Task Development and Item Analyses in Innovative Measures of Early Literacy

Alisha Wackerle-Hollman, Tracy Bradfield, Scott McConnell,

Anthony D. Albano & Michael C. Rodriguez

University of Minnesota

Task Development and Item Analyses in Innovative Measures of Early Literacy

Early literacy assessment with preschool populations offers resources for educational professionals to make knowledgeable decisions regarding instruction, intervention and early literacy environmental stimuli. Current assessment models available to early childhood professionals offer a wide array of forms and stylistic approaches to measuring preschool early literacy skills. Specifically, currently available assessments can offer information regarding how instruction may improve students' early literacy skills through mastery of concepts measures (e.g. Unit Tests), how early literacy intervention may benefit students with limited skills through intervention effectiveness measures (e.g. IGDIs, DIBELS), and how early literacy environments may be enriched to support attainment of related skills through general classroom environment measures (e.g. CLASS).

Current emphasis in the field of early literacy suggests that preschool classrooms are improving how they teach students early literacy skills by examining how students respond to classroom instruction and instruction. This approach, termed Response to Intervention (RTI), assesses student skills at an individual level and evaluates individual skill in response to universal instruction (tier 1) and targeted (tier 2) or intensive interventions (tier 3). Few measures are currently available that meet the needs of this systemic and dynamic approach, leaving teachers with limited resources for evaluating the fluid process of applying, increasing, and decreasing intervention intensity according to each student's individual need. An assessment system is needed that can identify skill level and related level of intervention (e.g. tier 1, tier 2 or tier 3), monitor progress while receiving intervention within the tier, and provide transitional assessments that can capture the fluid nature of an RTI system, including transitions between tier levels.

To address this current need in the field the Center for Response to Intervention in Early Childhood (CRTIEC) has taken on the challenge of creating a systemic assessment model for use within an RTI system. The model will have two components, first an identification component that can identify and isolate skill levels related to tier level services, and second, a progress monitoring component that can measure individual student progress while receiving tier level services. The purpose of this paper is to present the development of the measures for use in the identification component, the criteria used to evaluate each task, and the resulting revision process for improving each task, including related analyses that focus on correlational and descriptive statistics.

Theoretical Perspectives

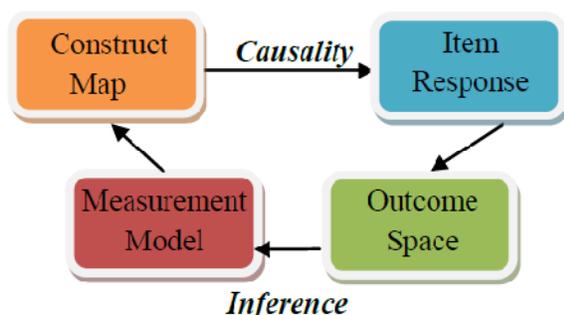To create tasks with utility for use in an RTI model, the measures must be psychometrically sound and grounded in construct validity (see related paper: *Construct Identification to Support Early Literacy Assessment,* Bradfield, Wackerle-Hollman, McConnell, Albano & Rodriguez, 2011). To ensure the tasks are a direct manifestation of the construct a measurement framework was employed. Wilson's item response modeling approach (2005) provides an innovative way to address early literacy assessment systems by creating an assessment to meet the needs of the data, rather than forcing the data to fit an existing

Figure 1. The four building blocks of an item response model approach to measurement construction (Wilson, 2005).

model (Wilson, 2003). Wilson's model encompasses four building blocks, including the *construct map, item response, outcome space,* and *measurement model* (see Figure 1).

The *construct map* represents the instrument as a manifestation of the construct of interest. The construct, early literacy, is defined through a construct map that according to Wilson includes three critical features: (a) a coherent and substantive definition for the content of the construct; (b) an underlying continuum which allows for comparisons to be made among the attributed content of the scale, or level of a trait, as more, equal, or less and (c) a "mapping" of the continuum in terms of the responses to the items, suggesting an order of item response features (Masters, Adams & Wilson, 1990; Wilson, 2003).

The *item response* is a real world translation of the construct to an item and is the focus of this paper. Because many levels of the construct of early literacy can be conceived, many items are constructed to represent different levels of the trait (Wilson, 2003). In this way, the items are realizations of the construct and the construct, as it is manifested in the individual, causes the item response.

Each item generates a response from the student that is classified into a specified category in the *outcome space* (Wilson, 2003). By specifying categories of the outcome space the characteristics of the response that are valued are defined. The outcome space also specifies the scoring rules acknowledging construct-relevant features of responses (Wilson, 2005).

Finally, to appropriately design an instrument, a *measurement model* that relates scores to constructs must be selected (Wilson, 2005). The measurement model must provide a methodology to use the information about students and their responses in a way that allows for coding in the outcome space to inform the location of the students and their responses on the construct map (Wilson, 2003). For the purposes of the assessment model described here the Rasch (Rasch, 1960) Item Response Theory (IRT) model was selected.

Measure Development

Initial construction of the measurement model was grounded in an extensive period of conceptualization through literature reviews, existing measure review, and emphasis of the first component of Wilson's model, demonstrating and supporting the construct as describe in the previous paper (i.e. *Construct Identification to Support Early Literacy Assessment;* Bradfield, et al., 2011*)*. Research was classified into four domains of early literacy including oral language, phonological awareness, alphabet knowledge and comprehension as previously described. To best define the domains of early literacy relevant within an RTI model an extensive literature review was completed. Results provided literature to construct operational definitions and identification of the relevant features important for capturing representative skills when creating an assessment tool through a comprehensive measure review. Based on the results of the literature and existing measure review, and the lack of availability of current measures that demonstrate utility in an RTI model to bridge the gap between early childhood and reading readiness at kindergarten, the creation of new tasks and items was most appropriate. As a result, new Individual Growth and Development Indicators (IGDIs 2.0) were created. To begin the process, Wilson's model, including the four component parts of measure creation (*Construct Map, Item Response, Outcome Space* and *Measurement Model*), was used, with description of item development reflecting the item response. New tasks were created that (a) capture each domain of early literacy in its entirety; (b) represent typical behaviors of children in preschool environments, and (c) are brief, engaging, and easy to interpret.

Developing Item Response and Outcome Space

To construct new items the measurement team consulted existing early literacy resources and content experts to maintain construct validity within the item sets. Similarly, to ensure that the content of items in each of the tasks was representative of the four domains and was of an appropriate level for preschool age students' lexicon, four methods of selection were used. First, words that appeared on three existing lists of oral language and vocabulary: *The Lexical Development Norms for Young Children* (The MacArthur-Bates communicative inventory, Dale & Fenson, 1996), *The Social World of Children: Learning to Talk* (Hart & Risley, 1999) and the Kindergarten list from *Words Worth Teaching* (Biemiller, 2009) were entered into a potential pool of word candidates. Words that appeared on all three lists were used in an effort to ensure the words were likely to be in a preschooler's vocabulary. Second, a review of eight published preschool curricula (see Kaminski et al. 2009 for a complete list of the review and process) yielded a list of target words presented during the pre-kindergarten year. These words were then cross-referenced with the candidate words from the existing published word lists, yielding a composite list of words represented in both common vocabulary assessments and word lists, as well as within widely published and accepted curricula. These words were then chosen to be applied in each measure developed for testing.

Oral language. To capture the skills represented in the domain of oral language, five tasks were created representing both expressive and receptive language ability. Research demonstrates oral language can be accurately assessed by using both expressive and receptive measures, accessing skills representative of both how students *produce* and *understand* words and vocabulary in their environment.

Expressive language tasks. Picture naming. During this task, modeled after the IGDI 1.0 task of the same name, a child is presented with images of common objects (one at a time) and asked to identify the item verbally, as quickly as possible. *Picture Naming* is a production task with student score recorded as the number of items correct in 1 minute.

Expressive categories: During *Expressive Categories* the child is given a category name (e.g. animals), and asked to verbally list items that fit within the identified category as quickly as possible. Initial development determined no time limit would be placed on this task until piloting for feasibility was completed. *Expressive Categories* is a production task with student score recorded as the number of items produced correctly representing the category suggested.

Receptive language tasks. *Point-to-picture*. During *Point-to-Picture* the child is asked to correctly identify a verbally presented prompt by pointing to the picture that represents the target word spoken by the administrator. *Point-to-Picture* has a multiple-choice format with student score recorded as the number of items correct in 1 minute.

Motor instructions. During this task the child performs a motor movement described in by the administrator (e.g. "touch your nose"). *Motor Instruction* is a production task with student score recorded as the total number of items correct in 1 minute.

Which one doesn't belong? During this task the child points to, or verbally identifies, the picture that doesn't fit categorically with two additional distracter items presented on a card. *Which One Doesn't Belong*? is a multiple-choice format task with student score recorded as the total number of items correct in 2 minutes. *Which One Doesn't Belong*? was originally developed as an oral language task, but after initial analyses was determined to be more accurately representative of comprehension.

Definitional vocabulary with and without pictures. During the *Definitional Vocabulary without Pictures* task the child is asked to confirm a definition of a common vocabulary word. For example, "Is the sun hot or cold?" In *Definitional Vocabulary with Pictures* the same questions are asked, however the child is provided with an image of the vocabulary word (e.g. "sun"). The image is included in this task to reduce the cognitive load required to answer the question by alleviating the child's need to remember the target vocabulary word. Both *Definitional Vocabulary* tasks are two-choice format tasks with student score recorded as the total number of items correct in 1 minute.

Picture Naming, *Point to Picture*, *Which One Doesn't Belong?* and *Definitional Vocabulary with Pictures* were all presented using 8" by 5.5" cards depicting a target stimulus and in some cases two distracters presented horizontally on each card. Each task included 12 items, with the exception of *Motor Instruction* which included 16. *Motor Instruction* items were written on cards, however they were provided simply for assessor convenience - the child did not interact with the cards. *Definitional Vocabulary without Pictures* tasks were presented verbally.

Phonological awareness. To capture the skills represented in the domain of phonological awareness four tasks were created representing detection and manipulation of sounds. Research suggests phonological awareness is most easily captured at the word and syllable level, advancing in complexity to rime and phoneme.

Detection of sounds. *Rhyming***:** During *Rhyming* the child identifies a word (via pointing to an image) that rhymes with a target word, given three choices illustrated pictorially. *Rhyming* is a multiple choice format task with student score recorded as the number correct in 2 minutes.

*Alliteration***:** During this task the child identifies a word that starts with the same sound as the target word, given three choices illustrated pictorially. Each item includes an example alliterative emphasizing a name or adjective that starts with the same initial sound as the target (e.g. Dan the Dog). Whenever possible, the target picture's name or adjective was monosyllabic. *Alliteration* is a multiple-choice format task with student score recorded as the number correct in 2 minutes.

Manipulation of sounds. *Sound blending***.** *Sound Blending* includes a continuum of word, syllable and phoneme blends (Cadigan, 2008). The child is asked to produce a word given a prompt, including word, syllable, or phoneme components. For example, the administrator may provide the syllables "/wa/ /ter/" and the child is expected to produce the word "water." *Sound Blending* is a production task with student score recorded as the number correct in 2 minutes.

*Syllable segmentation***.** *Syllable Segmentation* requires the child to correctly clap the syllable pattern of 2, 3, or 4-syllable words, presented verbally. For example, the word "elephant" is intended to elicit 3 syllables, represented by three consecutive claps produced by the child. *Syllable Segmentation* is a production task with student score recorded as the number correct in 2 minutes.

Rhyming and *Alliteration* were presented using 8" by 5.5" cards depicting a target stimulus and three distracters presented horizontally on each card. Both tasks included 45 items, 10 of which were used during pilot testing reported here. *Syllable Segmentation* and *Sound Blending* were presented verbally. Each task included 20 items.

Alphabet knowledge. To capture the skills represented in the domain of alphabet knowledge five tasks were created representing identification of both orthographic units (letters) and phonological units (sounds). The alphabet knowledge tasks were developed to assess children's general knowledge about the names and sounds of letters and their correspondence by providing letters or sounds in isolation and with distracter letters and symbols.

Orthographic tasks. *Letter identification*. During this task the child is asked to identify a letter by name given three choices including other identifiable letters. *Letter Identification* is a multiple-choice format task and student score is recorded as the number correct in 1 minute.

Letter naming. During *Letter Naming* the child is asked to produce a specific letter by name when provided with the prompt "What letter is this?" and simultaneously given the image of the letter presented in isolation. Letter Naming is a production task with student score recorded as the number correct in 1 minute.

Letter orientation. During this task the child is asked to identify a letter by name, given three choices, including rotations, symbols, and shape distracters. Symbols included numerals and mathematic symbols (e.g. Ø, ß, and ∞). *Letter Orientation* is a multiple-choice format intended to capture a child's familiarity with the appearance of the letters of the alphabet. Student score is recorded as the number correct in 1 minute.

Phonological tasks. *Sound identification*. During *Sound Identification* the child is presented with three letters and asked to identify which letter corresponds with a letter sound provided by the administrator (e.g. "Point to the letter that makes the /s/ sound"). *Sound Identification* is a multiple-choice format that intends to capture a student's knowledge of which letter symbol has a specific sound.

Sound naming: During *Sound Naming* the child is asked to identify a specific sound of a letter by producing the sound when provided with the image of a letter presented in isolation and the prompt: "What sound does this letter make?" *Sound Naming* is a production task that intends to capture a student's knowledge of the sound related to each letter. Student score is recorded as the number correct in 1 minute.

The *Letter Identification, Letter Naming, Letter Orientation, Sound Identification* and *Sound Naming* tasks were presented using 8" by 5.5" cards. Each set, with the exception of Letter Naming, included between 50 and 52 items, however for the purposes of this study 10 items were presented. Item content included all 26 letters of the alphabet and corresponding distracters presented in Zaner-Bloser font, when appropriate. Each card featured letters positioned centrally and horizontally in either upper or lowercase font. Lowercase target symbols were considered more difficult in ability level than uppercase targets (Adams, 1990). Distracter symbols were of three categories including letters that look very similar (e.g. e, o, and a), letters that look very different (e.g. "B", "Z", and "U") and cards where one distracter is similar to the target letter symbol and one that looks different (e.g. "F", "A", and "D"). For tasks utilizing letters in isolation, the lowercase letter symbols "l" and "t" were excluded because the symbols were not easily identifiable (e.g. the lowercase l looks similar to the uppercase I or the numeral 1).

Comprehension. To capture the skills related to the form of comprehension identified through the construct development process, six tasks were created that incorporated both inferencing and understanding concepts and ideas through basic comprehension. The comprehension tasks were developed to assess general knowledge about concepts and ideas through listening and picture/text comprehension, as well as added extension of thoughts through "wh" questions characteristic of inferencing.

Inferencing tasks. *Picture comprehension*. During this task children were presented with a dynamic illustration and asked a relevant question about the context of the image. For example, the illustration might contain a beach, a girl in a swimsuit, or a bucket and shovel and a towel. The child would be asked, "Where is she going?" Correct responses included statements such as: swimming, the beach, etc. Student score is recorded as the number correct in a 2 minute period.

Sequencing. During *Sequencing* the child is presented with a target illustration featuring an activity. The child is asked to examine the target illustration. On the same testing card two response illustrations are provided. After examining the target illustration the child is asked to select which of the two response illustration happens next. For example, the child may first be presented with an image of a boy fishing. Response illustrations might include the boy putting fish in a cooler and the boy preparing his fishing gear. Students are expected to sequence the activities by selecting what comes next. In this case, the boy would put the fish in a cooler after capturing them while fishing. Student score is recorded as the number correct in a 2 minute period.

Story titles task. During this task the administrator reads a brief story presented in rhyming poem format to the child. Children are then given three potential titles (verbally presented) and asked to select the title that best represents the story provided. Student score is recorded as the number correct in a 2 minute period. The *Story Titles Task* intends to capture comprehension by asking the child to make an inference about the nature of the story.

Basic comprehension tasks. *Sentence comprehension*. During the *Sentence Comprehension* task children are presented with a card featuring three illustrations. One of the illustrations accurately represents a sentence presented verbally and the two remaining images function as distracters, one with maximally different features and one with minimally different features. The child is asked to select the illustration that matches the target sentence. For example, a verbally presented prompt, "The rabbit is jumping over the fence," is paired with three illustrations, one of the rabbit jumping over the fence, one of the rabbit sitting near the fence and one with a cat far away from the fence. *Sentence Comprehension* captures comprehension skills by evaluating students' understanding of vocabulary words, including prepositional phrases and adjectives as well as syntactic sentence structure. It also taps a child's text comprehension skills as the child needs to examine the information represented pictorially in the image to match it to the information presented aloud by the administrator. Student score is recorded as the number correct in a 2 minute period.

Which one doesn't belong? The "*Which One Doesn't Belong?*" task was previously described in the Oral Language section of this paper. Transition to the comprehension domain occurred after analysis of relation with criterion measures demonstrated higher relations between this measure and comprehension tasks compared to oral language tasks.

Story comprehension. During this task children are asked to listen to a brief story available to general preschool populations through public media resources (e.g. bookstores, online, etc.) such as *The Carrot Seed,* or *The Birthday Box.* The administrator reads the book and asks the child a series of brief questions about the story, including questions about main characters, activities, and setting events. *Story Comprehension* score is recorded as the number of correct responses. This measure was not timed.

Sentence Comprehension and *Sequencing* were presented using 8.5" by 11" cards depicting a target stimulus and two distracters presented on each card. Similarly, *Picture Comprehension* was presented using 8.5" by 11" cards depicting a target stimulus, however no distracters were provided. All three tasks included 20 items used during pilot testing. *Story Comprehension* and the *Story Titles Task* were presented verbally and each task included 20 items.

Taken together, the task and item descriptions provided here support improved methodology for building new measures and promote further testing to provide new early literacy measures applicable within an RTI framework. Initial field testing of tasks with small cohorts (n =

50) of four-year old children yielded results indicating the following tasks may demonstrate promise in assessing preschool-age student's early literacy skills:

- Oral Language: *Picture Naming*, *Definitional Vocabulary*
- Phonological Awareness: *Rhyming*, *Alliteration*, *Sound Blending*
- Comprehension: *Sentence Comprehension*, *Which One Doesn't Belong?*
- Alphabet Knowledge: *Letter Identification*, *Letter Orientation*, *Letter Sounds*

Qualitative results indicated that these measures could be improved at the item level. As a result, item level revisions were carried out before testing with preschool age children.

Item Level Revisions

Following field testing each task was considered for item level revisions. To improve each task the research team first isolated the construct relevant content of each item. For example, within *Picture Naming* each item was evaluated for clarity, isolated content, and background noise. Because *Picture Naming* intends to capture knowledge of vocabulary, only isolated images of vocabulary words were provided. As a result any image that was not available in isolation on a white background was not used. Similarly, because vocabulary represents real world representations of a child's knowledge based on interaction with contents in their environment, all images used were real life photographs rather than illustrations or other renditions (e.g. clip art). Any image that could not clearly be identified as the target vocabulary word due to image quality was removed and any image that featured a word with an ambiguous target word was removed (e.g. glasses). Figure 2 presents before and after images of *Picture Naming* task items illustrating changes as a result of item level revisions. In the first example, "elephant" illustrates a revision to content as the original elephant (labeled "before") was a toy elephant that in real life interaction is fuzzy and small, which are not characteristics typical of a real elephant. The revised elephant (labeled as "after") features a true-to–life photograph of an elephant to reduce confusion and distortion of the concept of "elephant". The second example, "car" illustrates a revision to the content as the original car was a hand-drawn representation, and the revised car features a true-to-life photograph of a car. Finally, the third example, "glue" illustrates a revision of the background to limit distracting factors within the items, as the original glue was on a colored background, while the revised glue has been isolated on a white background for clarity and construct relevance.

Similarly, tasks with prompts were also evaluated and improved for clarity and interpretation. For example, for *Rhyming* the original prompt included the phrase: "Point to the one that sounds the same as (target word)." Because "sounds the same" can be confusing to the child as both *Alliteration* and *Rhyming* tasks are related to sounds, the prompt was changed to "Point to the one that rhymes with (target word)." Additionally, the original *Rhyming* task featured three choices, however the improved measure features two choices that are then provided both verbally and pictorially to the child. As with *Picture Naming,* all items were reviewed for content, including the removal of background images (see Figure 3). As illustrated, the "rat" was replaced with a photograph of a real life rat.

To further improve item level construct relevance, all potential features of items that could potentially contaminate the task were removed. For example, in *Rhyming,* any distracter that was alliterative with the target word was removed. Similarly other characteristics that were controlled for included color scheme (e.g. specific attention was given to the color of the target image and distracters to ensure that color of image did not influence the child's response, see Figure 3) and category (e.g. specific attention was given to whether the target image and either of the distracters represented a similar category of objects). As an example, the item cat, rat, phone

was not included because rat and cat were rhyming words, but also were both categories of animals; instead the image is displayed as rat, hat and phone; see Figure 3).

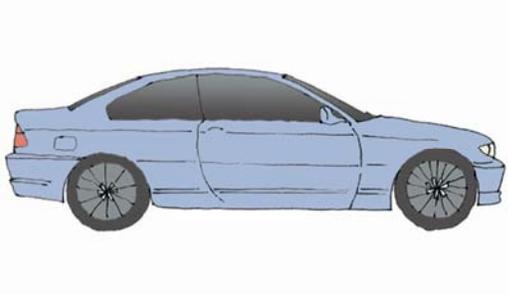Figure 2. *Before and After Item Level Revision Examples for Picture Naming*

Before Revisions	After Revisions
Picture Naming: Elephant	

Picture Naming: Car	

Picture Naming: Glue	

Figure 3: *Item Level Revisions for Rhyming*

Before Revisions	After Revisions
Rhyming: Rat, hat, phone Controlling for color, content, category and background	

Once item level revisions were completed, items were assembled into testing cadres and presented to children for validation including results yielding Rasch model analysis, descriptives, reliability, and criterion-related validity. Results from the field testing trials also indicated that the comprehension measures did not perform as well as the other domains, demonstrating a significant amount of zero scores (> 25%) and general lack of student understanding of the directions. In addition, based on students' difficulty in responding to the task and items, questions emerged regarding the conceptualization of the tasks/items as appropriate forms of the construct. As a result, it was determined that additional development, revision, and field testing would be necessary before analysis of their technical properties could be done in a valid and reliable manner.

Participants

Participants for initial validation testing included a cross sample of 756 children and families served by preschool programs at 4 sites in the Midwest and Pacific Northwest. Children received all language and literacy measures in the fall, winter and spring of the 2009 academic year.

Additional Measures

After initial field testing the reduced cadre of IGDI 2.0 measures as suggested above was included for testing. During testing procedures item level statistics were recorded and will be reported using Rasch modeling in the paper following this document titled: *Scaling Individual Growth and Development Indicators: Early Literacy Measurement* (Albano, Rodriguez, McConnell, Wackerle-Hollman & Bradfield, 2010). In addition to the IGDI 2.0 early literacy measures previously described, standardized criterion assessments were also administered. Each domain received at least one corresponding criterion measure: oral language included the Peabody Picture Vocabulary Test-4 (PPVT-4) and the Preschool Language Scale-3 (PLS-3), Alphabet Knowledge included the Test of Preschool Early Literacy (TOPEL) print awareness subtest, DIBELS- Letter Naming Subtest, and Phonological Awareness included the TOPEL phonological awareness subtest. Standardized test were presented in fall and spring testing seasons only. For the purposes of this paper only one criterion test will be reported per domain (PPVT-4 & TOPEL) for brevity.

Results

Descriptive results by measure and wave are presented in Table 1. During testing the *Sound Blending* tasks yielded qualitative and anecdotal results suggesting students did not clearly understand the task. Similarly, preliminary wave one results showed a large percentage of students received zeros (>25%). As a result, *Sound Blending* was removed from the cadre of measures and is not presented here.

Table 1. *Descriptive Results by Measure*

Measure	Wave	N	Mean	SD	% of zeros
Picture Naming	1	691	27.44	7.99	3.6%
	2	658	29.06	7.3	1.3%
	3	690	29.81	6.73	0.5%
Definitional Vocabulary with Pictures	1	693	15.74	3.9	2.3%
	2	632	16.83	3.41	0.9%
	3	671	17.26	2.85	1.1%
WODB	1	638	11.86	4.45	9.2%
	2	601	13.95	4.18	4.4%
	3	644	14.54	4.08	3.9%
Sound Identification	1	518	10.22	5.54	14.0%
	2	524	13.26	5.37	6.0%
	3	593	13.83	5.46	7.0%
Letter Orientation	1	518	13.47	5.59	9.0%
	2	518	15.89	4.32	2.0%
	3	583	16.57	3.96	2.0%
Alliteration-Revised	1	535	9.15	3.91	28.0%
	2	575	10.13	4.31	11.0%
	3	633	10.31	4.10	11.0%
Rhyming-Revised	1	538	10.49	4.87	25.0%
	2	560	12.36	5.03	14.0%
	3	653	12.47	5.11	10.0%

Descriptive results suggest all IGDI measure demonstrate appropriate means and standard deviations, with no measures reaching significant floor or ceiling effects. Some measures achieved proportionally high levels of zero scores (*Alliteration*, *Rhyming* and *Sound Identification*) indicating limited utility with students who demonstrate low levels of the corresponding domain.

Additionally, correlation coefficients were computed between measures and with criterion assessments. Correlations are reported in Tables 2 and 3, with scores reported for fall and spring waves only. Results suggest correlations were generally moderate within domains and weak between domains, supporting both convergent and discriminant validity.

Conclusions

The measures constructed and tested were created to fill a unique need within a RTI framework in early childhood. Because these measures are pioneers in determining the best approach, method and analysis for use within an RTI model, continued revision and refinement is likely needed. As such, the measures within this paper continue to be revised and tested with varying populations to increase applicability with preschool children across the nation. The analysis provides evidence for such activities, indeed the IGDI 2.0 measure presented here suggests that most, if not all measures should be pursued for additional testing with diverse populations and some measures may benefit from further item level refinement. Albano and colleagues (2011) will present and discuss the detailed information mined from this study, most specifically the item level statistics. These statistics provide a new and unique opportunity to examine preschool performance both at a person and item level, allowing corresponding

measures to produce highly sensitive items that may inform instruction and intervention matches between student skill level as tested on the IGDI 2.0 measures and RTI tier level.

Table 2. *Inter-Measure Correlations at Fall (1) and Spring (2)*

Measures	wave	Sound ID 1	Sound ID 3	Letter Orientation 1	Letter Orientation 3	Def. Vocab. 1	Def. Vocab. 3	Picture Naming 1	Picture Naming 3	WODB 1	WODB 3	Rhyming 1	Rhyming 3	Alliteration 1	Alliteration 3
Sound ID	1	1.00													
	3	0.46	1.00												
Letter	1	0.47	0.54	1.00											
Orientation	3	0.33	0.52	0.56	1.00										
Def.	1	0.44	0.35	0.45	0.35	1.00									
Vocab.	3	0.23	0.35	0.36	0.43	0.45	1.00								
Picture	1	0.37	0.27	0.36	0.27	0.51	0.48	1.00							
Naming	3	0.34	0.34	0.36	0.35	0.49	0.56	0.76	1.00						
WODB	1	0.36	0.32	0.45	0.30	0.44	0.33	0.42	0.37	1.00					
	3	0.38	0.41	0.46	0.41	0.44	0.40	0.44	0.42	0.56	1.00				
Rhyming	1	0.35	0.34	0.40	0.29	0.37	0.27	0.34	0.36	0.44	0.40	1.00			
	3	0.41	0.42	0.40	0.33	0.36	0.30	0.37	0.39	0.45	0.39	0.61	1.00		
Alliteration	1	0.28	0.29	0.33	0.28	0.33	0.20	0.27	0.31	0.35	0.30	0.40	0.37	1.00	
	3	0.40	0.38	0.34	0.27	0.37	0.23	0.29	0.30	0.36	0.36	0.41	0.45	0.43	1.00

Table 3. *IGDI 2.0 at fall (Wave 1) and spring (Wave 3) and Standard Criterion Measures (Wave 3)*

	wave	TOPEL print knowledge	TOPEL phonological awareness	TOPEL definitional vocabulary	PPVT-4
Sound ID	1	0.61	0.40	0.69	0.45
	3	0.57	0.32	0.50	0.38
Letter Orientation	1	0.55	0.39	0.31	0.54
	3	0.38	0.22	0.44	0.34
Def. Vocab.	1	0.47	0.44	0.26	0.55
	3	0.20	0.26	0.50	0.55
Picture Naming	1	0.37	0.48	0.86	0.69
	3	0.42	0.44	0.86	0.68
WODB	1	0.50	0.48	0.46	0.52
	3	0.38	0.40	0.54	0.52
Rhyming	1	0.45	0.47	0.48	0.51
	3	0.34	0.43	0.57	0.49
Alliteration	1	0.47	0.34	0.23	0.38
	3	0.45	0.37	0.52	0.39

References

Adams, M. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.

Albano, A. D., Rodriguez, M. C., McConnell, S., Bradfield, T., & Wackerle-Hollman, A. (2011, April). *Scaling with measures of early literacy*. Paper presented at the meeting of the National Council for Measurement in Education, New Orleans, LA.

Biemiller, A. (2009). *Words worth teaching*. McGraw Hill: TX.

Bradfield, T., Wackerle-Hollman, A., McConnell, S., Rodriguez, M. C. & Albano, A. D. (2011, April). *Construct identification to support early literacy measurement*. Paper presented at the meeting of the National Council for Measurement in Education, New Orleans, LA.

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavioral Research Methods, Instruments, & Computers,* 28, 125-127. *American children*. Baltimore: Paul H. Brookes.

Hart, B. & Risley, T. (1999) *The Social World of Children: Learning to Talk.* Paul H. Brooks Publishing Co., Baltimore, MD.

Masters, G. N., Adams, R. A., & Wilson, M. (1990). Charting student progress. In T. Husen & T. N. Postlethwaite (Eds.), *International encyclopedia of education: Research and studies* (Supplementary Vol. 2, pp. 628–634). Oxford, England: Pergamon.

Rasch, G. (1960). Probabilistic models for some intelligence and attainment tests. Chicago, IL: University of Chicago Press.

Rodriguez, M. C., Albano, A. D., McConnell, S., Wackerle-Hollman, A. & Bradfield, T. (2011, April). *Standard setting with innovative measures of early literacy: Contrasting Groups*. Paper presented at the meeting of the National Council for Measurement in Education, New Orleans, LA.

Wackerle-Hollman, A., Bradfield, T., McConnell, S., Albano, A. D., & Rodriguez, M. C. (2011, April). *Task development and item analysis in innovative measures of early literacy*. Paper presented at the meeting of the National Council for Measurement in Education, New Orleans, LA.

Wilson, M. (2003). On Choosing a Model for Measuring. *Methods of Psychological Research Online*, 8(3), 1-22

Wilson, M. (2005). *Constructing measures: an item response modeling approach*. Mahwah, NJ: Lawrence Erlbaum Associates.