

Journal of Psychoeducational Assessment

<http://jpa.sagepub.com/>

Construct and Predictive Validity Evidence for Curriculum-Based Measures of Early Literacy and Numeracy Skills in Kindergarten

Joseph Betts, Mary Pickart and Dave Heistad

Journal of Psychoeducational Assessment 2009 27: 83 originally published online 25 August 2008

DOI: 10.1177/0734282908323398

The online version of this article can be found at:

<http://jpa.sagepub.com/content/27/2/83>

Published by:



<http://www.sagepublications.com>

Additional services and information for *Journal of Psychoeducational Assessment* can be found at:

Email Alerts: <http://jpa.sagepub.com/cgi/alerts>

Subscriptions: <http://jpa.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations: <http://jpa.sagepub.com/content/27/2/83.refs.html>

>> [Version of Record](#) - Mar 16, 2009

[OnlineFirst Version of Record](#) - Aug 25, 2008

[What is This?](#)

Construct and Predictive Validity Evidence for Curriculum-Based Measures of Early Literacy and Numeracy Skills in Kindergarten

Joseph Betts

Center for Cultural Diversity & Minority Education, Madison, Wisconsin

Mary Pickart and Dave Heistad

Research, Evaluation, and Assessment Department, Minneapolis Public Schools, Minnesota

The assessment of early literacy and numeracy skills can provide useful and important information in pursuance of the goal to increase student academic achievement. At present, there have been promising results using curriculum-based measurement (CBM) for evaluating early literacy and early numeracy. There has been little research investigating the use of both early literacy and numeracy CBM-type skills assessments in a comprehensive battery. This research extends previous research by investigating the internal validity of an assessment battery of literacy and numeracy skills assessed at the end of kindergarten and external validity in the form of predictive validity for second-grade outcomes in reading and mathematics. Results suggest that correlated factors representing early literacy and numeracy skills are being measured. Evidence of convergent and incremental predictive validity with respect to later reading and mathematics outcomes was strong, but discriminant validity was not substantiated.

Keywords: *curriculum-based measurement; construct validity; predictive validity; assessment; early literacy*

There has been growing recognition that early assessment coupled with early interventions is necessary to provide a firm foundation for later student learning (Chaffee, Cunningham, Secord-Gilbert, Elbard, & Richards, 1990; Dunkle & Vismara, 2004; National Association for the Education of Young Children, 1997; National Association for the Education of Young Children & National Council for Teachers of Mathematics, 2003; Scanlon & Vellutino, 1996; West, Denton, & Hausken, 2000). The effective composition of assessments and interventions can be seen within the context of logical decision-making systems (Fuchs, 2003; VanDerHeyden & Snyder, 2006). These systems use assessments to detect students' difficulties early in their academic careers and inform interventions to stem the cumulative effects of skill deficiency. However, more extensive investigations of assessment instruments' technical adequacy in the area of early education need to be undertaken (Hojnoski & Missall, 2006).

Authors' Note: Please address correspondence to Joseph Betts, Center for Cultural Diversity and Minority Education, 29 High Point Woods Drive, #102, Madison, WI 53719; bett0088@umn.edu

The assessment of early literacy and numeracy skills has taken many forms. Some examples are the use of individually administered, fluency-based tests such as the Dynamic Indicators of Basic Early Literacy (Kaminski & Good, 1996) or the Individual Growth and Development Indicators (McConnell, McEvoy, & Priest, 2002), group-administered tests such as the Group Reading Assessment and Diagnostic Evaluation (American Guidance Service, 2001), or computerized adaptive tests such as STAR Early Literacy (Renaissance Learning, 2005). Although the approaches differ in their blueprints, they all attempt to assess general outcomes and to evaluate changes in student achievement over time.

At present, much research on the assessment of early literacy skills using a curriculum-based measurement (CBM) methodology and fluency-based assessment procedures (e.g. Dynamic Indicators of Basic Early Literacy or Individual Growth and Development Indicators) has shown good evidence of score reliability and validity under differing circumstances and with respect to diverse outcomes (Daly, Wright, Kelly, & Martens, 1997; Good, Simmons, & Kame'enui, 2001; Kaminski & Good, 1996; Marston, Pickart, Reschly, Muyskens, Heistad, & Tindal, 2007; McConnell, McEvoy, & Priest, 2002; McConnell, Priest, Davis, & McEvoy, 2002; Munn, 1994). Preliminary evidence of good psychometric properties of measures of early numeracy or early mathematics skills are also emerging (Chard et al., 2005; Clarke & Shinn, 2004; Daly, Wright, Kelly, & Martens, 1997; Daley & Lefevre, 1997; Joyce & Wolking, 1987; Munn, 1994; VanDerHeyden et al., 2004; VanDerHeyden, Witt, Naquin, & Noell, 2001). However, much of the research base for literacy and numeracy has remained independent.

Given these encouraging results, the examination of a comprehensive battery of skills developed from the perspective of CBM that incorporates both literacy and numeracy would be important. A battery that combines both literacy and numeracy components has the potential to increase overall information while also providing a comprehensive assessment of important formative skills that set a foundation for later learning. One such battery is the Minneapolis Kindergarten Assessment (MKA; Minneapolis Public Schools, 2004).

In this research, we investigated two aspects of validity, internal and external, of the MKA. Internal validity addressed the relationship between the subtests measuring literacy and numeracy skills, score reliability, and a set of confirmatory factor analyses to evaluate the underlying measurement model. External validity addressed the extent to which the MKA scores obtained at the end of kindergarten relate to later outcomes in reading and mathematics measured at the end of second grade.

Method

Participants

The participants in this study were kindergarten students in a large, urban, midwestern school district. Originally, all 3,174 kindergarten students were assessed. However, because of attrition, 2,180 students (69%) were evaluated at the end of second grade on both math and reading outcomes.

To evaluate the extent to which the students lost by attrition might bias results, each student's kindergarten scores were summed to get a total score. An additional dichotomous

variable was coded to represent whether the student had a reading score and a math score at second grade (coded 0 if a student had both second-grade scores and coded as 1 if one or both of the second-grade scores were missing). The dichotomous variable was used as a grouping variable in an analysis of variance with MKA total score as outcome. The results suggested that missingness of second-grade achievement data was not significantly related to kindergarten MKA scores, $F(1, 2971) = 1.09, p = .30$. This result indicated that the group lost to attrition was probably not different, with respect to kindergarten skills, from the group not lost to attrition. The students with complete cases could be considered a random subsample of the original sample and the statistics derived unbiased (Little & Rubin, 2002; Lohr, 1999; Schafer, 1997).

Ethnicity of participants was as follows: 39% African American, 30% European American, 16% Hispanic American, 11% Asian American, and 4% American Indian. Twenty-eight percent of the students were identified as limited English proficient with the following primary home languages: 48% Spanish, 26% Hmong, and 15% Somali; the other 11% consisted of more than 70 different languages. About 60% of the students were eligible for free or reduced-price lunch. Seven percent of the children had Individualized Education Plans.

Measures

Minneapolis Kindergarten Assessment. The MKA (Minneapolis Public Schools, 2004; Pickart, Betts, Sheran, & Heistad, 2005) is a standardized, individually administered assessment measuring common early literacy and numeracy skills developed during kindergarten. In this research, we used the rhyming, alliteration, letter naming, and letter sounds subtests for the early literacy subtests and the number sense, patterning/functions, and spatial sense/measurement subtests for the early numeracy subtests. The early literacy subtests are fluency-based tests scored by the total number of correct responses within a specific time span (1 min for letter naming and letter sounds and 2 min for rhyming and alliteration). The early numeracy tests are untimed subtests on which the total score consists of the total number of correct answers. Raw scores or normalized scores with a mean of 100 and standard deviation of 15 can be derived. The norm-referenced scores were used in this analysis because the norming population was similar to the study sample. Evidence for reliability was reported to be strong, with coefficients greater than .80 for both test-retest and internal consistency across multiple ethnic groups (Pickart et al., 2005). Concurrent and predictive validity evidence was reported to be strong, with correlations of .75 and .84, respectively (Pickart et al., 2005).

Northwest Achievement Levels Test (NALT). The NALT (Northwest Evaluation Association, 2003) has specific academic achievement tests of reading and mathematics. Both the reading and the math tests are scored using modern item response theory and provide an overall achievement score, the RIT scale score, that was used in this analysis. This assessment was given in a group-administered setting, and each student was administered a leveled form of the test that was tailored to his or her ability level. Validity estimates for the second grade with respect to the Stanford Achievement Test (Harcourt Educational Measurement, 1997) found correlations for reading scores and math scores, separately, to

be .86 and .80, respectively. Marginal reliability for second grade was above .90, and test-retest reliabilities were consistently above .70 for both reading and mathematics.

Procedures

Kindergarten students were assessed during May by a trained assessor using standardized procedures outlined in the technical manual (Minneapolis Public Schools, 2004). The cohort of assessors was retired teachers who completed 2 hours of administration training on the MKA. Training was done before all seasonal administration occasions (during the fall, winter, and spring) and consisted of both group instruction and didactic modules consisting of videotaped MKA administration examples. In addition, all assessors score the videotaped examples, and results are checked for reliability with a criterion of 90% agreement to the standard. The NALT reading and math tests were administered during April 2005. All teachers were responsible for complying with administration rules and procedures, and special days were set aside for all of the students to participate in the assessment.

Statistical Analysis

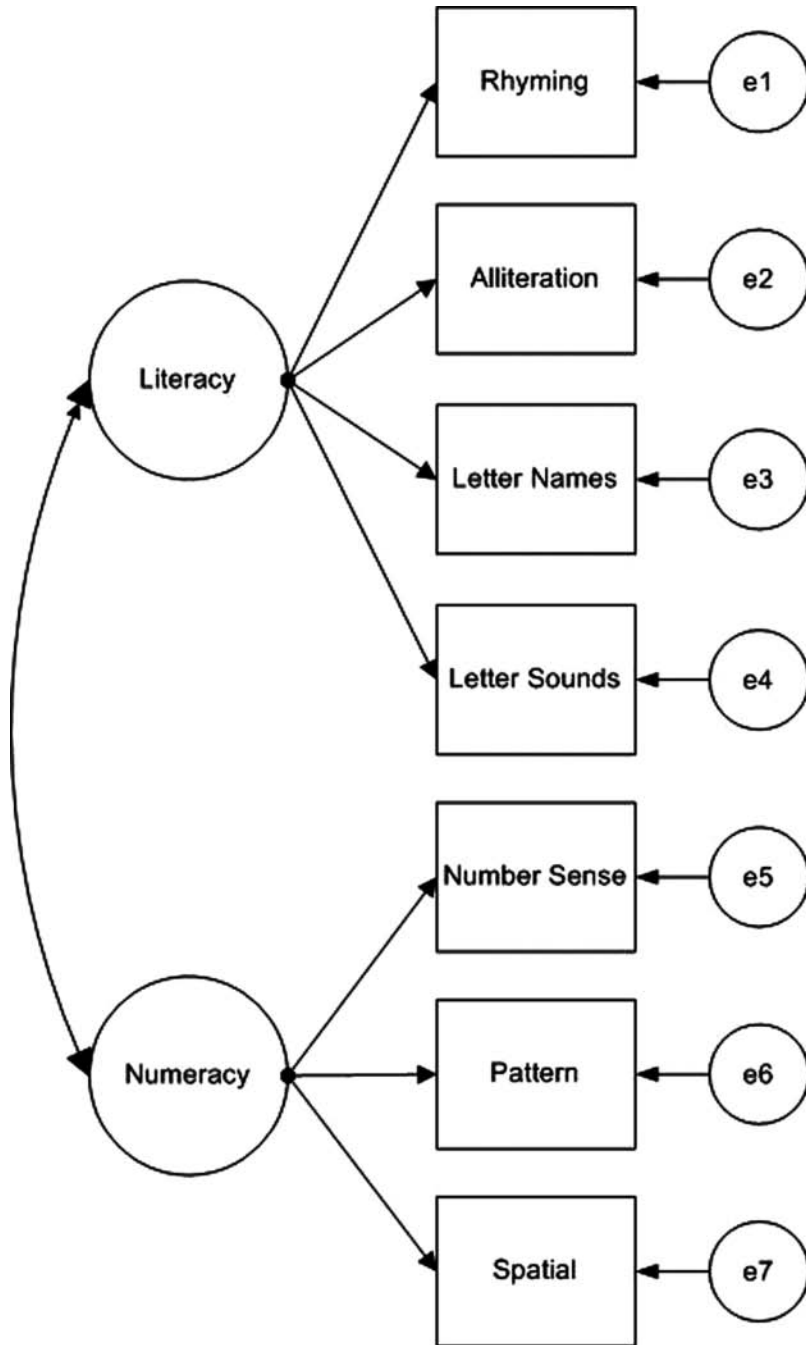
The first part of this research investigated internal validity by exploring the internal relationships of the MKA subtests. The reliability of scores obtained in the sample was computed using the Gilmer-Feldt coefficient (Feldt & Brennan, 1989; Gilmer & Feldt, 1983). We chose the Gilmer-Feldt coefficient because the MKA subtests were considered congeneric measures, and therefore traditional internal consistency reliability estimates, such as Cronbach's alpha, would not be appropriate.

An evaluation of the intercorrelations between subtests of the battery that were expected to measure the same domain (literacy or numeracy) was undertaken. Then an evaluation of the cross-correlations of subtests measuring one purported domain (e.g., literacy) with the measures of the other domain (e.g., numeracy) was undertaken. We expected that intercorrelations would be larger, on average, than cross-correlations, providing evidence of convergent and divergent validity.

Next, a more thorough analysis of the interrelations between all the subtests was undertaken using a confirmatory factor analysis (Jöreskog & Sörbom, 2001; McDonald, 1985; Mulaik, 1972). Three models were investigated. First, we examined a common factor model (Model 1), which posited a single common factor accounting for the relationship between all subtests. Second, we investigated a two-factor (Literacy and Numeracy) model (Model 2) with correlated factors and simple structure on the basis of the theoretical separation of literacy and numeracy skills (see Figure 1). The third model was a direct extension of Model 2 that allowed for the assumptions of uncorrelated residuals and simple structure to be relaxed.

To evaluate the fit of the data to the proposed confirmatory factor analysis models, we used the chi-square test of model fit, Tucker-Lewis Index, Comparative Fit Index, and root-mean-square error of approximation. Nonsignificant chi-square test statistics indicate good fit of the model to the data, as do Comparative Fit Index and Tucker-Lewis Index values at or above 0.94, and root-mean-square error of approximation values at or lower than 0.05

Figure 1
Two-Factor Correlated Model Diagram (Model 2)



can be considered good-fitting models (Hu & Bentler, 1999; Schumacker & Lomax, 1996). To evaluate the relative fit between the three different models, we used the Akaike and Bayesian information criteria, with lower values indicating better fit (Schumacker & Lomax, 1996).

The second part of this research investigated the external validity of the MKA battery measured in kindergarten with measures of reading and mathematics measured at the end of second grade. Correlations between MKA subtests and second-grade outcomes were computed. Evidence of convergent validity was evaluated by correlations between literacy variables and reading and between numeracy variables and math outcomes. Divergent validity was investigated with respect to correlations between literacy variables and math outcomes and numeracy variables with reading outcomes.

To evaluate the relative importance of the MKA subtests on second-grade reading and mathematics, the reading and mathematics scores were regressed onto all the MKA subtests. However, the regression was done in two steps to evaluate the incremental validity associated with adding the set of MKA subtests from the cross-domain to predicting second-grade outcomes. The second-grade reading scores were first regressed onto the literacy subtests, and then the set of numeracy subtests was added to evaluate the incremental increase in validity. Likewise, second-grade math scores were regressed on the set of early numeracy subtests first, and then the set of literacy subtests were added. Standardized regression coefficients were computed to evaluate the relative importance in predicting second-grade achievement scores when all MKA subtests were entered into the model.

Results

Internal Validity

Descriptive statistics suggested students' results were similar to expectations, with means close to 100 and standard deviations close to 15 (Table 1). All subtests had skew and kurtosis statistics below unity and should not bias results. Reliability of the MKA was estimated to be .84, which was similar to previously reported estimates. Correlations between the MKA subtests were all significant ($p < .01$; Table 1). The range of intercorrelations between the early literacy measures was .34 to .62 with a median of .53. The range of intercorrelations between the early numeracy subtests was .44 to .61 with a median of .55. The cross-correlations between the literacy and numeracy subtests ranged from .36 to .58 with a median of .44. Given that the cross-correlations were very similar to the intercorrelations, evidence of divergence was weak; however, the median of the cross-correlations was somewhat lower than the median of the intercorrelations.

Results of the confirmatory factor analysis models (Table 2) indicated that the common factor model (Model 1) had poor fit of the data to the hypothesized model on all model fit indices. Results for Model 2 suggested poor (root-mean-square error of approximation = 0.11) to moderate fit (Comparative Fit Index = 0.94). Model 3 appeared to fit the data very well with all fit indices except chi-square model fit, indicating excellent model fit. In addition, Model 3 had the lowest Akaike and Bayesian information criteria values, indicating better fit relative to the other hypothesized models. Therefore, Model 3 was deemed the

Table 1
Descriptive Statistics and Correlations Between Kindergarten
Subtests and Second Grade Achievement Measures

Subtest	Kindergarten							Second Grade	
	1	2	3	4	5	6	7	Reading	Math
1. Rhyming	—	.62	.42	.34	.44	.39	.36	.56	.53
2. Alliteration		—	.55	.50	.52	.44	.40	.58	.56
3. Letter Names			—	.61	.58	.45	.47	.57	.53
4. Letter Sounds				—	.51	.39	.41	.40	.34
5. Number Sense					—	.61	.55	.56	.53
6. Pattern						—	.44	.49	.48
7. Spatial							—	.42	.37
<i>M</i>	101	101	101	100	101	102	101	188	195
<i>SD</i>	15	15	15	15	14	14	14	16	15

Note: All correlations were significant at $p < .001$.

Table 2
Results of Confirmatory Factor Analysis With the Three Models

Fit Index	Model 1	Model 2	Model 3
χ^2	613.84	374.85	17.62
<i>df</i>	14	13	8
<i>p</i>	<.001	<.001	.02
Comparative Fit Index	0.91	0.94	0.99
Tucker-Lewis Index	0.86	0.91	0.99
Root-mean-square error of approximation	0.14	0.11	0.02
Akaike information criterion	36,877	36,640	36,293
Bayesian information criterion	36,957	36,726	36,406

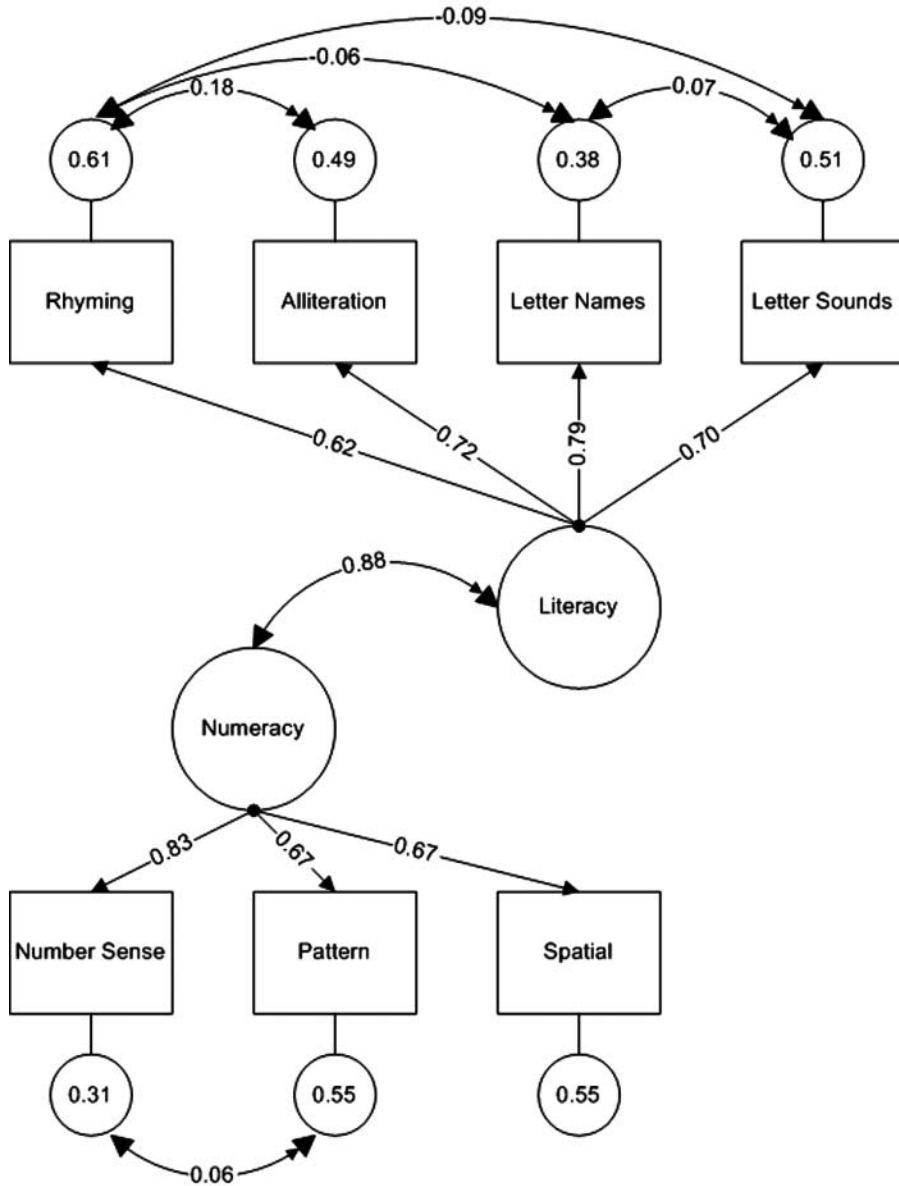
best-fitting model (see Figure 2 with coefficients given in their standardized form) and suggested that a two-factor simple structure model with correlated factors and correlated residuals fit the data well.

External Validity

Correlations between the MKA subtests and achievement tests were all significant ($p < .01$; Table 1). The correlation between second-grade reading and mathematics was .79. There did not appear to be a large amount of divergent validity because most of the kindergarten variables correlated just as highly with both reading and math. All subtests appeared to correlate more highly with reading than with math.

When predicting mathematics outcomes, the numeracy measures accounted for 33% of the variance ($R^2 = .33$), and when the literacy measures were added, this increased to 46%. This was a statistically significant increase, $\Delta F(4, 2172) = 138.45$, $p < .001$, and medium to large practical significance ($f^2 = 0.24$; Cohen, 1988). Prediction of reading outcomes was

Figure 2
Confirmatory Factor Analysis of the Two-Factor Correlated Model



also significantly increased, $\Delta F(3, 2172) = 60.34, p < .001$, with the addition of the numeracy measures to the literacy measures and resulted in an increase in R^2 from .48 to .52, which was of small practical effect ($f^2 = 0.08$; Cohen, 1988). These results indicate that the set of both early literacy and numeracy variables are significant predictors of later achievement, provide significant incremental validity, and account for a substantial amount of the variability

Table 3
Standardized Regression Results From
Regressing Second-Grade Achievement
Scores on MKA Subtests

MKA Subtest	NALT	
	Reading	Math
Rhyming	0.23*	0.21*
Alliteration	0.19*	0.21*
Letter Name	0.25*	0.21*
Letter Sound	0.07*	0.11*
Number Sense	0.16*	0.17*
Pattern	0.12*	0.15*
Spatial	0.03	0.01

Note: Standardized beta weights for each of the reading and math variable regressed on the MKA variables. MKA = Minneapolis Kindergarten Assessment; NALT = Northwest Achievement Levels Test

* $p \leq .001$.

in later achievement. The standardized regression weights for the model with all subtests in the prediction (see Table 3) indicated that all of the MKA variables except for the spatial/measurement subtest significantly predicted both later reading and math achievement, separately.

Discussion

This research took a step forward in understanding the properties of a battery of early literacy and numeracy measures at kindergarten with respect to internal and external validity. Results suggested that the MKA provided reliable scores for early literacy and early numeracy skills. Two factors related to literacy and numeracy skills were found to be distinct but highly correlated. Predictive validity with respect to both reading and mathematics outcomes at the end of second grade was strong. In addition, incremental validity indicated that both literacy and numeracy variables accounted for significant levels of variance in both reading and mathematics.

These results provided preliminary confirmatory evidence of the relationship between early literacy and numeracy skills and later outcomes, but the results also suggested that the relationship might not be clear cut. Convergent validity evidence was found, but divergent validity was weak for both internal and external validity. For instance, the cross-correlations between literacy and numeracy were substantial and did not indicate a strong separation between the domains even though the factor analytic results found good model fit to the correlated two-factor model. A consideration here was the high correlation between the two domains.

In addition, a lack of predictive discrimination was observed, with literacy and numeracy factors substantially adding to the prediction of both reading and mathematics. Interestingly, the incremental validity increase with the addition of the literacy variables to the prediction of later mathematics outcomes was substantial. Although numeracy variables significantly increased the proportion of variance accounted for in later reading above the literacy variables, the increase was much smaller.

The lack of divergence, both for internal and external validity, might indicate that the differentiation of reading and mathematics skills are less pronounced when measuring the early component skills on which the later academic achievement correspond. It might be that during early education the development of a broad range of skills from learning letters and numbers to learning sounds and counting will provide a general benefit for later reading and mathematics outcomes. Overall, these results indicate that a multivariate approach to making predictions about later achievement standing might be useful.

If the goal of early assessment and intervention is to increase student skill on components known to correlate with both reading and mathematics achievement, this research provides initial evidence that the measurement of the precursor skills can have beneficial payoffs with respect to predicting later outcomes. This research suggests that early assessments of literacy and numeracy skills can provide insights into student skill levels and important predictive information about future achievement. This predictive validity is interesting because the outcome variables in this research were administered two grades into the future. As both literacy and numeracy skills provided predictive utility to later reading and mathematics, they could both be useful in making determinations of benchmark criteria for identifying students in need of more intense monitoring or interventions.

In addition to providing insights into specific literacy and numeracy skills for which interventions can be targeted, the potential for using both sets of kindergarten variables, early literacy and early numeracy, in predictive models provides a potential for using a compensatory model of student achievement for making benchmark scores. As opposed to developing benchmark criteria based on only one of the MKA subtests, such as Letter Sounds, the full battery can be used to better predict students for whom remediation should be helpful. In addition to a compensatory predictive model that uses scores from all subtests, the potential for developing useful profiles could be investigated.

Limitations of the Present Study

Because of the scale of this research, it was difficult to obtain estimates of interrater reliability. Although we obtained reliability estimates with respect to student scores, it would be useful to examine the extent to which raters tended to rate responses in a similar manner. We suspect that interrater reliability would be substantial because of administrators having been retired teachers and also having participated in substantial training and scoring instruction.

Although the sample used in this research was representative of the district from which the students were obtained, it is not representative of the nation. Therefore, results may not generalize to all possible populations. In addition, the subtests used in the MKA might appear similar to other types of fluency-based tests, but the results might not generalize to

those other measures. Lack of generalizability should guide future use in student populations deemed to be significantly different from the sample in this research. All results should be replicated in different populations, and further evidence of validity is needed with respect to those different populations.

Possible Future Research

Future research focused on analyzing the possible reasons for the poor divergent validity evidence found in this study should be relevant. One consideration for addressing the issue of high levels of cross-correlations between the early literacy and early numeracy measures could be related to a background variable that is not accounted for in the present model. One potential variable could be a general cognitive ability variable like g , which has been found to underlay most mental tasks (Jensen, 1998). Future research might attempt to make a measure of g continuous with the measurement of the early literacy and numeracy skills and to partition out the variance associated with that factor. Thus, correlations between the early academic measures of literacy and numeracy could be done with the effects of that background variable partitioned out.

Research on the use of fluency-based tests with kindergarten students has shown positive results. However, the development of assessments for early numeracy skills is just emerging. This research suggests that more work is needed in the development of tasks that measure important areas of early mathematical development. Future research might attempt to better specify those early skills that contribute to the development of mathematical skills. The construct of early numeracy could be more fully developed and tasks developed to better sample this important domain of early education. Considering the concern about student mathematics performance, early assessments to engender early support could be highly beneficial. A thorough investigation of the most important numeracy skills to address in early education on par with the colossal undertaking of the National Reading Panel (National Institute of Child Health and Human Development, 2000) for reading appears to be needed. The results of this research suggest that measuring only literacy skills at the kindergarten level may provide an incomplete story of children's early academic development.

Overall, this research provides evidence to support the early assessment of both literacy and numeracy skills. However, continued research is vital both for constructing valid assessments and also for linking those results to useful instructional supports to foster student skill development. Highly reliable and valid assessment results can engender better decision making with respect to identifying students at risk for later difficulties. With a good early warning signal, it may be possible to ameliorate difficulties before they are too entrenched.

References

- American Guidance Service. (2001). *Group Reading Assessment and Diagnostic Evaluation (GRADE)*. Circle Pines, MN: American Guidance Services.
- Chaffee, C. A., Cunningham, C. E., Secord-Gilbert, M., Elbard, H., & Richards, J. (1990). Screening effectiveness of the Minnesota Child Development Inventory Expressive and Receptive Language Scales: Sensitivity,

- specificity, and predictive value. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 2, 80-85.
- Chard, D., Clarke, B., Baker, S., Otterstedt, J., Braun, D., & Katz, R. (2005). Using measures of number sense to screen for difficulties in mathematics: Preliminary findings. *Assessment for Effective Intervention*, 30, 3—14.
- Clarke, B., & Shinn, M. (2004). A preliminary investigation into the identification and development of early mathematics curriculum-based measurement. *School Psychology Review*, 33, 234—248.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Daley, K., & Lefevre, J. (1997, April). *The development of numeracy: A longitudinal study of counting and arithmetic*. Poster presented at the Biennial Meetings of the Society for Research in Child Development, Washington, DC.
- Daly, E., Wright, J., Kelly, S., & Martens, B. (1997). Measures of early academic skills: Reliability and validity with a first grade sample. *School Psychology Quarterly*, 12, 268-280.
- Dunkle, M., & Vismara, L. (2004). Developmental checkups: They're good, they're cheap and they're almost never done. What is wrong with this picture? *Exceptional Parent*, 34, 160-161.
- Feldt, L., & Brennan, R. (1989). Reliability. In R. Linn (Ed.), *Educational measurement* (3rd ed., pp. 105-146). New York: American Council on Education/Macmillan.
- Fuchs, L. (2003). Assessing intervention responsiveness: Conceptual and technical issues. *Learning Disabilities: Research & Practice*, 18, 172-186.
- Gilmer, J., & Feldt, L. (1983). Reliability estimation for a test with parts of unknown lengths. *Psychometrika*, 48, 99-111.
- Good, R. H., Simmons, D. C., & Kame'enui, E. J. (2001). The importance of decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third grade high-stakes outcomes. *Scientific Studies of Reading*, 5, 257-288.
- Harcourt Educational Measurement. (1997). *Stanford Achievement Test series* (9th ed., technical data report). San Antonio, TX: Author.
- Hojnoski, R., & Missall, K. (2006). Addressing school readiness: Expanding school psychology in early education. *School Psychology Review*, 35, 602-614.
- Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55.
- Jensen, A. (1998). *The g-factor: The science of mental ability*. Westport, CT: Praeger.
- Jöreskog, K., & Sörbom, D. (2001). *LISREL 8: User's reference guide* (2nd ed.). Lincolnwood, IL: Scientific Software International.
- Joyce, B., & Wolking, W. (1987). Standardized tests and timed curriculum-based assessments: A comparison of two methods for screening high-risk students. *Journal of Psychoeducational Assessment*, 5, 185-193.
- Kaminski, R. A., & Good, R. H. (1996). Towards a technology for assessing basic early literacy skills. *School Psychology Review*, 25, 215-227.
- Little, R., & Rubin, D. (2002). *Statistical analysis with missing data* (2nd ed.). New York: Wiley-Interscience.
- Lohr, S. (1999). *Sampling: Design and analysis*. New York: Duxbury Press.
- Marston, D., Pickart, M., Reschly, A., Muyskens, P., Heistad, D., & Tindal, G. (2007). Early literacy measures for improving student reading achievement: Translating research into practice. *Exceptionality*, 15, 97-117.
- 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, 3-14.
- 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* (4th ed., pp. 1231-1246). Washington, DC: National Association of School Psychologists.
- McDonald, R. (1985). *Factor analysis and related methods*. Hillsdale, NJ: Lawrence Erlbaum.
- Minneapolis Public Schools. (2004). *Minneapolis kindergarten assessment*. Minneapolis, MN: Minneapolis Public School Research, Evaluation & Assessment Division.
- Mulaik, S. (1972). *The foundations of factor analysis*. New York: McGraw-Hill.
- Munn, P. (1994). The early development of literacy and numeracy skills. *European Early Childhood Education Research Journal*, 2, 5-18.

- National Association for the Education of Young Children. (1997). *Developmentally appropriate practice in early childhood programs serving children from birth through age 8*. Retrieved September 1, 2006, from <http://www.naeyc.org/about/positions/pdf/PSDAP98.PDF>
- National Association for the Education of Young Children & National Council for Teachers of Mathematics. (2003). *Early childhood mathematics: Promoting good beginnings*. Retrieved September 1, 2006, from http://www.naeyc.org/about/positions/pdf/Mathematics_Exec.pdf
- National Institute of Child Health & Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups* (NIH Pub. No. 00-4754). Washington, DC: U.S. Government Printing Office.
- Northwest Evaluation Association. (2003). *Technical manual for the NWEA Measures of Academic Progress and Achievement Level Tests*. Portland, OR: Northwest Evaluation Association.
- Pickart, M., Betts, J., Sheran, C., & Heistad, D. (2005). *Minneapolis Kindergarten Assessment. Beginning and end of kindergarten assessment: Technical manual*. Minneapolis, MN: Minneapolis Public Schools.
- Renaissance Learning. (2005). *STAR Early Literacy: Computer-adaptive diagnostic assessment*. Wisconsin Rapids, WI: Author.
- Scanlon, D., & Vellutino, F. (1996). Prerequisite skills, early instruction, and success in first-grade reading: Selected results from a longitudinal study. *Mental Retardation and Developmental Research Reviews*, 2, 54-63.
- Schafer, J. (1997). *Analysis of incomplete multivariate data*. New York: Chapman & Hall/CRC.
- Schumacker, R., & Lomax, R. (1996). *A beginner's guide to structural equation modeling*. Mahwah, NJ: Lawrence Erlbaum.
- VanDerHeyden, A., Broussard, C., Fabre, M., Stanley, J., Legendre, J., & Creppell, R. (2004). Development and validation of curriculum-based measures of math performance for preschool children. *Journal of Early Intervention*, 27, 27-41.
- VanDerHeyden, A., & Snyder, P. (2006). Integrating frameworks from early childhood intervention and school psychology to accelerate growth for all young children. *School Psychology Review*, 35, 519-534.
- VanDerHeyden, A., Witt, J., Naquin, G., & Noell, G. (2001). The reliability and validity of curriculum-based measurement readiness probes for kindergarten students. *School Psychology Review*, 30, 363-382.
- West, J., Denton, K., & Hausken, E. (2000). *America's kindergartners* (NCES 2000-070). Washington, DC: National Center for Educational Statistics.