A Comparison of AIMSWeb TEL and PALS-K Reading Assessments

Reese Butterfuss & Allison Gehring, University of Wisconsin-Eau Claire
Faculty Mentor: Mary Beth Tusing, PhD, NCSP

Introduction

Response to Intervention (RtI) practices in educational settings aim to identify students who are at risk for academic struggles. Typically, RtI tries to identify 20% of students as at risk (Good, Simmons, & Smith, 2008). It is crucial to provide at-risk students with early interventions to prevent student failure (Hughes & Dexter, 2014). Standardized screening assessments are used in identifying students who are at-risk for failure. To successfully identify at-risk students, screeners must be reliable, valid, accurate, and inexpensive (Graney et al. 2010).

Wisconsin has mandated the Phonological Awareness Literacy Screening--Kindergarten (PALS, University of Virginia) to assess reading-related skills during students' kindergarten year. Although PALS-K is mandated, schools use other assessments as well. One early literacy screener commonly used in schools prior to the PALS mandate is the AIMSweb Tests of Early Literacy (Edformation). Both assessments claim to assess important early literacy skills that are predictive of future reading performance.

Gaither (2008) compared PALS and AIMSweb to determine which screener correctly identified students who were at-risk; the results found that the two assessments were strongly correlated with each other. Moreover, students identified as at-risk by PALS were also identified as at-risk by AIMSweb. These results suggested that the results favored AIMSweb as a more practical assessment because it is more economical in terms of time and money to administer.

More research is needed to examine the usefulness of the PALS screener in an RtI system.

Thus, our study addressed two questions:

1. How well does each screener correctly predict which children will perform below expectation (i.e. "at-risk" children) on Grade 1 reading measures?
2. Which kindergarten subtests are most predictive of first-grade reading performance?

Method

Participants. 66 (39 girls, 27 boys) school-aged children from a school district in Wisconsin. Specific demographic data was not available to describe our sample. However, according to the Wisconsin Department of Public Instruction, the overall school district reported that the student body is 1.1% Asian, 2.5% Black, 2.0% Hispanic.

Procedure. We accessed existing assessment data archivally. Students had completed early literacy screeners (PALS-K and AIMSWeb TEL) in the fall of kindergarten. The same students completed AIMSWeb TEL in the winter of their first-grade year. We used children's assessment scores and conducted analyses to answer the following questions:

Question 1: Classification Accuracy. Using the assessment tool's criterion scores for determining risk status (cut-scores described below), we conducted crosstabulations to compare students' benchmark status from kindergarten assessments to first grade reading performance. These analyses allowed us to determine the predictive accuracy (sensitivity and specificity) of AIMSWeb TEL and PALS-K for first-grade reading outcomes (R-CBM).

- *Sensitivity* answers the question of whether the assessment is indicating the correct students as failing to meet benchmark.

Sensitivity = true positives / (true positives + false negatives)

- *Specificity* answers the question of whether the assessment is indicating only those students who fail to meet benchmark.

Specificity = true negatives / (False positives + true negatives)

Question 2: Predictive Validity. We conducted a regression analysis to determine the extent to which screening score assessments predicted winter first-grade reading outcomes. We considered the regression coefficients of each screener to determine the relative contribution of each screening score to first-grade reading outcomes.

Measures.

Predictor variables.

AIMSWeb TEL is used as a screening tool to determine which students may be at risk for future reading difficulties. Only one measure is administered in the fall of kindergarten:

- *LNF: students present the names of visually presented letters for one minute.*

The cut-score defined by AIMSWeb as identifying the students who are most at-risk is 3. This cut score is set at the 15th percentile when compared to kindergarteners across the nation.

PALS-K consists of a Phonological Awareness Task (rhyme and beginning sound awareness), Literacy Tasks (alphabet knowledge), a Letter-Sound Awareness, and Concept of Word and Recognition. The test takes approximately *30 min*. to administer.

- *Rhyme and Beginning Sound Awareness: identify the words and sounds that rhyme in an orally given selection.*
- *Alphabet Awareness: students name the letters in the alphabet in both uppercase and lowercase print. For Letter-Sound Awareness: students touch each letter and say the sound it represents.*
- *Concept of Word: students point to the words as a teacher reads them aloud.*
- *Spelling: students spell consonant-vowel-consonant words.*
- *Recognition: students must read words from a word list.*

The cut-score defined by PALS-K as identifying the students most at-risk is 28. This score was determined by the rhyme awareness, spelling, Concept of Word List, Beginning Sound Awareness, Alphabet Awareness, and Letter-Sound Awareness.

Outcome Variables.

The outcome measure for students was R-CBM administered in the fall and winter of first grade. R-CBM is a brief, individually administered test of oral reading fluency. Students are scored based on the total number of words read correctly in 1 minute.

Results

Question 1: Classification Accuracy

PALS-K and R-CBM (Winter Grade 1)

		Tier 2 R-CBM (Winter Grade 1)	
		At Risk	Not at Risk
PALS-K Tier 2 Benchmark Score	At Risk	4	1
	Not at Risk	11	41

Sensitivity = .27, Specificity = .98, Hit Rate = .79

AIMSWeb and R-CBM (Winter Grade 1)

		Tier 2 R-CBM (Winter Grade 1)	
		At Risk	Not at Risk
LNF (Fall Kindergarten) Tier 2	At Risk	1	0
	Not at Risk	13	42

Sensitivity = .07, Specificity = 1.0 , Hit Rate = .77

Summary of Classification Accuracy

	PALS-K	AIMSWeb TEL
Sensitivity	.27	.07
Specificity	.98	1.0
Hit Rate	.79	.77

- PALS-K more accurately identifies at-risk children than AIMSWeb TEL.
- PALS-K shows higher sensitivity (.27) than AIMSWeb TEL (.07), meaning PALS-K identifies the most children as at-risk.
- AIMSWeb TEL shows slightly higher specificity (1.0) compared to PALS-K (.98). This means AIMSWeb is identifying *only* those students who are most at-risk

Question 2: Predictive Validity

Regression Analysis Comparing PALS-K and AIMSWeb in Predicting R-CBM

R-CBM (n = 57)

Predictor	Standard β	Partial r	sr^2
AIMSWeb LNF	.42*	.36	.38
PALS-K Sum Score	.30	.26	.08

*$p < .01$

- Simultaneous regression analyses revealed a significant beta weight for AIMSWeb LNF (β = .42). This means AIMSWeb TEL contributes more to R-CBM than PALS-K.
- Semipartial correlation coefficients revealed that AIMSWeb LNF accounted for 8% of the variance in R-CBM when added to PALS-K, and PALS-K adds only 4% of the variance to R-CBM when added to AIMSWeb TEL.

Discussion

Our research suggests that PALS, the assessment currently mandated in Wisconsin schools, may not be the most pragmatic assessment we could use, as AIMSWeb is costs less time and money to administer. In a multiple-gating strategy, AIMSWeb could be used as a quick "first-gate" assessment to indicate the children who need the most intensive intervention. The next "gate" could be a more comprehensive assessment, like PALS.

Implication for Question 1. Schools may need to consider an appropriate threshold of performance to define risk status in individual schools. Because schools may differ in score variability and students differ in skills, adjusting cut-scores may help correctly identify at-risk students. The poor sensitivity shown by both assessments indicates that both assessments are under-identifying at-risk students in our sample with the current cut-scores. In future research, we could analyze the level of performance on each assessment that's most predictive of the future outcome; instead of using identified cut-scores, it may be important for schools to consider how well screeners identify risk status and determine cut scores locally.

Implications for Question 2. Overall, AIMSWeb and PALS-K accounted for 45% of the total variance in R-CBM, (F = 21.9, $p <$ .001). Beta weights tell you how much each the predictors contribute to R-CBM when the shared variance is removed. Although the cut-scores suggest PALS is more accurate, considering the two variables together indicates that AIMSWeb is more closely associated with R-CBM. Furthermore, AIMSWeb adds more to the prediction, and was significant (β = .42), whereas PALS was not (β = .30). The semipartial correlation coefficients denote how much more AIMSWeb and PALS-K add to the R-CBM scores when considered together. AIMSWeb adds 8% more to the prediction of R-CBM, whereas PALS only adds 3.9%. Thus, it seems as though AIMSWeb is the better predictor of R-CBM, however the cut score may be too low. Thus, schools may need to consider adjusting cutscores based on student performance.

Potential limitations for this project include our relatively small sample size and incomplete data. Our sample consisted of 66 children, which may be insufficient if we are to generalize to school-age populations in general. Also, we do not know how well our sample generalizes to Wisconsin schools as a whole because of a lack of data from the school district.

In future research, we could ask similar questions using a larger, representative archival dataset.

Acknowledgments

Funding for this project was provided by the Office of Research and Sponsored Programs

References

- Gaither, A. C. (2008). *A comparison of two early education literacy benchmark assessments*.
- Good, R., Simmons, D., & Smith, S. (2001). *Reading Research Anthology: The Why of Reading Instruction*. Novato, CA: Arena P
- Graney, S., Martinez, R. S., Missall, K. N., & Aricak, O. (2010). Universal screening of reading in late elementary school: R-CBM versus CBM maze. *Remedial and Special Education, 31,* 368-377. doi:10.1177/0741932509338371
- Hughes, C., & Dexter, D. D. (2014). Universal Screening Within a Response-to-Intervention Model. *Universal Screening Within a RTI Model*. Retrieved April 17, 2014, from http://www.rtinetwork.org/learn/research/universal-screening-within-a-rti-model