Psychology in the Schools, Vol. 42(4), 2005
Published online in Wiley InterScience (www.interscience.wiley.com).

DOI: 10.1002/pits.20096

CONCURRENT AND PREDICTIVE VALIDITY OF THE PHELPS KINDERGARTEN READINESS SCALE-II

JENNIFER DUNCAN AND ERIN M. RAFTER

State University of New York at Buffalo

The purpose of this research was to establish the concurrent and predictive validity of the Phelps Kindergarten Readiness Scale, Second Edition (PKRS-II; L. Phelps, 2003). Seventy-four kindergarten students of diverse ethnic backgrounds enrolled in a northeastern suburban school participated in the study. The concurrent administration of the PKRS-II and the Woodcock-Johnson III Brief Intellectual Ability Scale (R.W. Woodcock, K.S. McGrew, & N. Mather, 2001a) occurred in the fall of the kindergarten year. To assess predictive validity, the Woodcock Johnson III Tests of Achievement (R.W. Woodcock, K.S. McGrew, & N. Mather, 2001b) was administered in the spring of that year. All concurrent and predictive correlations were significant. Based on the results of this study, the PKRS-II may be used with confidence to screen for children who may be at risk for academic difficulties. © 2005 Wiley Periodicals, Inc.

Numerous national task forces, commissions, and initiatives have emphasized the importance of children's early years (e.g., Diamond, Reagan, & Bandyk, 2000). Likewise, there are national education goals that stress the importance of children being ready to learn when they enter school. Yet, there are many interpretations regarding what school readiness means. Kagan (1994) stated that readiness includes two constructs: readiness to learn and readiness for school. Readiness to learn includes the developmental processes whereas readiness for school refers to the specific set of skills a child must have before he or she is ready to enter kindergarten. Despite competing definitions, the concept of readiness for school continues to hold a strong place in national discussions (Diamond et al., 2000). This is likely because mastery of readiness skills has been shown to correlate with performance during subsequent instruction on related academic tasks (VanDerHeyden, Witt, Naquin, & Noell, 2001).

Most placement decisions for kindergarten are based on age. When students meet the age requirement set by each school district, they may then be evaluated for school readiness (Augustyniak, Cook-Cottone, & Calabrese, 2004). In most states, education laws require that screening procedures be used with each new student and usually specify that diagnostic screening include a health examination and determination of language and cognitive skills (Costenbader, Rohrer, & DiFonzo, 2000). Most educators agree that screening assessments are an important component in the process of identifying young children who are at risk for subsequent educational difficulties (Scott & Delgado, 2003). Likewise, readiness tests are usually the initial evaluative tool utilized to identify children who may need more in-depth assessments (Gumpel, 1999; Scott & Delgado, 2003). When used for the identification of children with possible handicapping conditions, the results may assist in instructional planning (Pianta & McCoy, 1997; Shepard, 1997; Witt, Naquin, & Noell, 2001).

The identification of children with learning difficulties at an early age presents a critical educational issue. That is, the early identification of, and hence intervention with, children having mild to moderate learning difficulties may greatly enhance their long-term academic expectations (Scott & Delgado, 2003). Yet, a major concern when testing young children is the developmental flux that occurs at this age, making prediction of future academic achievement complicated. Screening measures must be reliable and valid to decrease the incidence of false positives (i.e., students

Correspondence to: Jennifer Duncan, 409 Baldy Hall, SUNY-Buffalo, Buffalo, NY 14260–1000. E-mail: jduncan@buffalo.edu

who are identified as at risk, yet are not) and false negatives (i.e., students who could benefit from further assessment, but were not identified; May & Kundert, 1997; Shepard, 1997). In fact, the challenging search for precise identification measures that identify the right children have spanned the past 75 years (Bishop, 2003).

The selection of instruments used in kindergarten screening should be based on adequate standardization, reliability, and validity (Costenbader et al., 2000). The predictive validity of a measure is a key issue for readiness tests. If a school is screening for possible learning difficulties, the measure that is chosen needs to be at least moderately, if not strongly, predictive of later academic achievement (Oosthuizen, Van Rensburg, & Jordaan, 1997). To this end, the purpose of this study was to evaluate the concurrent and predictive validity of the Phelps Kindergarten Readiness Scale-II.

Method

Participants

Participants were 76 kindergarten children (52% female, 48% male) attending a western New York school district. The sample consisted of 48 Caucasians, 18 African Americans, 7 Latinos, 2 Asians, and 1 American Indian. Students ranged in age from 5 years 2 months to 6 years 1 month ($M = 5.6$, $SD = .45$) at the time of initial testing.

Measures

Phelps Kindergarten Readiness Scale-II. The original Phelps Kindergarten Readiness Scale (PKRS; Phelps, 1991) was developed to assess the academic readiness of children preparing to enter kindergarten. The PKRS measures three domains: Verbal Processing, Perceptual Processing, and Auditory Processing. These domains were selected based on research that identified areas predictive of later academic achievement and consist of subtests designed to assess problem-solving skills rather than test the retrieval of previously learned facts (Phelps, 1991). Each processing domain yields a scaled score ($M = 10$, $SD = 3$). The three processing domains, when added together, result in a Total Readiness Score. The Total Readiness Score is a standard score ($M = 100$, $SD = 15$).

The Verbal Processing Domain consists of three subtests: Vocabulary, Verbal Reasoning, and Analogies. This domain assesses a child's knowledge of and ability to decipher the meaning of words and the ability to understand and recognize verbal relationships. The Perceptual Processing Domain combines two subtests (Visual Discrimination, Perceptual Motor) that evaluate a child's ability to compare shapes visually and to reproduce shapes that become increasingly more difficult. The final domain is Auditory Processing, as measured by three subtests: Auditory Discrimination, Auditory Digit Memory, and Memory for Sentences/Stories. These subtests assess a child's ability to discriminate between sounds and to recall auditory material.

The PKRS was normed on 554 children across the United States with ethnic, gender, and socioeconomic status matching the 1980 U.S. Census Data (Phelps, 1991). Alpha reliability coefficients for the PKRS ranged from .91 to .97, with the Total Readiness Score coefficient being .93. The test-retest reliability coefficient for the entire test (98 items) was .87, with a range of .61 to .87 for the individual domains. The concurrent validity of the PKRS was established with the Missouri Kindergarten Inventory of Developmental Skills (KIDS; Missouri Department of Elementary and Secondary Education, 1981), which yielded a correlation of .75 between the PKRS and the KIDS (Phelps, 1991). The predictive validity of the PKRS was established with the Woodcock-Johnson III Tests of Achievement (WJ III ACH; Woodcock et al., 2001b) at a 9-month interval with correlations ranging from .52 to .78. The predictive validity also was determined with the Otis-

Lennon School Ability Test at a 2.75-year interval (.62, $p < .0001$) (Otis & Lennon, 1996), the Stanford Achievement Test at a 2.75-year interval (with correlations ranging from .62–.69) (Harcourt Brace Educational Measurement, 1997), and the Tennessee Comprehensive Assessment Program at a 2.5-year interval (with correlations ranging from .32–.45) (Tennessee State Department of Education, 2003).

The PKRS was restandardized to update and recalibrate the norms (PKRS-II; Phelps, 2003). The restandardization sample was selected based upon ethnicity, age, gender, parental education, and geographic location to approximate the 2000 Census data (Phelps, 1991) and consisted of 510 children ranging in age from 4 years 5 months to 6 years 1 month ($M = 5.25$ years, $SD = .41$). As with the original PKRS, the PKRS-II has these domains and a Total Readiness Score. Reliability coefficients for the PKRS-II range from .75 to .93.

WJ III Brief Intellectual Ability (WJ III BIA). The WJ III Tests of Cognitive Abilities (Woodcock et al., 2001a) provide a measure of intelligence with 10 subtests. Of these, 3 subtests comprise the BIA measure (Verbal Comprehension, Concept Formation, Visual Matching). The BIA provides a short, but reliable, measure of intelligence and is appropriate for screening purposes. The alpha coefficients for the 3 tests in the WJ III BIA with 5-year-old children are: Verbal Comprehension, $r = .89$, Concept Formation, $r = .94$, and Visual Matching, $r = .93$. The test-retest coefficient for the WJ III BIA was .94 (McGrew & Woodcock, 2001).

WJ III ACH. Three broad area scores from the WJ III ACH were utilized: Broad Reading, Broad Math, and Broad Written Language. The area of Broad Reading consists of the Reading Fluency, Passage Comprehension, and Letter-Word Identification subtests, which draw on the child's reading speed, comprehension of language, and decoding skills. The area of Broad Math includes the Math Calculation, Math Concepts and Math Fluency subtests, which evaluate a child's math achievement, number fluency, and quantitative reasoning. The area of Broad Written Language comprises the Writing Samples, Spelling, and Sentence Completion subtests, which assess a child's writing ability, spelling ability, and writing speed. The WJ III test-retest correlations of the broad clusters range from .92 to .96.

Procedure

After obtaining written informed consent from the parent(s) of each participating child, two advanced doctoral school psychology students administered the PKRS-II and the WJ III BIA in counterbalanced order over a 3-week period in September and October of the kindergarten year. The assessments occurred over 2 half-hour sessions during regular school hours. The WJ III BIA standard score and the PKRS-II three domain scores (Verbal, Perceptual, Auditory) and Total Readiness Score were calculated for each child. Eight months later, during the months of May and June, the WJ III ACH was administered. After testing was completed, children received a Wal-Mart gift card for their participation.

Results and Discussion

The concurrent and predictive validity coefficients are illustrated in Tables 1 and 2. The pattern of moderate intercorrelations between the PKRS-II and the WJ III BIA imply that tests of intellectual ability and academic readiness are related, but inherently different. The PKRS-II and WJ III ACH correlations suggest that the screening test is related to subsequent academic functioning in kindergarten.

The author of the PKRS-II (Phelps, 1997) has recommended that the Total Readiness Score be used when making screening decisions because of increased reliability and validity. These data would support that recommendation. The strength of the correlational data indicated that the use

Table 1
Means, SDs, and Concurrent Validity Coefficients

	PKRS-II			
	Verbal	Perceptual	Auditory	Total Readiness
M 11.08	11.32	11.57	105.96	
SD 2.07	2.43	2.70	11.51	
W-J III BIA	.60	.41	.52	.59
M 103.45				
SD 16.18				

Note. PKRS-II = Phelps Kindergarten Readiness Scale (2nd ed.); W-J III BIA = Woodcock-Johnson III Brief Intellectual Ability Score.
All correlations significant at $p \leq .0001$.

of the PKRS-II provides a valid and appropriate measure for kindergarten screening and readiness. Nonetheless, it is important to note that the PKRS-II is to be utilized *solely* as a screening measure that may identify children who require a more comprehensive assessment and possible close monitoring or special services.

The strengths of this study include the broad ethnicity of the sample. Nonetheless, it is important to note research limitations. First, the sample size was relatively small ($N = 76$) and included one school district in Western New York. Further, this study calculated concurrent and predictive validity with two other measures. Future research should include a larger sample size, broader range of school districts, and utilization of other measures such as classroom performance. A cross validation of the test also would be beneficial. Additionally, future researchers could investigate

Table 2
Means, SDs, and Predictive Validity Coefficients

	PKRS-II			
	Verbal	Perceptual	Auditory	Total Readiness
W-J III ACH				
Broad Reading	.32	.32	.37	.39
M 108.18				
SD 17.35				
Broad Math	.34	.58	.52	.53
M 103.91				
SD 15.83				
Broad Written Language	.29	.49	.45	.46
M 112.48				
SD 16.46				

Note. PKRS-II = Phelps Kindergarten Readiness Scale (2nd ed.); W-J III ACH = Woodcock-Johnson III Tests of Achievement.
.29 = $p \leq .05$.
.32–.39 = $p \leq .01$.
.45–.58 = $p \leq .001$.

the longitudinal predictive validity of the PKRS-II by assessing students as they progress through the primary grades.

In conclusion, despite some limitations, the present findings indicate good concurrent and predictive validity of the PKRS-II. School professionals may use the PKRS-II with confidence to screen children entering school. Early detection of academic risk may help schools intervene quickly to increase the likelihood of successful academic progression. Nonetheless, school personnel are cautioned to utilize this scale only as a screening tool and not as a diagnostic measure.

References

Augustyniak, K.M., Cook-Cottone, C.P., & Calabrese, N. (2004). The predictive validity of the Phelps Kindergarten Readiness Scale. Psychology in the Schools, 41, 509–516.

Bishop, A.G. (2003). Prediction of first-grade reading achievement: A comparison of fall and winter kindergarten screenings. Learning Disability Quarterly, 26, 189–200.

Costenbader, V., Rohrer, A.M., & DiFonzo, N. (2000). Kindergarten screening: A survey of current practice. Psychology in the Schools, 37, 323–332.

Diamond, K.E., Reagan, A.J., & Bandyk, J.E. (2000). Parents' conceptions of kindergarten readiness: Relationships with race, ethnicity, and development. Journal of Educational Research, 94, 93–100.

Gumpel, T.P. (1999). Use of item response theory to develop a measure of first grade readiness. Psychology in the Schools, 36, 285–293.

Harcourt Brace Educational Measurement. (1997). Stanford Achievement Test. San Antonio, TX: Harcourt Brace Educational Measurement.

Kagan, S.L. (1994). Readying school for young children: Problems and priorities. Phi Delta Kappan, 76, 226–233.

May, D.C., & Kundert, D.K. (1997). School readiness practices and children at-risk: Examining the issues. Psychology in the Schools, 34, 73–84.

McGrew, K.S., & Woodcock, R.W. (2001). Technical Manual. Woodcock-Johnson III. Itasca, IL: Riverside.

Missouri Department of Elementary & Secondary Education. (1981). Missouri Kindergarten Inventory of Developmental Skills Guidebook. Jefferson City, MO: Author.

Oosthuizen, S., Van Rensburg, E., & Jordaan, C. (1997). Comparison of the predictive validity of the Aptitude Test for School Beginners and of the Metropolitan Readiness Tests. Perceptual and Motor Skills, 84, 59–65.

Otis, A.S., & Lennon, R.J. (1996). Otis-Lennon School Ability Test. San Antonio, TX: Harcourt Brace Educational Measurement.

Phelps, L. (1991). Phelps Kindergarten Readiness Scale. Brandon, VT: Psychology Press.

Phelps, L. (1997). Phelps Kindergarten Readiness Scale: Manual. Brandon, VT: Psychology Press.

Phelps, L. (2003). Phelps Kindergarten Readiness Scale: Manual. Brandon, VT: Psychology Press.

Pianta, R.C., & McCoy, S.J. (1997). The first day of school: The predictive validity of early school screening. Journal of Applied Developmental Psychology, 18, 1–22.

Scott, M.S., & Delgado, C.F. (2003). Using educational placement in third grade to select and validate a preschool screening measure. Psychology in the Schools, 40, 565–582.

Shepard, L.A. (1997). Children not ready to learn? The invalidity of school readiness testing. Psychology in the Schools, 34, 85–97.

Tennessee State Department of Education. (2003). Tennessee Comprehensive Assessment Program. Monterey, CA: CTB, McGraw-Hill.

VanDerHeyden, A.M., Witt, J.C., Naquin, G., & Noell, G. (2001). The reliability and validity of curriculum-based measurement readiness probes for kindergarten students. School Psychology Review, 30, 363–382.

Witt, J.C., Naquin, G., & Noell, G. (2001). The reliability and validity of curriculum-based measurement readiness probes for kindergarten students. School Psychology Review, 30, 363–382.

Woodcock, R.W., McGrew, K.S., & Mather, N. (2001a). Woodcock-Johnson III Tests of Cognitive Abilities. Itasca, IL: Riverside.

Woodcock, R.W., McGrew, K.S., & Mather, N. (2001b). Woodcock-Johnson III Tests of Achievement. Itasca, IL: Riverside.