two states. *Journal of Early Intervention, 32*, 310–334.

Horst, J., Samuelson, L., Kucker, S., & McMurray, B. (2011). What's new? Children prefer novelty in referent selection. *Cognition, 118*, 234–244.

Jenkins, J. R., Hudson, R. F., & Johnson, E. S. (2007). Screening for at-risk readers in a response to intervention framework. *School Psychology Review, 36*, 582–600.

Kaminski, R. A., Powell-Smith, K. A., Hommel, A., McMahon, R., & Aguayo, K. B. (2015). Development of a Tier 3 curriculum to teach early literacy skills. *Journal of Early Intervention, 36*(1), 313–332. doi:10.1177/1053815115581210

Kane, M. T. (1992). An argument-based approach to validation. *Psychological Bulletin, 112*, 527–535.

Kane, M. T. (2013). The argument-based approach to validation. *School Psychology Review, 42*, 448–457.

Kelley, E. S., & Goldstein, H. (2015). Building a Tier 2 intervention: A glimpse behind the data. *Journal of Early Intervention, 36*(4), 292–312. doi:10.1177/1053815115581657

Kratochwill, T. R., Clements, M. A., & Kalymon, K. M. (2007). Response to intervention: Conceptual and methodological issues in implementation. In S. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of response to intervention* (pp. 25–52). New York, NY: Springer.

McConnell, S. R., Bradfield, T., & Wackerle-Hollman, A. (2014). Universal early literacy screening. In R. Kettler, T. Glover, C. Albers, & K. A. Feeney-Kettler (Eds.), *Universal screening in educational settings: Identification, implications, and interpretation* (pp. 141–170). Washington, DC: American Psychological Association.

McConnell, S. R., & Greenwood, C. R. (2013). General outcome measures in early childhood and the Individual Growth and Development Indicators. In V. Buysse & E. Peisner-Feinberg (Eds.), *Handbook of response to intervention in early childhood* (pp. 143–154). Baltimore, MD: Brookes.

McConnell, S. R., & Missall, K. N. (2008). Best practices in monitoring progress for preschool children. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology* (5th ed., pp. 561–573). Washington, DC: National Association of School Psychologists.

McConnell, S. R., Priest, J. S., Davis, S. D., & McEvoy, M. A. (2002). Best practices in measuring growth and development for preschool children. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (4th ed., pp. 1231–1246). Washington, DC: National Association of School Psychologists.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 12–103). New York, NY: American Council on Education.

National Early Literacy Panel. (2009). *Developing early literacy: Report of the National Early Literacy Panel—A scientific synthesis of early literacy development and implications for intervention.* Jessup, MD: Author.

Newton, P. E., & Shaw, S. D. (2013). Standards for talking and thinking about validity. *Psychological Methods, 18*, 301-319. doi:10.1037/a0032969

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

Rodriguez, M. C. (2010). *Building a validity framework for second-generation IGDIs* (Technical reports of the Center for Response to Intervention in Early Childhood). Minneapolis: University of Minnesota.

Reliability and Validity

Reliability Analyses

Standard Error of Measurement

The iPad version of myIGDIs utilizes a computerized adaptive testing model (CAT; Linacre, 1998) that improves student ability estimates and decreases administration time. As one of the most desired properties in CAT administration, the precision of test-taker ability estimate is improved compared with the traditional paper and pencil fixed format assessments due to the fact that CAT delivers each item based on the ability estimate from the test taker's previous response. In other words, each administered item in the CAT assessment is the one that matches best with the given ability estimate of test takers.

Standard Errors of Measurement (or SEM) for all five English language and early literacy IGDIs are summarized in Table 1. Distributions for Alliteration, Picture Naming, and Rhyming are similar with a mean of 0.38 and standard deviation around 0.017, whereas the SEM distribution of Sound Identification and Which One Doesn't Belong are more variable with higher means. Minimum values of SEM across the five measures are approximately the same. For all five measures, ability estimates with best precision for each measure range from –0.94 to 0.32. Additionally, the ability estimate with the lowest precision level for each measure ranges from 1.94 to 3.85, ability ranges located where there are fewer items.

Table 1: Descriptive Summaries of Standard Errors of Measurement in Progress Monitoring Assessments across Five Measures

			SEM Range			
Measure	**Mean**	**SD**	**Min SEM**	**Ability**	**Max SEM**	**Ability**
Alliteration	0.381	0.018	0.368	–0.48	0.442	2.05
Picture Naming	0.382	0.016	0.368	–0.06	0.424	3.40
Rhyming	0.380	0.017	0.368	0.20	0.443	1.94
Sound Identification	0.408	0.045	0.370	0.32	0.544	2.81
Which One Doesn't Belong	0.471	0.094	0.368	–0.94	0.621	3.85

Note: All scores expressed in Rasch units.

Scatter plots between final ability estimates and corresponding SEM across five measures are presented in Figure 4. A number of observations can be made from Figure 4. First, the scatter points are relatively more clustered for Alliteration, Picture Naming, and Rhyming than the distributions in Sound Identification and Which One Doesn't Belong. Second, for Picture Naming and Sound Identification, low-performing students (relative to the location of the items) were identified by the progress monitoring assessments, whereas low-performing students were missing for the other three measures. It is important to note that students see different item sets in progress monitoring assessments since the CAT model is implemented and the distribution of final ability estimates (and SEM) are dependent on the ability of the sample of participants.

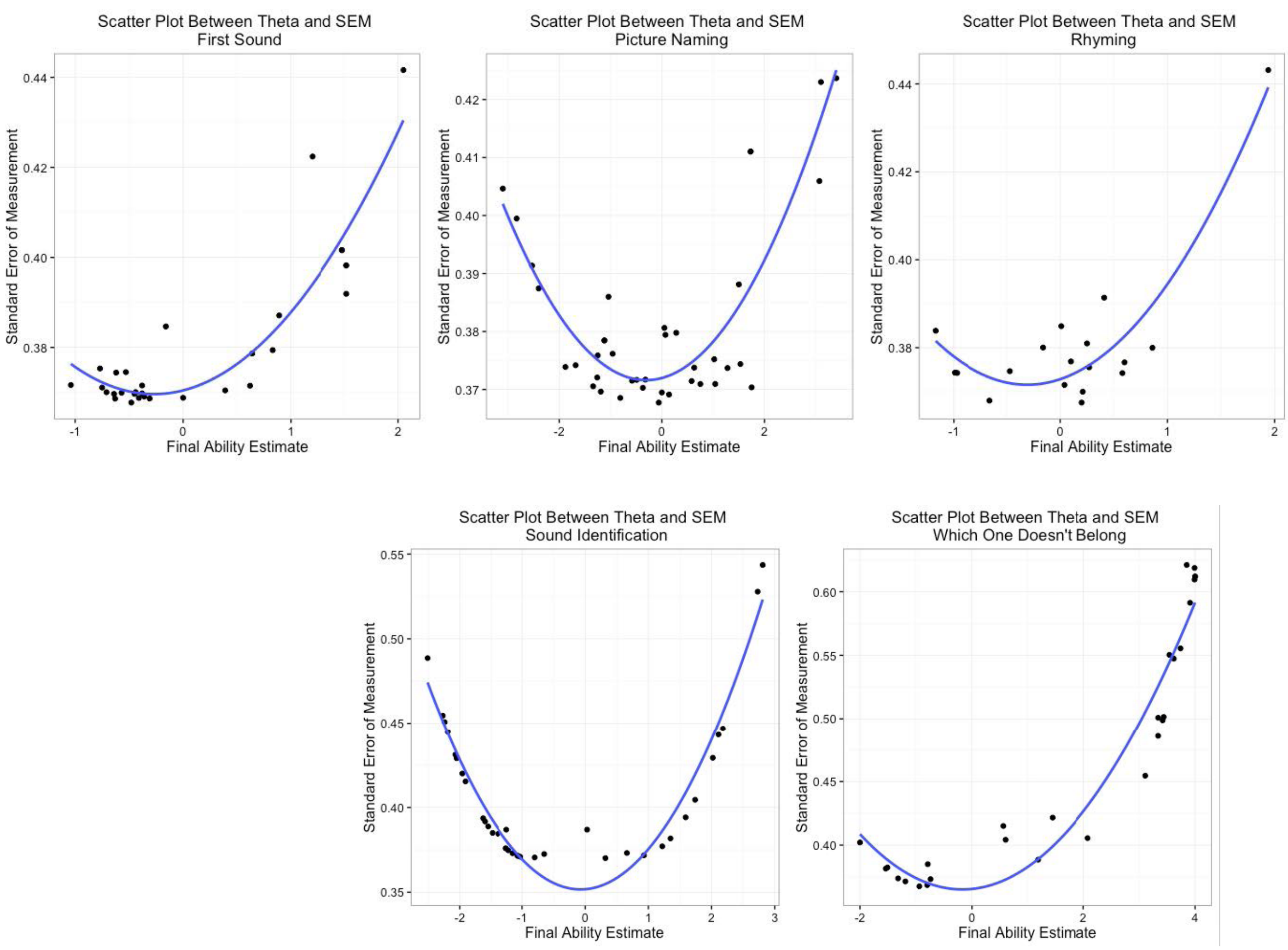

Figure 3: Scatter plots between final ability estimate and corresponding standard error of measurement across five measures.

Confirmatory Factor Analysis

To better understand the structure of the IGDI measures, and to provide evidence of unidimensionality to support the use of the Rasch model, UM researchers conducted two forms of confirmatory factor analysis (CFA). The first tests the fit of the data to a unidimensional model independently, and second to test a two-factor model allowing the factors to correlate. For each model, two fit indices are reported, including the comparative fit index (CFI) where good fit is found with values greater than 0.95 and the root mean squared error of approximation (RMSEA) where good fit is found with values less than 0.08 (Brown, 2006). The first independent unidimensional models fit very well. For Alliteration, the CFI was 0.986 and RMSEA was 0.026. For Rhyming, the CFA was 0.922 and RMSEA was 0.072. The combined model allowing the two measures of PA to correlate yielded a CFI of 0.961 and RMSEA of 0.038, with a correlation between the factor scores of Alliteration and Rhyming (removing measurement error) of 0.75 (56% common variance between the constructs of Alliteration and Rhyming). The CFA results indicate adequate to excellent fit. UM researchers also completed CFAs for the Picture Naming measure, where results were 0.980 and 0.04 respectively; for the WODB measure, where results were 0.913 and 0.074; and finally for Sound Identification, where results were 0.971 and 0.054.

Internal Consistency

Reliability estimates for the card- and iPad-based item pools were obtained from data collected from the IGDI 2.0 screening measures administered during the fall of 2012. Screening measures consisted of 15 items each. Data was collected from different sites (Lake Cristal, Monona, Anchorage, South Washington County, Mounds View, and Spring Lake Park).

Reliabilities were estimated using jMetrik 2.1 (Meyer, 2011). The reliability estimates are based on the congeneric measurement model, which allows each item to load on the common factor at different levels and allows item error variances to vary freely (each item can be measured with a different level of precision). This is the most flexible measurement model and most appropriate for measures with few items.

Table 2: Reliabilities for Screening Measures (Fall 2012)

	Congeneric	N
Sound Identification	0.81	622
Picture Naming	0.74	926
Rhyming	0.90	462
Which One Doesn't Belong	0.81	727

Classification Accuracy

Original developers evaluated the classification accuracy of screening decisions made with items included in the card-based assessments. In general, cut scores (expressed in Rasch units) increase across seasons, sensitivity rates are high-moderate (above 0.7) and specificity rates slightly lower. Results, by season, are presented in the following table.

Table 3: Cut Score, Sensitivity, Specificity, and Area Under the Curve by IGDI 2.0 by Season

	Fall				Winter				Spring			
	Cut	**Sens**	**Spec**	**AUC**	**Cut**	**Sens**	**Spec**	**AUC**	**Cut**	**Sens**	**Spec**	**AUC**
Picture Naming	0.73	72	64	76.6	1.42	73	63	73.1	1.14	73	69	76.8
Rhyming	0.24	73	59	69.4	1.05	75	62	69.3	1.60	72	58	71.0
Alliteration	0.42	73	52	64.9	0.77	73	57	69.7	1.21	72	61	68.4
Sound ID	0.28	71	51	66.8	1.08	72	56	67.1	2.25	71	40	61.7
Which One Doesn't Belong	−0.07	71	57	69.6	0.13	77	57	70.6	0.57	75	38	58.7

Note: Sens = sensitivity; Spec = specificity; AUC = area under the curve (%).

Validity Analyses

Criterion Correlations

A variety of concurrent, construct, and predictive validity studies have been completed across the history of research on Individual Growth and Development Indicators. Two sets of correlations, collected in recent research, summarize these relations.

Further evidence to support extrapolation inferences is documented through concurrent construct-related validity correlations with field-recognized standardized assessments that capture performance in the intended domain.

In a study conducted in the winter of 2011, the IGDI 2.0 measures were administered to a sample of preschool students between the ages of three to five years, served in a variety of early care and education settings in the metropolitan Twin Cities area and Columbus, Ohio. Concurrent correlations with standardized measures of matching constructs were as follows:

Table 4: IGDI Correlations with Criterion Measures

	Sound ID	Alliteration	Rhyming	Picture Naming	Which One Doesn't Belong
CELF SS_raw					0.67
CELF WS_raw					0.68
CELF EV_raw					0.71
PPVT_raw				0.68	
TOPEL PA_raw		0.61	0.49		
TOPEL PK_raw	0.71				

Note: *n* = 53 to 58.

CELF SS = Comprehensive Evaluation of Language Fundamentals Sentence Structure subtest

CELF WS = Comprehensive Evaluation of Language Fundamentals Word Structure subtest

CELF EV = Comprehensive Evaluation of Language Fundamentals Expressive Vocabulary subtest

PPVT = Peabody Picture Vocabulary Test (4th Edition)

TOPEL PA = Test of Preschool Early Literacy Phonological Awareness subtest

TOPEL PK = Test of Preschool Early Literacy Print Knowledge subtest

Sound Identification is capturing skills representative of the construct of alphabet and print knowledge in a similar manner to the Test of Preschool Early Literacy Skills – Print Knowledge subtest (TOPEL PK), a standard criterion for this domain. Alliteration and Rhyming show some convergence, but while they are related, they are definitely tapping a somewhat different skill set than the Test of Preschool Early Literacy Skills – Phonological Awareness subtest (TOPEL PA). The Picture Naming IGDI is capturing skills representative of oral language development in a similar manner to the Peabody Picture Vocabulary Tests (4th Edition) and Comprehensive Evaluation of Language Fundamentals Expressive Vocabulary subtest (CELF EV), which are well-regarded criterion measures of the construct. Finally, Which One Doesn't Belong (WODB) appears to be capturing skills representative of the comprehension construct similarly to the Comprehensive Evaluation of Language Fundamentals Sentence Structure and Word Structure subtests (CELF SS and CELF WS). Overall, examination of the criterion correlations between the IGDIs and related gold-standard measures provides strong evidence of the IGDIs construct-related validity to support our extrapolation inferences.

Predictive Relation to Kindergarten and Beyond

The University of Minnesota team evaluated relations between preschool IGDI performance and early elementary performance for all versions of IGDIs. An early evaluation (Missall et al., 2007) demonstrated strong and stable correlations between PK4 assessments with the original card-based form of IGDIs and measures of alphabetic principle and phonological awareness in kindergarten, and reading fluency in first grade.

More recently, they have examined relations between PK4 performance on the current card-based form of IGDIs and kindergarten performance on various measures of alphabetic principle, concepts of print, and phonological awareness, all measures associated with FastBridge Learning. While it is important to note that kindergarten measures in this evaluation do not sample child language—a common, but perhaps not fully appropriate, approach to assessment of reading in the very early grades—relations are strong and consistent.

Table 5: Correlations, IGDIs in PK4 and FAST in Kindergarten

	Letter Naming Per Min.	Letter Sounds Per Min.	Onset Sounds	Concepts of Print	FAST Composite
Picture Naming	0.30	0.36	0.36	0.33	0.30
Rhyming	0.34	0.33	0.51	0.36	0.38
Sound Identification	0.60	0.59	0.48	0.35	0.42
Which One Doesn't Belong	0.30	0.32	0.37	0.33	0.32
Alliteration	0.53	0.51	0.68	0.42	0.54
IGDI Composite	0.45	0.44	0.54	0.35	0.44

Preschool Growth and Prediction of Kindergarten Performance

Using the same sample that provided relations between PK4 IGDIs and FAST alphabetic principle and phonological awareness measures, the university-based researchers assessed IGDIs' sensitivity to growth to evaluate progress students make on *IGDI* measures over the course of one school year, and the relation between student characteristics, initial performance, and growth on *IGDI* tasks. Data were analyzed for 943 children served in publicly funded PreK programs in Iowa. These findings suggest that students in preschool classrooms continue to gain important early literacy skills. Though demographic information affects their initial (fall) ability in early literacy skills, all groups made progress at the same rate. Future analyses should address what classroom and instructional procedures drive growth and whether certain practices could help students who start out lower "catch" their peers by the end of the year.

Table 6: Picture Naming

Picture Naming			Null	Age	Female	FRL
Initial status	Intercept	B00	1.05	0.58	0.94	1.10
	Var1	B01		0.07	0.23	−0.37
Rate of change	Intercept	B10	0.62	0.68	0.62	0.60
	Var1	B11		−0.01	−0.01	0.12

Table 7: Rhyming

Rhyming			Null	Age	Female	FRL
Initial status	Intercept	B00	0.53	−0.24	0.26	0.63
	Var1	B01		0.12	0.55	−0.80
Rate of change	Intercept	B10	1.18	1.24	1.18	1.17
	Var1	B11		−0.01	−0.01	0.09

Table 8: Sound Identification

Sound Identification			Null	Age	Female	FRL
Initial status	Intercept	B00	0.53	−0.24	0.26	0.63
	Var1	B01		0.12	0.55	−0.80
Rate of change	Intercept	B10	1.18	1.24	1.18	1.17
	Var1	B11		−0.01	−0.01	0.09

Table 9: Which One Doesn’t Belong?

Which One Doesn’t Belong?			Null	Age	Female	FRL
Initial status	Intercept	B00	−0.19	−0.85	−0.48	−0.14
	Var1	B01		0.10	0.57	−0.45
Rate of change	Intercept	B10	0.81	0.92	0.86	0.80
	Var1	B11		−0.02	−0.10	0.09

Table 10: Alliteration

Alliteration			Null	Age	Female	FRL
Initial status	Intercept	B00	2.47	1.94	2.24	2.53
	Var1	B01		0.08	0.45	−0.57
Rate of change	Intercept	B10	0.23	0.12	0.29	0.24
	Var1	B11		0.02	−0.11	−0.03

Research on Individual Growth and Development Indicators for Preschool Children—Selected References

Individual Growth and Development Indicators have been focus of research for the past 15 years. A selection of reports, including peer-reviewed publications and technical reports and presentations, is provided here.

Peer-Reviewed

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., Greenwood, C. R., Goldstein, H., McConnell, S. R., Kaminski, R., Bradfield, T. A., … Atwater, J. (2016). Advances in Multi-tiered Systems of Support for Prekindergarten Children: Lessons Learned from 5 Years of Research and Development from the Center for Response to Intervention in Early Childhood. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of Response to Intervention* (pp. 587–606). Springer US.

Carta, J. J., Greenwood, C. R., Walker, D., Kaminski, R., Good, R., McConnell, S., & McEvoy, M. (2002a). Individual growth and development indicators (IGDIs): Assessment that guides intervention for young children. *Young Exceptional Children Monograph Series, 4*(15–28).

Greenwood, C. R., Carta, J. J., & Walker, D. (2002, November). *Individual growth and development indicators (IGDI's): Tools for assessing intervention results for infants and toddlers.* Kansas City.

Greenwood, C. R., Carta, J. J., & Walker, D. (2005). Individual growth and development indicators (IGDIs): Tools for assessing intervention results for infants and toddlers. In B. Heward & et al. (Eds.), *Focus on Behavior Analysis in Education: Achievements, Challenges, and Opportunities (Chapter 6)* (pp. 103–124). Columbus, OH: Pearson/Prentice-Hall.

Leichman, E. S. (2014). Predictive validity of the individual growth and development indicators: Examination of students who attend head start. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 75*(4–B(E)), No-Specified.

McConnell, S. R., Bradfield, T. A., & Wackerle-Hollman, A. K. (2014). Early childhood literacy screening. In R. Kettler, T. Glover, C. Albers, & K. A. Feeney-Kettler (Eds.), *Universal screening in educational settings:*

Identification, implications, and interpretation (pp. 141–170). Washington, D.C.: American Psychological Association.

McConnell, S. R., & Greenwood, C. R. (2013). General outcome measures in early childhood and the Individual Growth and Development Indicators. In V. Buysse & E. Peisner-Feinberg (Eds.), *Handbook of Response to Intervention in early childhood* (pp. 143–154). Baltimore MN: Paul H. Brookes.

McConnell, S. R., McEvoy, M. A., & Priest, J. S. (2002). "Growing" Measures for Monitoring Progress in Early Childhood Education: A Research and Development Process for Individual Growth and Development Indicators. *Assessment for Effective Intervention, 27*(4), 3–14.

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). https://doi.org/10.1177/2332858416653756

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. R., Wackerle-Hollman, A., Bradfield, T. A., & Rodriguez, M. C. (2012). *Individual Growth and Development Indicators of Early Literacy* (2nd ed.). St Paul MN: Early Learning Labs, Inc.

McCormick, C. E., & Haack, R. (2010). Early literacy Individual Growth and Development Indicators (ELIGDIs) as predictors of reading skills in kindergarten through second grade. *International Journal of Psychology: A Biosocial Approach, 7*, 29–40.

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 and Young Children,* (21), 241–253.

Missall, K. N., & McConnell, S. R. (2010). Early literacy and language IGDIs for preschool-aged children. In J. J. Carta & C. R. Greenwood (Eds.), *Individual Growth and Development Indicators: Tools for monitoring progress and measuring growth in very young children.* (pp. 181–201). Baltimore, MD: Paul H. Brookes.

Missall, K. N., Reschly, A., Betts, J., McConnell, S. R., Heistad, D., Pickart, M., … Marston, D. (2007). Examination of the predictive validity of preschool early literacy skills. *School Psychology Review, 36*(3), 433–452.

Moyle, M. J., Heilmann, J., & Berman, S. S. (2013). Assessment of Early Developing Phonological Awareness Skills: A Comparison of the Preschool Individual Growth and Development Indicators and the Phonological Awareness and Literacy Screening–PreK. *Early Education & Development, 24*(5), 668–686. https://doi.org/10.1080/10409289.2012.725620

Priest, J., Davis, K. N., McConnell, S., McEvoy, M., & Shin, J. (1999). Individual growth and development indicators of preschoolers' expressing meaning skills: Follow that trajectory! In *Annual meeting of the Division for Early Childhood, Council for Exceptional Children, Washington, DC*.

Rodriguez, M. I., & Wackerle-Hollman, A. K. (2015). Bridging the Gap: Individual Growth and Development Indicators—The Which One Doesn't Belong Task. *Assessment for Effective Intervention, 40*(2), 96–98. https://doi.org/10.1177/1534508414556505

Roseth, C. J., Missall, K. N., & McConnell, S. R. (2012). Early Literacy Individual Growth and Development Indicators (EL-IGDIs): Growth trajectories using a large, internet-based sample. *Journal of School Psychology, 50*, 483–501.

Wackerle-Hollman, A., Duran, L., & Rodriguez, M. (2015). *Individual Growth and Development Indicators–Español.* St Paul, MN: Early Learning Labs.

Wackerle-Hollman, A. K., Rodriguez, M. I., Bradfield, T. A., Rodriguez, M. C., & McConnell, S. R. (2015). Development of early measures of comprehension: Innovation in Individual Growth and Development Indicators. *Assessment for Effective Intervention, 40*(2), 81–95. https://doi.org/10.1177/1534508414551404

Wackerle-Hollman, A. K., Schmitt, B. A., Bradfield, T. A., Rodriguez, M. C., & McConnell, S. R. (2015). Redefining individual growth and development indicators: Phonological awareness. *Journal of Learning Disabilities, 48*(5), 495–510.

Walker, D., Carta, J. J., Greenwood, C. R., & Buzhardt, J. (2008). The use of Individual Growth and Developmental Indicators for progress monitoring and intervention decision-making in early education. *Exceptionality, 16*(1), 33–47.

Whitaker, L. H. (2013). Implementation of response to intervention (RTI) and Individual Growth Development Indicators (IGDIS): The effects on emergent literacy for preschoolers with and without disabilities. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 74*(1–A(E)).

Presentations and Technical Reports

Bradfield, T. A., Besner, A., Wackerle-Hollman, A. K., Schmitt, B., McConnell, S. R., & Jones, K. (2010). Moving beyond picturing naming: The next generation of oral language IGDIs. Presented at the 7th Biennial Conference on Research Innovations in Early Intervention (CRIEI).

Bradfield, T. A., McConnell, S. R., & Wackerle-Hollman, A. K. (2014). Confident decision-making: Validation of an Early Childhood Response to Intervention decision-making framework. Presented at the Conference for Research Innovation in Early Intervention.

Buzhardt, J., Greenwood, C. R., McConnell, S. R., Wackerle-Hollman, A. K., & Walker, D. (2016, December). *The IGDI Enterprise: Development and use of Individual Growth and Development Indicators to support early intervention and early learning of infants, toddlers, and preschoolers.* Invited Presentation presented at the Institute of Education Sciences Project Directors' Meeting, Washington D.C.

Early Childhood Research Institute on Measuring Growth and Development. (1998a). *Research and development of individual growth and development indicators for children between birth and age eight* (No. 4). Minneapolis, MN: Center for Early Education and Development, University of Minnesota. Retrieved from http://ici2.umn.edu/ecri/ecrirpt4.pdf

Early Childhood Research Institute on Measuring Growth and Development. (1998b). *Research and development of individual growth and development indicators for children between birth and age eight* (Technical Report 4) (Technical Reports of the Early Childhood Research Institute on Measuring Growth and Development). Minneapolis, MN: University of Minnesota.

Early Childhood Research Institute on Measuring Growth and Development. (2003). Individual Growth and Development Indicators [Website]. Retrieved October 25, 2003, from www.getgotgo.net

Early Childhood Research Institute on Measuring Growth and Development. (2011). I*ndividual Growth and Development Indicators: Early Literacy.* St. Paul, MN: Early Learning Labs, Inc. Retrieved from www.myigdis.com

Jacobs, H., Wang, Q., Wackerle-Hollman, A. K., Rodriguez, M. C., & McConnell, S. R. (2016, February). *Computer adaptive testing adjustments in an early literacy assessment.* Presented at the Conference on Research Innovation in Early Intervention, San Diego, CA.

McConnell, S. R., & Kincaid, A. (2016, February). *Creating Meaningful Scale Scores for Individual Growth and Development Indicators of Early Literacy.* Presented at the Conference on Research Innovation in Early Intervention, San Diego, CA.

McConnell, S. R., McEvoy, M. A., Carta, J. J., Greenwood, C. R., Kaminski, R. A., Good, R. H. I., … Goldberg, P. (1998). *Research and development of individual growth indicators for children between birth to age eight.* (pp. 1–33). Minneapolis: University of Minnesota.

McConnell, S. R., Missall, M., Rodriguez, M., & Wackerle-Hollman, A. (2010). Monitoring progress on progress monitoring: Recent innovations in the design and evaluation Of Individual Growth And Development Indicators. Presented at the 7th biennial Conference on Research Innovations in Early Intervention (CRIEI).

McConnell, S. R., Missall, M., Rodriguez, M., & Wackerle-Hollman, A. (2010). Monitoring progress on progress monitoring: Recent innovations in the design and evaluation of Individual Growth and Development Indicators. Presented at the 7th biennial Conference on Research Innovations in Early Intervention (CRIEI).

McConnell, S. R., Rodriguez, M., Wackerle-Hollman, A., Bradfield, T., & Palma, J. (2013). Individual Growth and Development Indicators of language and early literacy development: Technical manual.

McConnell, S. R., Wackerle, A., Roverud, J., Wagner, A., Hays, A., & Missall, K. N. (2007). Recent research in early literacy assessment and intervention. Presented at the National Association of School Psychologists.

McConnell, S. R., & Wackerle-Hollman, A. K. (2015, July). *Supporting innovative assessment and data- based decision making with young children: Individual growth and development indicators—automated performance evaluation of early literacy and language.* Paper presented within a symposium: Innovative measures of language and literacy at the presented at the 22nd Annual meeting of the Society for the Scientific Study of Reading, Kona, Hawaii.

Missall, K. N., & McConnell, S. R. (2004). Technical Report: Psychometric characteristics of individual growth and development indicators–Picture Naming, Rhyming & Alliteration. *Minneapolis, MN: Center for Early Education and Development.*

Schardt, A., Brandes, D., Wackerle-Hollman, A., & McConnell, S. (2017). *Defining a language and early literacy domain for assessment of three-year-olds: Alphabet Knowledge* (Technical Report No. #2). Minneapolis, MN: University of Minnesota.

Schardt, A., Schuster, K., Brandes, D., Wackerle-Hollman, A., & McConnell, S. R. (2017). *Defining a language and early literacy domain for assessment of three-year-olds: Phonological Awareness* (Technical Report No. 1). Minneapolis, MN: University of Minnesota.

Schuster, K., Cai, L., McConnell, S., & Wackerle-Hollman, A. (2017). *Defining a domain of assessment for language and early literacy in three-year-olds: Oral Language* (Technical Report No. 3). Minneapoilis MN: University of Minnesota.

Schuster, K., Schardt, A., Brandes, D., Cai, L., Albano, A. D., Wackerle-Hollman, A., & McConnell, S. R. (2017). *Prototype data collection formats for assessing language and early literacy in young children* (Technical Report No. 4). Minneapolis, MN: University of Minnesota.

Wackerle-Hollman, A., Brunner, S., Duran, L., McConnell, S., Palma, J., Kohlmeier, T., … Rodriguez, M. (2012). *The development of early literacy skills in bilingual and Spanish-speaking preschool-age children: A literature review.* Minneapolis MN: University of Minnesota.

Wackerle-Hollman, A., Duran, L., Rodriguez, M. C., Palma, J., & Brunner, S. (2014). Technical Report 3: Evaluating S-IGDI Measures: Iterative Decision-Making in the Development Process.

Wackerle-Hollman, A. K., & McConnell, S. R. (2007). General Outcome Measures of Early Literacy Development: Evidence of Validity and Reliability Across Ages. Presented at the National Association of School Psychologists.

Wackerle-Hollman, A., Schmidt, B., Bradfield, T. A., & McConnell, S. R. (2010). How do we best measuring phonemic awareness? The development of new Individual Growth and Development Indicators (IGDIs). Presented at the 7th Biennial Conference on Research Innovations in Early Intervention (CRIEI).

Wagner, A., & Davis, S. (2007). Individual Growth and Development Indicators (IGDIs) as a Basis for Projected Growth in an Early Literacy Intervention. Presented at the National Association of School Psychologists.

Score Definitions

Up to three different scores are visible to users of myIGDIs. When assessment is administered with the seasonal card sets, users will see Cards, Scale Scores, and Tier Assignments. iPad users will see only Scale Scores and Tier Assignments. (Note that all Scale Scores and Tier Assignments are based on Rasch scoring. However, due to the rather technical and nonintuitive features of these Rasch values, they are typically not presented in user materials.)

Cards

Seasonal screening with cards is based on 15 items, selected to optimize evaluation of child performance relative to the benchmark for that season. During administration, examiners note the number of errors the child produces in this standard 15-item set, and typically enters the total correct items per measure into the myIGDIs online data system. Note that the difficulty of items varies for the same measure in different seasons, and for different measures in the same season. As a result, this “cards” score cannot be used to compare performance across measures or across seasons.

Scale Scores

Seasonal cards measures and all assessments using iPads are converted to IGDI Scale Scores. Scale scores currently range from approximately 40 to 60, and can be used to compare scores obtained on each measure, including a single child’s performance across time (e.g., are scores increasing?) or different children’s performance at the same time (e.g., which students in my class are scoring relatively lower?).

Tier Assignments

As described in the earlier section of this manual on Development of IGDIs, child performance is summarized into one of three levels, or tiers for differentiated intervention. Children scoring reliably above an empirically identified cut score are considered candidates for Tier I intervention; this assignment is noted by a green circle in the classroom summary report and indicates that the child is making good progress toward a long-term goal and is a good candidate to continue receiving the general classroom intervention.

Children scoring reliably below that empirically identified cut score are considered candidates for Tier II/III intervention; this assignment is noted by a red circle in the classroom summary report and indicates that assessment suggests the child is both not making adequate progress and is likely a good candidate for supplemental or intensive intervention. Finally, children whose score is neither reliably above nor reliably below the empirical cut score are noted by an orange circle in the classroom summary report; this indicates that more information is needed, typically from the teacher or other specialists, to best determine the appropriate level of intervention for that child.

Conversion Tables

Two conversion tables are presented below. The first presents information for assigning students assessed with cards to one of three tier assignments, and the second documents those scale scores used for tier assignment when assessed with the computer-adaptive version completed with iPads.

Table 11: Screening Benchmarks for Card Administrations

Tier		Fall			Winter			Spring		
		II/III	Cut	I	II/III	Cut	I	II/III	Cut	I
Picture Naming	Cards	0–5	6–10	11–15	0–5	6–10	11–15	0–5	6–10	11–15
	Scale Score	40–46	47–49	50–55	40–46	47–49	50–56	40–47	48–50	51–56
Rhyming	Cards	0–6	7–11	12–15	0–5	6–10	11–15	0–7	8–12	13–15
	Scale Score	37–44	45–47	48–53	39–45	46–48	49–55	39–47	48–50	51–55
Sound ID	Cards	0–6	7–11	12–15	0–5	6–10	11–15	0–9	10–13	14–15
	Scale Score	39–46	47–49	50–55	41–48	49–51	52–57	42–50	51–54	55–57
Which One Doesn’t Belong	Cards	0–5	6–10	11–15	0–5	6–10	11–15	0–5	6–10	11–15
	Scale Score	40–46	47–49	50–55	40–47	48–49	50–56	41–47	48–50	51–57
Alliteration	Cards	n/a	n/a	n/a	0–5	6–10	11–15	0–5	6–10	11–15
	Scale Score	n/a	n/a	n/a	38–45	46–47	48–54	38–45	46–48	49–54

Table 12: Screening Benchmarks for iPad Administrations

Tier	Fall			Winter			Spring		
	II/III	Cut	I	II/III	Cut	I	II/III	Cut	I
Picture Naming	30–46	47–49	50–62	31–47	48–50	51–61	34–47	48–49	50–59
Rhyming	36–45	46–47	48–55	38–45	46–47	48–55	36–47	48–49	50–54
Sound ID	37–45	46–47	48–58	38–48	49–50	51–57	40–50	51–52	53–55
Which One Doesn’t Belong	36–46	47–49	50–59	39–47	48–49	50–57	37–48	49–50	51–57
Alliteration	n/a	n/a	n/a	37–48	49–51	52–54	35–48	49–51	52–55

- **Tier I:** Strong Progress—scores in this range indicate with confidence that the child is making adequate progress with the intervention currently provided.
- **Cut Range:** More Information Needed—scores in this range indicate that more information is needed to determine the most appropriate level of intervention, and Tier Status, for the child.
- **Tier II/III:** At-risk Progress—scores in this range indicate with confidence the child is not making necessary progress and will likely benefit from further instructional/intervention support. Other information and teacher judgment will help determine whether Tier 2 or Tier 3 services are most appropriate.

Table 13: Scale Score Conversions

Assessment	Season	Card Count	Scale Score	Rasch Score
Picture Naming	Fall	0	40	−3.14
		1	42	−1.89
		2	44	−1.12
		3	45	−0.64
		4	45	−0.26
		5	46	0.06
		6	47	0.35
		7	47	0.62
		8	48	0.89
		9	48	1.16
		10	49	1.45
		11	50	1.77
		12	50	2.15
		13	51	2.63
		14	53	3.4
		15	55	4.65
	Winter	0	40	−3.01
		1	42	−1.76
		2	44	−0.99
		3	45	−0.5
		4	46	−0.13

Assessment	Season	Card Count	Scale Score	Rasch Score
Picture Naming (continued)	Winter (continued)	5	46	0.19
		6	47	0.48
		7	47	0.76
		8	48	1.02
		9	49	1.3
		10	49	1.59
		11	50	1.91
		12	51	2.28
		13	52	2.77
		14	53	3.54
		15	56	4.8
	Spring	0	40	−2.77
		1	43	−1.51
		2	44	−0.74
		3	45	−0.25
		4	46	0.12
		5	47	0.45
		6	47	0.74
		7	48	1.01
		8	49	1.28
		9	49	1.56
		10	50	1.85
		11	50	2.17
		12	51	2.55
		13	52	3.04
		14	54	3.81
		15	56	5.06
Rhyming	Fall	0	37	−3.86
		1	40	−2.61
		2	41	−1.84
		3	42	−1.35
		4	43	−0.98

Assessment	Season	Card Count	Scale Score	Rasch Score
Rhyming (continued)	Fall (continued)	5	44	−0.66
		6	44	−0.37
		7	45	−0.1
		8	45	0.17
		9	46	0.44
		10	46	0.73
		11	47	1.05
		12	48	1.42
		13	49	1.91
		14	50	2.68
		15	53	3.93
	Winter	0	39	−2.99
		1	41	−1.73
		2	43	−0.96
		3	44	−0.47
		4	45	−0.09
		5	45	0.23
		6	46	0.52
		7	46	0.79
		8	47	1.07
		9	48	1.34
		10	48	1.63
		11	49	1.95
		12	50	2.33
		13	50	2.82
		14	52	3.59
		15	55	4.85
	Spring	0	39	−2.91
		1	42	−1.65
		2	43	−0.88
		3	44	−0.39
		4	45	−0.01
		5	45	0.31

Assessment	Season	Card Count	Scale Score	Rasch Score
Rhyming (continued)	Spring (continued)	6	46	0.6
		7	47	0.88
		8	47	1.15
		9	48	1.43
		10	48	1.72
		11	49	2.04
		12	50	2.42
		13	51	2.91
		14	52	3.69
		15	55	4.94
Sound ID	Fall	0	39	–3.98
		1	41	–2.72
		2	43	–1.96
		3	44	–1.47
		4	45	–1.09
		5	45	–0.78
		6	46	–0.49
		7	46	–0.22
		8	47	0.05
		9	48	0.32
		10	48	0.61
		11	49	0.93
		12	50	1.3
		13	50	1.79
		14	52	2.56
		15	55	3.81
	Winter	0	41	–2.83
		1	44	–1.58
		2	45	–0.8
		3	46	–0.31
		4	47	0.07
		5	48	0.39

Assessment	Season	Card Count	Scale Score	Rasch Score
Sound ID (continued)	Winter (continued)	6	48	0.68
		7	49	0.96
		8	49	1.23
		9	50	1.51
		10	51	1.8
		11	51	2.13
		12	52	2.51
		13	53	3.01
		14	54	3.78
		15	57	5.04
	Spring	0	42	−2.7
		1	44	−1.45
		2	46	−0.67
		3	47	−0.18
		4	47	0.2
		5	48	0.52
		6	49	0.82
		7	49	1.1
		8	50	1.37
		9	50	1.65
		10	51	1.95
		11	51	2.28
		12	52	2.67
		13	53	3.16
		14	55	3.94
		15	57	5.2
Which One Doesn't Belong	Fall	0	40	−3.98
		1	42	−2.73
		2	44	−1.96
		3	45	−1.47
		4	45	−1.1
		5	46	−0.78
		6	47	−0.49

Assessment	Season	Card Count	Scale Score	Rasch Score
Which One Doesn't Belong (continued)	Fall (continued)	7	47	−0.22
		8	48	0.05
		9	48	0.32
		10	49	0.61
		11	50	0.93
		12	50	1.31
		13	51	1.79
		14	53	2.56
		15	55	3.82
	Winter	0	40	−3.74
		1	43	−2.49
		2	44	−1.72
		3	45	−1.23
		4	46	−0.86
		5	47	−0.54
		6	47	−0.25
		7	48	0.02
		8	48	0.29
		9	49	0.56
		10	49	0.85
		11	50	1.17
		12	51	1.54
		13	52	2.03
		14	53	2.8
		15	56	4.05
	Spring	0	41	−3.28
		1	44	−2.03
		2	45	−1.26
		3	46	−0.77
		4	47	−0.4
		5	47	−0.08
		6	48	0.21

Assessment	Season	Card Count	Scale Score	Rasch Score
Which One Doesn't Belong (continued)	Spring (continued)	7	49	0.49
		8	49	0.76
		9	50	1.03
		10	50	1.32
		11	51	1.64
		12	52	2.02
		13	53	2.51
		14	54	3.28
		15	57	4.53
Alliteration	Winter	0	38	−3.1
		1	40	−1.84
		2	42	−1.07
		3	43	−0.58
		4	44	−0.21
		5	44	0.11
		6	45	0.41
		7	45	0.68
		8	46	0.95
		9	47	1.22
		10	47	1.51
		11	48	1.84
		12	49	2.21
		13	50	2.7
		14	51	3.47
		15	54	4.73
	Spring	0	38	−2.93
		1	41	−1.66
		2	42	−0.87
		3	43	−0.36
		4	44	0.03
		5	45	0.36
		6	45	0.66

Assessment	Season	Card Count	Scale Score	Rasch Score
Alliteration (continued)	Spring (continued)	7	46	0.95
		8	47	1.23
		9	47	1.51
		10	48	1.81
		11	48	2.13
		12	49	2.52
		13	50	3.01
		14	52	3.79
		15	54	5.05

About the myIGDIs Authors

Dr. Scott McConnell earned his PhD at the University of Oregon and is currently a Professor of Educational Psychology and Child Psychology at the University of Minnesota. In 2008–2009, Dr. McConnell served as the Fesler-Lampert Chair in Urban and Regional Affairs and also devotes time to the University's Northside Initiative. In addition, Dr. McConnell also led Minnesota's efforts in collaboration between the University of Kansas, Ohio State University, and Dynamic Measurement Group (DMG) called the Center on Response to Intervention in Early Childhood (CRtIEC).

Dr. Tracy Bradfield was a research associate with the Center for Early Education and Development (CEED) at the University of Minnesota, with specific responsibility for project coordination and management of activities of the Center for Response to Intervention in Early Childhood (CRtIEC), including research activities to support ongoing Individual Growth and Development Indicator (IGDI) measure development and development of a decision-making framework to support score interpretation.

Dr. Alisha Wackerle-Hollman is a research assistant professor of Educational Psychology at the University of Minnesota. Dr. Wackerle-Hollman currently holds the role of project coordinator for the Center for Response to Intervention in Early Childhood. Dr. Wackerle-Hollman has contributed to IGDI development since 2005 and currently works on the research and development of future measures, including identification and progress monitoring measure development.

Dr. Michael Rodriguez is a professor in the Department of Educational Psychology at the University of Minnesota. Dr. Rodriguez's research focus is on understanding the psychometric properties of tests. This work has included research on the effects of item formats and the use of constructed-response versus multiple-choice items. Dr. Rodriguez currently provides advisory support to the National Board of Professional Teaching Standards, the National Assessment of Educational Progress (ETS), the US Defense Department on military personnel testing, the Association of American Medical Colleges, and the Buros Center for Testing.

Index

About Renaissance

We create assessment and practice solutions that put learning analytics to work for educators, saving hours of prep time while making truly personalized learning possible. Schools nationwide use our solutions to analyze students' abilities and guide high-quality instruction. We help teachers teach better, students learn better, and school administrators lead better—all to improve academic outcomes.

23Jun2020