Published in final edited form as:

J Learn Disabil. 2011 July ; 44(4): 313–321. doi:10.1177/0022219411407861.

Comparing Two Forms of Dynamic Assessment and Traditional Assessment of Preschool Phonological Awareness

Patricia Thatcher Kantor, Richard K. Wagner, Joseph K. Torgesen, and Carol A. Rashotte
Florida State University and Florida Center for Reading Research

Abstract

The goal of the current study was to compare two forms of dynamic assessment and standard assessment of preschool children's phonological awareness. The first form of dynamic assessment was a form of scaffolding in which item formats were modified in response to an error so as to make the task easier or more explicit. The second form of dynamic assessment was direct instruction of the phonological awareness tasks. The results indicate that preschool children's phonological awareness can be assessed using standard assessment procedures, provided the items require processing units larger than the individual phoneme. No advantage was found in reliability or validity for either dynamic assessment condition relative to the standard assessment condition. Dynamic assessment does not appear to improve reliability or validity of phonological awareness assessments when preschool children are given tasks that they can perform using standard administration procedures.

Response-to-intervention (RTI) models of identification of children with reading disability have been proposed as an answer to the “wait to fail” aspect of the traditional IQ-achievement discrepancy model (Fletcher, Lyon, Fuchs, & Barnes, 2007). Because of floor effects on typical reading tests and the fact that reading instruction was not begun in earnest until first grade, a severe discrepancy between IQ and achievement rarely emerged until second grade. However, if RTI models require documented failure to respond to effective general reading instruction followed by documented failure to respond to focused intervention, it is not clear that earlier identification will occur with RTI models. Nevertheless, it is likely to be the case that a given child receives intervention sooner under an RTI model compared to a traditional model in that intervention occurs after failure to respond to effective general reading instruction, prior to ultimate identification.

Two potential solutions to the wait-to-fail aspect of both traditional and RTI models are the target of the present study. The first is using emergent measures of early literacy as a way to get below the floor effects found on typical reading tests. The second is the use of dynamic assessment to “speed up” an RTI-based identification model.

One of the major advances in our understanding of the development of early reading is that emergent literacy skills precede the development of full-blown reading (Snow, Burns, & Griffin, 1998; Whitehurst & Lonigan, 1988). For example, measures of print awareness provide an assessment of knowledge and skills that are precursors of reading. Additionally, although preschool children have difficulty with phonemic awareness tasks, they can do phonological awareness tasks that are based on units larger than individual phonemes such as compound words, syllables, and onset-rime units (Lonigan, Burgess, Anthony, & Barker, 1998; Lonigan, Wagner, Torgesen, & Rashotte, 2007; Torgesen, Wagner, Rashotte, Burgess, & Hecht, 1997). Phonological awareness tasks based on larger units show continuity with later phonemic awareness and serve as unique predictors of later decoding when given in preschool (Lonigan, Burgess, & Anthony, 2000). Similarly, measures of phonological

awareness given in kindergarten predict later reading (Anthony and Francis, 2005; Hogan, Catts, & Little, 2005; Wagner, Torgesen, & Rashotte, 1994; Wagner et al, 1997).

Given the longitudinal stability of phonological processing skills and their predictive relations with subsequent reading, tasks that assess preschool phonological processing would appear to be useful for screening children at risk for later reading failure. This is especially true if the results of early assessment can lead to early intervention that reduces risk for subsequent reading failure. Prevention of reading difficulties is likely more efficient and cost effective than remediation of reading difficulties (Berninger et al., 2002; Coyne, Kame'enui, Simmons, & Harn, 2004; Torgesen, 2000).

However, it can be difficult to assess very young children reliably and validly (Caffrey, Fuchs, & Fuchs, 2008; Gredler, 1997; Lonigan, 2006; Shepard, Kagan, & Wurtz, 1998). A potential solution to inherent difficulties in assessing preschool-age children may lie in use of dynamic assessment techniques (Lidz & Pena, 2009; Swanson & Lussier, 2001). Dynamic assessment (DA) utilizes the interaction of the examiner and examinee to obtain a measure of learning potential (Grigorenko & Sternberg, 1998). In DA testing, the examiner provides feedback to the examinee, as well as offers instruction to address examinee struggles. For example, when a student is having difficulty on a test item, an examiner may modify the format, provide additional trials, provide a useful strategy, or offer a cue, hint, or prompt (Swanson & Lussier, 2001). Another form of DA is a pretest/teach/posttest model of assessment (Grigorenko, 2009). Here, children are taught how to perform the assessment tasks directly. In this sense, DA can be conceptualized as an abbreviated form of RTI. A small set of studies have looked at the use of dynamic versus traditional assessment in young children. Spector (1992) examined use of a DA of phonemic awareness as a predictor of later reading ability. Spector found that the dynamic phoneme segmentation task given in the fall of kindergarten was a better predictor of kindergarten reading progress than the traditional measures. Spector further found that the dynamic phoneme segmentation accounted for more variance in a measure of spring phonemic awareness than any of the other phonemic awareness measures administered in the fall. Olswang and Bain (1996) investigated how well static and dynamic measures correlate with change in language production of twenty-one preschool children with specific expressive language impairment. Results showed that DA was the most highly correlated with short term language growth. Day, Englehardt, Maxwell, & Bolig (1997) looked at whether dynamic measures might improve the predictive ability of fluid and crystallized ability tasks in preschoolers. Day et al. utilized structural equation modeling to compare models of the relationships between pre-training skills, learning-ease transfer, and post-training performance. The results showed that the structural model that best fit the data included paths from pretest and learning assessments to posttests in each domain. Thus, the use of both pretests and learning assessments were better predictors of a child's ability than either alone.

In a meta-analysis of studies on DA, Swanson and Lussier (2001) found large effect sizes for DA over traditional assessment. However, the effect sizes were reduced when sample size was considered and when pre-test/post-test designs were corrected for biases related to retesting. They found that the magnitude of effect sizes was dependent on type of ability group, sample size, type of assessment, and domain of assessment. Overall, Swanson and Lussier's results suggest that DA may provide estimates of processing potential that are not found with traditional testing. Additionally, they concluded that DA training results in better performance than traditional assessment.

Caffrey et al. (2008) focused their review of DA on predictive validity. They found correlations of similar magnitude between DA and traditional tests with measures of future achievement. However, they also found research that supported DA as a tool for identifying

students who respond to intervention and research that reported that DA contributed to achievement prediction above and beyond traditional testing. Additionally, it was found that DA was a better predictor for students with disabilities rather than students without disabilities or typically developing students. Caffrey et al. concluded that DA may not be a substitute for traditional tests, but may be utilized to provide additional information.

Given the small amount of research on DA in young readers, as well as questions over whether DA represents a valid solution to addressing potential difficulties in accurately assessing this population, the goal of the current study was to compare two forms of DA and traditional assessment of preschool phonological awareness. The first form of DA was a form of scaffolding in which item formats were modified in response to an error so as to make the task easier or more explicit. The second form of DA was direct instruction of the phonological awareness tasks. We compared the reliabilities of phonological awareness tasks across the three conditions. We also compared the concurrent validity of the assessments using print awareness as the criterion variable to be predicted. It was predicted that the two forms of DA would show greater reliability and validity compared to the traditional assessment.

Method

Participants

Participants were 123 (78 male) preschool children with mean age = 4.41, ($SD = 0.66$). Participants attended nine different private preschools in a small southeastern city. The majority of the children (70.7%) were Caucasian, with 22.0% African American and 7.3% other ethnicities (primarily Hispanic and Asian). The primary language for all children was English. Socioeconomic status (SES) of the sample was estimated to be middle to upper-middle class based on the location of the preschools and that these schools were private institutions.

Measures

Phonological awareness measures: Three researcher-developed tasks were used to measure phonological awareness: elision, blending, and initial sound matching. Each task included multiple choice and free response sections. Each new task was preceded by practice trials to teach the children the task. Corrective feedback was given during these practice trials but not during the test trials. Many of the items used pictures to reduce memory load. Each item was scored 1 if the answer was correct and 0 if it was incorrect.

The *elision task* required children to say a word after dropping part of it. The task consisted of 36 multiple choice and free response items plus two practice items, with the items progressing from larger phonological units (e.g. whole words in compound nouns) to smaller units (e.g. syllables and sub-syllables). The two practice items and multiple choice items were presented verbally and with pictures. The free response items were presented verbally only. In both the multiple choice and free response items, the first six items required deleting single syllable words from a compound word to form a new word. The next six in both the multiple choice and free response required deleting a syllable or phoneme from a word to make a new word, and the final six in both required deleting a smaller syllable or single letter sound from a word to make a new word. For the multiple choice items, those presented verbally and with pictures, the examiner (a) showed the child four pictures, (b) named them while pointing to each of them (e.g. “Look at these pictures. Table. Box. Brush. Tooth.”), (c) asked the child to say the compound (e.g. “My word is toothbrush, say toothbrush.”), and then (d) asked the child to delete part of it and point to the picture that represented what word is left (e.g. “If you take away brush from the word

toothbrush, what word do have? Point to it.”). For the free response items, the child was instructed that he or she was going to answer some questions without pictures and to listen carefully (e.g. “Say tear. Now say tear without /t/.”). All pictured items were administered unless the child clearly did not understand the task. For the more difficult non-pictured, free response items, a ceiling of 6 errors in a row was imposed.

The *blending task* required children to combine word elements to form a word. The blending task consisted of 36 multiple choice and free response items plus two practice items. The two practice items and multiple choice items were presented verbally and with pictures. The free response items were presented verbally only. In both the multiple choice and free response items, the first six items required blending single syllable words to form a new compound word. The next six in both the multiple choice and free response required blending two syllables or phonemes to make a new word, and the final six in both required blending two smaller syllables or single letter sounds to make a new word. For the multiple choice items, those presented verbally and with pictures, the examiner showed the child four pictures, named them while pointing to each of them (e.g. “Look at these pictures. Cupcake. Doormat. Hotdog. Basket.”), asked the child what word would be formed if they were said together (e.g. “Listen carefully to what I say and point to the word you hear. What word do these make? Cup—Cake.”). For the free response items the child was instructed that he or she was going to answer some questions without pictures and to listen carefully. The pause between word elements was approximately three-fourths of a second for word elements that were syllables or larger units and approximately one-half second for sub-syllabic units (e.g. “What word do these sounds make? F—ox.”). The ceiling rules were identical to those described for the elision task.

The *initial sound matching* task required that children identify words that start with the same sound as the word presented. The initial sound matching task consisted of 30 multiple choice and free response items plus five practice items. Four practice items and multiple choice items were presented verbally and with pictures. One practice item and the free response items were presented verbally only. For the multiple choice items, those presented verbally and with pictures, the examiner (a) showed the child a picture, (b) named it while pointing to it (e.g. “This is a sandbox.”), (c) named and pointed to the pictures of the response items (e.g. “Lunchbox. Sandwich. Mailbox.”), and (d) asked the child which word starts with the same sound as the first word (e.g. “Which word starts with the same sound as sandbox?”). For the free response items, the child was instructed that he or she was going to answer some questions without pictures, to listen carefully, and tell the examiner a word that starts with the same sound as the word presented (e.g. “Tell me a word that starts with the same sound as ball.”). The ceiling was to give all items unless the child clearly did not know the answers.

Print awareness measure: The print awareness task required the child to demonstrate basic knowledge of print concepts, letters, and words and included multiple choice and free response items that asked the child to point to aspects of print, identify letters and words, and name and say sound sounds of specific letters. The print awareness task administered was an early version of the Print Knowledge subtest of the Test of Preschool Early Literacy (TOPEL, Lonigan, Wagner, Torgesen, & Rashotte, 2007) that included 54 items. Only the 36 items in the current published version of the TOPEL Print Knowledge subtest were used in analyses in order to make results more comparable to other studies that use the TOPEL. The TOPEL Print Knowledge subtest shows adequate reliability (internal consistency, $\alpha = .96$ for 4 year olds) and validity ($r = .77$ with the TERA-3 Alphabet subtest, $r = .55$ with the TERA-3 Reading Quotient (Test of Early Reading Ability – Third Edition, Reid, Hresko, & Hammill, 2001) and $r = .57$ with the Get Ready to Read! Screening Tool (Whitehurst & Lonigan, 2001)). For the multiple choice items, the examiner showed the child a series of

pictures (e.g. “64. 213. SMK. 589.”) and asked the child a question (e.g. “Which are letters?”). For the free response items the child was shown a picture of a letter and asked to name the letter (e.g. “What is the name of this letter?”) or asked what sound of noise the letter made (e.g. “What sound/noise does this letter make?”). Each item was scored 1 if the answer was correct and 0 if it was incorrect.

Procedure

Students were randomly assigned to one of three test administration groups: standard administration ($N = 43$), dynamic-supported administration ($N = 40$), and dynamic-instructed administration ($N = 40$). Members of the standard administration group were given the normal version of the phonological awareness measures. Members of the dynamic-supported administration group were given items with support that made them easier if an error was made. Members of the dynamic-instructed administration group were given training in how to do the phonological awareness tasks.

The design resulted in students taking phonological awareness measures three times with a one week interval between completed test administrations, with students given the Print Awareness Measure at Time 3 only. All groups were given the phonological awareness measures in the normal manner at Time 1. Depending on condition, test administration at time 2 and 3 varied across groups.

Standard administration group—Phonological awareness measures at Time 2 and 3 were given in the same manner as Time 1.

Dynamic-supported administration group—At Times 2 and 3, if the student answered correctly when first asked, he/she was given a score of 1. If the student did not answer correctly, he/she was told “Listen again” and then readministered the same item with a hint provided. A correct response to the readministered item earned a score of .5. For the elision task, the hint consisted of inserting a half-second pause at the point at which the word element to be dropped began. For the blending task, the hint consisted of reducing pause interval between word components by one-quarter second. For the initial sound matching task, the target word or sound was said with greater emphasis during the second attempt on the key word element that provided the match.

Dynamic-instructed administration group—During the week prior to the second task administration, students in this group attended three 10-15 minute sessions where they were provided instruction on the three phonological awareness tasks. The instruction was designed to increase the child's phonological awareness knowledge. A meta-analysis of 52 research studies has shown that training in phonological awareness skills is more effective in improving later reading than training which does not include manipulation of phonemes (National Reading Panel, 2000). Instruction followed a carefully prepared script. Session 1 covered “blending and elision with compound words,” session 2 covered “blending and elision with two syllable words,” and session 3 covered “first sound matching.” Instruction included visual aids, such as puzzle cards with one part of a compound word on one card (e.g. “cow”) and the other part on the other (e.g. “boy”). A puppet was used to demonstrate taking the words apart. Students in the dynamic-instructed administration group were then given the phonological awareness measures at Times 2 and 3 following the standard administration procedures.

Children were tested individually by trained examiners. Examiners also served as instructors for students in the dynamic-instructed group but did not test and instruct the same child. Testing took place in the children's preschools in a quiet location that was dedicated to

testing. Examiners were allowed to divide the testing into multiple sessions as needed depending on the child, with the battery being completed within a one-week period.

Results

Data Issues and Descriptive Statistics

Prior to analysis, the data were screened for missing and extreme values. Data from two subjects were deleted due to excessive missing data, 80 percent or greater missing, leaving standard group $n = 42$, dynamic-supported group $n = 39$, and dynamic-instructed group $n = 41$. Remaining missing data (range 0 - 6% across variables) were filled in using Schafer's (1999) NORM program to create 20 multiply imputed complete data sets.

Extreme values were determined using the median $\pm$ two interquartile range as criterion. Outliers were brought to the boundary (Tabachnik & Fidell, 2007). The following variables were adjusted: T1 initial sound matching (3 data points) and T3 blending (6 data points). Visual inspection of scatterplots revealed no bivariate outliers. Evaluation of skewness statistics indicated the presence of only minor deviations from normal distributions. Skewness statistics indicated the presence of slight ceiling effects on the blending task for the dynamic-supported group at Time 2 (skewness = -1.37 (.38)) and at Time 3 (skewness = -1.25 (.38)). There was no indication of floor effects for any of the tasks and any time point.

Means, standard deviations, and internal consistency reliabilities are reported in Table 1. A phonological awareness composite was formed by averaging z-scores for each of the three tasks based on means and standard deviations from the EM parameter estimates of the total sample at Time 1. To make sure the instruction worked as intended for the dynamic-instructed group, a manipulation check in the form of a mean comparison for each of the tasks at Time 1 compared to Time 2 was conducted. Results indicated a significant mean increase for the elision task $t [39] = 4.09, p < .01$ and blending task $t [39] = 2.27, p = .03$ and for the phonological composite $t [39] = 4.43, p < .01$. However, there was not a significant mean increase for the initial sound matching task $t [39] = 1.64, p = .11$.

Correlations between phonological awareness and print awareness measures are reported in Table 2. As noted in the tables, measures of internal consistency were generally moderate to high, $\alpha = .88$ to $.97$ for the phonological composite, $\alpha = .85$ to $.96$ for the elision task, $\alpha = .92$ to $.96$ for the blending task, $\alpha = .64$ to $.95$ for the initial sound matching task, and $\alpha = .94$ to $.96$ for the print awareness task. This was true for both standard and dynamic forms of the assessments. Furthermore, it did not appear to be the case that the dynamic forms of assessment yielded greater reliabilities than did the standard assessments.

Comparing Validity Coefficients

Significance tests for comparing validity coefficients were used to compare validity coefficients after correction for differential reliability. We did this to avoid the situation of differences in validity coefficients reflecting differential reliability of the assessments. However, because there were few substantial differences in reliabilities, the overall pattern of results is identical whether validity coefficients were corrected or not.

Two comparisons of validity coefficients were used to examine effects of dynamic assessment procedures. Because the two DA groups were given standard assessments at the first time point, effects of DA should be observable in increases of validity coefficients within groups across time. Effects of DA should also be observable in comparing validity coefficients across groups, with higher validity coefficients expected for the DA groups relative to the standard assessment group. Validity coefficients between the phonological awareness composite and print awareness measure within subjects across time are presented

in Table 3, and those between subjects across group are presented in Table 4. Differences in validity coefficients across groups were calculated to determine impact of the two forms of DA on the concurrent validity of the tasks and then a Z-value and corresponding significance was calculated for each difference. To compare validity coefficients within subjects across time, we followed Meng, Rosenthal, and Rubin's (1992) procedures for comparing correlated correlations.

Phonological awareness composite—Results indicated a significant increase in the validity coefficients of the phonological awareness composite and print awareness task from Time 1 to Time 2 for all three groups and from Time 1 to Time 3 for the dynamic-instructed group. There was a significant decrease in the validity coefficients from Time 1 to Time 3 for the standard group and from Time 2 to Time 3 for the dynamic-supported group.

In comparing the validity coefficients across groups, there was a significant difference between validity coefficients of the phonological awareness composite and print awareness measure in favor of the dynamic-supported group over the standard group at times 1 and 2. There were no significant differences between the standard and dynamic-instructed group nor between the dynamic-supported and dynamic-instructed group at any of the time points.

Elision—Results indicated a significant increase in validity coefficients for elision from Time 1 to Time 2 for the standard and dynamic-supported group, from Time 1 to Time 3 for the standard group, and from Time 2 to Time 3 for the dynamic-instructed group.

In comparing validity coefficients across group, significant differences between validity coefficients in favor of the dynamic-supported group over the standard group were found at all 3 times. In addition, a significant difference in validity coefficients was found between the dynamic-instructed group and standard group at Time 1 only in favor of the dynamic-instructed group. There were no significant differences between the dynamic-supported and dynamic-instructed groups at any time.

Blending—Results indicated a significant increase in the validity coefficients of the blending and print awareness measures at each time for the dynamic-instructed group. There was a significant decrease in the validity coefficients from Time 2 to Time 3 for the dynamic-supported group and there were no significant differences in validity coefficients for the standard group.

Comparing validity coefficients across the groups indicated no significant differences between the validity coefficients between blending and print awareness for any of the groups and any time point.

Initial sound matching—Results indicated a significant increase in the validity coefficients for initial sound matching and print awareness measures from Time 1 to Time 2 for all three groups. There was also a significant decrease in the validity coefficients from Time 2 to 3 for all of the groups, and a significant decrease from Time 1 to Time 3 for the dynamic-instructed group.

In comparing validity coefficients across groups, there was a significant difference in favor of the dynamic-instructed group at Time 1, but there were no significant differences between any of the groups at any of the other times.

In summary, the only consistent difference in validity coefficients was larger validity coefficients for the dynamic-supported group compared to the standard group. However, this difference cannot be attributable to the difference in assessment conditions because it also

existed at Time 1, the occasion on which all groups were given the measures under standard conditions.

Discussion

The goal of the current study was to compare two forms of dynamic assessment—one based on scaffolding and the other based on instruction—and traditional assessment of phonological awareness. In particular, we wanted to determine whether either form of DA improved the reliability and validity of assessments of phonological awareness relative to standard administration.

Beginning with reliability, the fact that reliabilities were as high as they were for the standard administration group does not support the view that the phonological awareness of preschool age children cannot be assessed reliably using standard task administration formats. The median reliability of the phonological composite for the standard administration group was .94. That the means of the phonological awareness tasks were well above chance, and that the inspection of distributions and comparing the magnitudes of the standard deviations to the means suggested no obvious floor effects, provide further evidence that preschool children were able to do the phonological awareness tasks we gave them. One important reason for their successful performance is that the tasks included items that required phonological processing larger than the individual phoneme, and also pictured multiple choice items that required only a pointing response.

The results did not support the idea that reliability would be better for DA compared to standard task administration. The median reliability of the phonological composite for the dynamic-supported administration group of .96 and for the dynamic-instructed administration group of .95 are comparable to the reliability of .94 previously noted for the standard administration group.

Turning to validity, our results did not support the expectation of higher validity coefficients for the two DA conditions compared to the standard administration condition. Our results are in line with the review by Caffrey et al. (2008) that found similar correlations for traditional and DAs with later measures of academic achievement. The advantage in validity coefficients for the dynamic-supported administration group over the standard administration group was evident at Time 1, when all groups were given phonological awareness tasks using standard administration. Our explanation is that our moderately-sized samples were not completely equivalent despite our random assignment of students to conditions.

Regarding implications of our results for response-to-intervention (RTI) models, we began this article by describing two potential approaches to circumventing the “wait-to-fail” nature of the traditional IQ-discrepancy model and, we suggest, an RTI model. These were reliance on emergent measures of early literacy as a way to get below the floor effects found on typical reading tests and use of DA to “speed up” an RTI-based identification model. Our results provide more support for reliance on emergent measures of early literacy than for DA as a way to speed up the RTI process at the preschool level. However, it is important to emphasize the preliminary nature of our results, and to recognize the fact that use of DA to speed up the RTI process might work better in first grade than it does in preschool, even if our results are confirmed by other studies.

Several limitations to the current study should be noted. First, the present study is limited in its generalizability. The children in the present study were from middle and higher socioeconomic populations. Additionally, the use of multiple test administrations in a short period may limit the generalizability to instances when using the tasks at a single time.

Futhermore, the short duration of the study may not have allowed time for group differences to emerge. Another potential limitation is the lack of information on fidelity of administration for the dynamic-supported group. It may be the case that the differences in timing for increasing time intervals for the elision task and reducing time intervals for the blending task were not properly adhered to in all cases and thus it is difficult to determine if the lack of group differences represents a true difference or is due to lack of adherence to testing procedures. Also, the current study's conclusions are based on concurrent validity as the print awareness measure was administered at Time 3. Given these limitations, replication of the current findings is desired. Additionally, future studies comparing DA techniques should include broader samples of children, as well as struggling readers. Additional studies should also consider following children longitudinally in order to measure predictive validity.

In summary, preschool-age children can do phonological awareness tasks under standard administration conditions when items require processing of units larger than individual phonemes. DA may not improve reliability or validity of assessments when children are given tasks that they can perform using standard administration procedures.

Acknowledgments

Support for carrying out this research was provided by grant P50 HD052120 from the National Institute of Child Health and Human Development and by Predoctoral Interdisciplinary Research Training Grant R305B04074 from the Institute of Education Sciences.

References

- Anthony JL, Francis DJ. Development of phonological awareness. *Current Directions in Psychological Science*. 2005; 14(5):255–259.
- Berninger VW, Abbott RD, Vermeulen K, Ogier S, Brooksher R, Zook D, et al. Comparison of faster and slower responders to early intervention in reading: Differentiating features of their language profiles. *Learning Disabilities Quarterly*. 2002; 25:59–76.
- Caffrey E, Fuchs D, Fuchs LS. The predictive validity of dynamic assessment. *The Journal of Special Education*. 2008; 41(4):254–270.
- Coyne MD, Kame'enui E, Simmons DC, Harn BA. Beginning reading intervention as inoculation or insulin: First-grade reading performance of strong responders to kindergarten instruction. *Journal of Learning Disabilities*. 2004; 37:90–104. [PubMed: 15493232]
- Day JD, Englehardt JL, Maxwell SE, Bolig EE. Comparison of static and dynamic assessment procedures and their relation to independent performance. *Journal of Educational Psychology*. 1997; 89(2):358–368.
- Fletcher, JM.; Lyon, GR.; Fuchs, LS.; Barnes, MA. *Learning disabilities: From identification to intervention*. Guildford; New York: 2007.
- Gredler GR. Issues in early childhood screening and assessment. *Psychology in Schools*. 1997; 34:99–105.
- Grigorenko EL, Sternberg RJ. Dynamic testing. *Psychological Bulletin*. 1998; 124(1):75–111.
- Grigorenko EL. Dynamic assessment and response to intervention: Two sides of one coin. *Journal of Learning Disabilities*. 2009; 42(2):111–132. [PubMed: 19073895]
- Hogan TP, Catts HW, Little TD. The relationship between phonological awareness and reading: Implications for the assessment of phonological awareness. *Language, Speech, and Hearing Services in Schools*. 2005; 36:285–293.
- Lidz CS, Pena ED. Response to intervention and dynamic assessment: Do we just appear to be speaking the same language? *Seminars in Speech and Language*. 2009; 30(2):121–133. [PubMed: 19399697]
- Lonigan CJ. Development, assessment, and promotion of preliteracy skills. *Early Education and Development*. 2006; 17(1):91–114.

- Lonigan CJ, Burgess SR, Anthony JL. Development of emergent literacy and early reading skills in preschool children: Evidence from a latent-variable longitudinal study. *Developmental Psychology*. 2000; 36(5):596–613. [PubMed: 10976600]
- Lonigan CJ, Burgess SR, Anthony JL, Barker TA. Development of phonological sensitivity in 2- to 5-year-old children. *Journal of Educational Psychology*. 1998; 90(2):294–311.
- Lonigan, CJ.; Wagner, RK.; Torgesen, JK.; Rashotte, CA. *Test of Preschool Early Literacy*. Pro-Ed; Austin, TX: 2007.
- Meng X-L, Rosenthal R, Rubin DB. Comparing correlated correlation coefficients. *Psychological Bulletin*. 1992; 111(1):172–175.
- National Reading Panel. Report of the National Reading Panel: Teaching children to read (No. 00-4769). National Institute of Child Health and Human Development; Washington, D. C.: 2000.
- Olswang LB, Bain BA. Assessment information for predicting upcoming change in language production. *Journal of Speech and Hearing Research*. 1996; 39:414–423. [PubMed: 8729929]
- Reid, D.; Hresko, W.; Hammill, D. *Test of Early Reading Ability – Third Edition*. Pro-ED; Austin, TX: 2001.
- Schafer, JL. NORM: Multiple imputation of incomplete multivariate data under a normal model, version 2. 1999. Software for Windows 95/98/NT, available from <http://www.stat.psu.edu/~jls/misoftwa.html>
- Spector JE. Predicting progress in beginning reading: dynamic assessment of phonemic awareness. *Journal of Educational Psychology*. 1992; 84(3):353–363.
- Shepard, L.; Kagan, SL.; Wurtz, E., editors. *Principles and recommendations for early childhood assessments*. National Education Goals Panel; Washington, D. C.: 1998.
- Snow, CE.; Burns, MS.; Griffin, P. *Preventing reading difficulties in young children*. National Academy Press; Washington, D. C.: 1998.
- Swanson HL, Lussier CM. A selective synthesis of the experimental literature on dynamic assessment. *Review of Educational Research*. 2001; 71(2):321–363.
- Tabachnick, BG.; Fidell, LS. *Using multivariate statistics*. 5th ed.. Allyn and Bacon; New York: 2007.
- Torgesen JK. Individual differences in response to early interventions in reading: The lingering problem of treatment resisters. *Learning Disabilities Research and Practice*. 2000; 15:55–64.
- Torgesen JK, Wagner RK, Rashotte CA, Burgess S, Hecht S. Contributions of phonological awareness and rapid automatic naming ability to the growth of word reading skills in second- to fifth-grade children. *Scientific Studies of Reading*. 1997; 1(2):161–185.
- Wagner RK, Torgesen JK, Rashotte CA. Development of reading-related phonological processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. *Developmental Psychology*. 1994; 30(1):73–87.
- Wagner RK, Torgesen JK, Rashotte CA, Hecht SA, Barker TA, Burgess SR, et al. Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: A 5-year longitudinal study. *Developmental Psychology*. 1997; 33(3):468–479. [PubMed: 9149925]
- Whitehurst GJ, Lonigan CJ. Child development and emergent literacy. *Child Development*. 1998; 69(3):848–872. [PubMed: 9680688]
- Whitehurst, GJ.; Lonigan, CJ. *Get Ready to Read! Screening Tool*. National Center for Learning Disabilities; New York: 2001.

Table 1

Descriptive Statistics

Variable Name	All participants (n = 121)			Regular group (n = 42)			Supported group (n = 39)			Instructed group (n = 40)		
	M	SD	α	M	SD	α	M	SD	α	M	SD	α
Age (years)	4.42	0.66		4.43	0.77		4.38	0.64		4.46	0.57	
Print Awareness	18.87	10.68		15.68	9.75	.94	20.87	11.25	.96	20.30	10.62	.95
Time 1												
Elision	14.78	6.75		13.38	5.69	.85	15.72	8.07	.92	15.33	6.37	.88
Blending	22.85	8.84		22.50	8.84	.94	22.67	10.19	.96	23.40	7.66	.92
Initial Sound Matching	13.69	4.80		12.43	3.91	.64	14.31	5.48	.84	14.43	4.84	.75
Phonological Awareness Composite	0.00	0.85		-0.17	0.75	.94	0.08	0.99	.96	0.10	0.79	.93
Time 2												
Elision	17.19	7.92		14.48	6.01	.86	18.83	9.14	.96	18.43	7.94	.92
Blending	26.18	8.32		24.48	8.26	.93	27.56	8.39	.96	26.63	8.30	.94
Initial Sound Matching	15.84	5.89		13.22	4.01	.65	19.15	6.70	.92	15.33	5.34	.82
Phonological Awareness Composite	0.39	0.97		0.01	0.73	.93	0.76	1.10	.88	0.44	0.94	.96
Time 3												
Elision	18.52	8.83		15.23	7.24	.92	21.15	9.55	.96	19.43	8.75	.94
Blending	28.71	6.76		27.55	6.91	.94	29.83	6.47	.94	28.86	6.76	.93
Initial Sound Matching	16.62	6.49		14.07	5.00	.95	20.05	7.11	.93	15.90	6.06	.87
Phonological Awareness Composite	0.61	1.00		0.23	0.78	.95	1.02	1.09	.97	0.61	0.97	.96

Note. α = Cronbach's alpha of internal consistency.

Table 2

Validity coefficients among phonological awareness variables and print awareness variables by group

Variable	Regular	Supported	Instructed
Phonological Composite T1	.47 (.51)	.81 (.85)	.73 (.77)
Phonological Composite T2	.65 (.69)	.86 (.89)	.81 (.85)
Phonological Composite T3	.62 (.66)	.83 (.86)	.82 (.86)
Elision T1	.30 (.34)	.78 (.92)	.67 (.73)
Elision T2	.50 (.56)	.86 (.90)	.68 (.73)
Elision T3	.48 (.52)	.82 (.86)	.75 (.79)
Blending T1	.51 (.54)	.66 (.69)	.43 (.46)
Blending T2	.46 (.49)	.61 (.64)	.66 (.69)
Blending T3	.44 (.47)	.52 (.55)	.71 (.76)
Initial Sound Matching T1	.38 (.49)	.63 (.71)	.71 (.84)
Initial Sound Matching T2	.64 (.82)	.80 (.85)	.77 (.88)
Initial Sound Matching T3	.56 (.65)	.78 (.83)	.70 (.77)

Note. Regular group, $N = 42$. Supported Group, $N = 39$. Instructed Group, $N = 40$. Reliability corrected validity coefficients are in parenthesis. All validity coefficients are statistically significant, $p < .05$. Full correlation matrix is available by contacting the first author.

Table 3

Z-test for Comparing Correlated Validity coefficients across time

Variable	Phonological Composite		Elision		Blending		Initial Sound Matching	
	Z	p	Z	p	Z	p	Z	p
Regular Group								
Time 2 vs. Time 1	4.99	0.00**	2.82	0.00**	-1.13	0.13	36.43	0.00**
Time 3 vs. Time 1	-1.85	0.03*	2.02	0.02*	-0.76	0.22	1.22	0.11
Time 3 vs. Time 2	0.67	0.25	-0.99	0.16	-0.17	0.43	-3.38	0.00**
Supported Group								
Time 2 vs. Time 1	4.12	0.00**	3.09	0.00**	-0.57	0.28	2.92	0.00**
Time 3 vs. Time 1	0.21	0.42	0.76	0.22	-1.34	0.09	2.24	0.01*
Time 3 vs. Time 2	-10.87 ^a	0.00**	-1.64	0.05	-4.53	0.00**	-0.72	0.24
Instructed Group								
Time 2 vs. Time 1	1.91	0.03*	0.08	0.47	1.73	0.04*	2.7	0.00**
Time 3 vs. Time 1	2.13	0.02*	1.22	0.11	2.44	0.01*	-1.78	0.04*
Time 3 vs. Time 2	0.99	0.16	12.85 ^a	0.00**	2.16	0.02*	-5.1	0.00**

Note.

*
p < .05.

**
p < .01.

^aTo calculate, a correlation of 0.99 was used in place of the correlation between Time 2 and Time 3 Phonological Awareness Composite as reliability corrected validity coefficient exceeded 1.00.

Table 4

Z-test for Comparing Independent Validity coefficients across Group

Variable	Phonological Composite		Elision		Blending		Initial Sound Matching	
	Z	p	Z	p	Z	p	Z	p
Time 1								
Supported vs. Regular	2.11	0.02 *	2.50	0.01 *	0.75	0.23	1.06	0.15
Instructed vs. Regular	1.46	0.07	1.76	0.04 *	-0.31	0.38	2.12	0.02 *
Supported vs. Instructed	0.65	0.26	0.75	0.23	1.04	0.15	-1.04	0.15
Time 2								
Supported vs. Regular	1.80	0.04 *	2.61	0.00 **	0.69	0.24	0.27	0.39
Instructed vs. Regular	1.22	0.11	0.93	0.18	1.01	0.16	0.67	0.25
Supported vs. Instructed	0.57	0.28	1.67	0.05	-0.30	0.38	-0.39	0.35
Time 3								
Supported vs. Regular	1.47	0.07	2.16	0.02 *	0.31	0.38	1.26	0.10
Instructed vs. Regular	1.55	0.06	1.54	0.06	1.48	0.07	0.78	0.22
Supported vs. Instructed	-0.07	0.47	0.63	0.26	-1.14	0.13	0.48	0.31

Note.

*
 $p < .05$.**
 $p < .01$.