Published in final edited form as:

J Pers Eval Educ. 2007 December 1; 20(3): 247–259. doi:10.1007/s11092-008-9055-x.

Development and Validation of a Principal Implementation Practices Measure: The Principal Implementation Questionnaire

Stephen M. Nettles and Yaacov Petscher

Florida Center for Reading Research, Florida State University

Abstract

Measurement of principal implementation behaviors has proved difficult to researchers in educational leadership due to a lack of consensus on the operational definitions of leadership constructs. The Principal Implementation Questionnaire (PIQ) was developed and validated with the intention of providing clarity in the assessment of principal leadership behaviors in the implementation of effective reading programs. Constructs were operationally defined within the context of the population of interest, with subsequent item writing centered around the constructs. A resulting calibration sample of principals from Florida *Reading First* schools was used to test the hypothesized measurement model to determine how well the items were described by the proposed factors. Results from LISREL analyses revealed a well-fitted model, based on numerous fit indices.

Keywords

Education leadership; Principal; Reading; Instrument development

There are few ideas in education less controversial than the notion that reading is the gateway to other academic knowledge. As such, reading is the focus of many national, state and local education accountability initiatives. Educators, parents and national leaders recognize that the limited reading skills of many children across our nation are an issue of national importance. Unfortunately, this understanding has not done much to halt the increase in the population of struggling readers in the United States. The 2005 National Assessment of Educational Progress (NAEP) indicated that approximately 36% of fourth grade students cannot read at a basic level, and approximately 69% cannot read at a level deemed “proficient”. When considering a typically cited student subgroup, 54% of low-income fourth grade students cannot read at basic levels. In addition, student performance trends over a 10-year period show insignificant progress for average readers and even less progress for the lowest-performing readers (Perie et al. 2005). To better inform the debate on this national crisis, there has been extensive research focused on the prevention of reading difficulties in children in recent years (Connor et al. 2007; Torgesen 2000). This body of research has provided a clearer understanding of the factors critical to addressing the needs of struggling readers in the early grades (K-3). With a base knowledge of what is

needed to help children learn to read has also come a clearer understanding of the challenges of implementing effective reading programs in our schools.

The effects of principals on school climate, school effectiveness and student achievement have also been of interest for several decades (Ediger 1999; Hallinger and Heck 1996; Levine and Lezotte 1990; Witziers et al. 2003). Furthermore, the role of the principal as the curriculum leader, guide, model and mentor (Ediger 2000) has been addressed by many scholars in the field (Levine and Lezotte 1990; Sammons et al. 1995) and identify the principal as one of the key factors that determines the success of various instructional interventions in a school.

While it is clear that principals influence overall school effectiveness, the specific behaviors associated with student achievement is unclear (Hallinger and Heck 1996, 1998; Pitner 1988; Snyder et al. 1996; Witziers et al. 2003). Thus, the effects of instructional leadership on individual student learning remain a complex and unanswered question (Bossert et al. 1982; Hallinger and Heck 1996, 1998; Heck 1993). Moreover, since reading is the primary skill affecting student development in all cognitive and social domains (Snow et al. 1991, 1998), modern principals are expected to be knowledgeable in methods of reading instruction.

Researchers have yet to agree upon appropriate methods of assessing leadership behaviors, as leadership is a multifaceted and contextually-based construct (Hallinger and Heck 1996; Witziers et al. 2003). A lack of consensus in operationalizing leadership has impeded efforts to adequately capture the effects of leadership on student achievement (Bossert 1988; Hoy and Ferguson 1985; Witziers et al. 2003). As comprehensive assessment of principal leadership practices is daunting—especially in relation to student achievement—it is all the more imperative that valid and reliable measures be developed to measure principals' impact on student learning. To date, there is a noteworthy lack of these instruments being used to evaluate the instructional leadership competencies of school principals in the area of reading.

Context of Instrument Development

The National Reading Panel report, *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* (National Reading Panel 2000) synthesized relevant research (over 100,000 studies) identifying standards that schools should follow. *Reading First* is a program that was conceived as a result of these findings.

Reading First is a federal program authorized under Title I, Part B, Subpart 1 of the *No Child Left Behind Act of 2001* (NCLB 2002) as a national initiative to improve the reading skills of students in kindergarten through third grade. The program is focused on effective classroom instruction supported by scientifically-based reading research, and provides a large amount of funding (approximately \$1 billion for FY 2006 and \$10 billion over the 10-year program) to states for activities targeting reading improvement in young students (United States Department of Education 2006). *Reading First* offers the opportunity for every state to receive funds, 20% of which can be used at the state level, with the remainder

being distributed to local education agencies (LEAs) to be used for the improvement of reading achievement in grades K-3.

The implementation of *Reading First* requirements places additional responsibilities on schools, especially principals. *Reading First*, like NCLB, carries with it high-stakes consequences for failure in return for the funding it provides to schools. Schools must set annual reading achievement goals that will facilitate growth to meet the *Reading First* requirement of all students reading on grade level by 2013. This goal has placed tremendous pressure on American schools to improve, and has garnered much attention from practitioners and policymakers alike.

While each *Reading First* state—and school—is unique in some respects, most face similar challenges in addressing the instructional needs of students, especially our most struggling readers. These issues include: (a) effectively identifying students that need additional support in reading, (b) selecting effective instructional programs based on scientific reading research, (c) meeting the diverse instructional needs of certain student subgroups (e.g., minority students, economically disadvantaged students, students with disabilities, and students with limited English proficiency), and (d) effectively managing these efforts at the state, district and, most importantly, the school level. In each of these areas, Florida is one of the nation's most challenging states. Given the large number of schools and students served under *Reading First* across the state, the variation in student diversity and the associated difficulties of providing adequate resources creates considerable challenges for educational leaders. The magnitude of the challenges at the instructional, and especially the administrative level, makes Florida an excellent example when considering national *Reading First* issues.

Under the Florida *Reading First* plan, if a school does not meet its achievement goals for reading at or above grade level by the end of the third year of *Reading First* implementation (2005–2006), the school district will then be asked to make a determination about whether the school should be continued as part of their *Reading First* grant for the next 3 years. Therefore, it is critical that each *Reading First* principal makes implementing effective reading programs a high priority, or risk losing significant funding (a minimum of \$40,000 per school, per year; FDOE 2002).

As part of the *Reading First* initiative, principals of participating schools are asked to participate in end-of-year evaluations of implementation practices. In prior years, this took the form of a qualitative survey with numerous free response areas for principals to describe the status of *Reading First* implementation and the practices that were used in this effort during the academic year. A revision process was initiated by the researchers in an attempt to collect more precise data on these practices that could be used for empirical analysis. The resulting instrument, the Principal Implementation Questionnaire (PIQ), was developed and validated to serve this purpose.

Initial Instrument Design

The PIQ was constructed to quantitatively assess the quality of principal implementation practices in relation to effective reading programs. The PIQ was initially written as a 105-item web-based inventory that was given to school principals to inquire about the implementation of *Reading First* activities in their school during the 2004–2005 school year. There are five dimensions of the PIQ that should be completed in approximately 25 min by school principals: (a) Core Reading Program, (b) Professional Development, (c) Leadership/Organizational Practices, (d) Assessment, and (e) Intervention. These five dimensions were selected and developed based on priorities set forth by state and federal *Reading First* policies, as well as a research-based understanding of the school-level elements needed to promote student achievement in reading. Further, each of these dimensions is focused on elements of overall reading program implementation that can be influenced by principal behaviors. Operational definitions for these five dimensions are as follows:

Core Reading Program

Core reading programs are defined as scientifically-based reading research curricula that are utilized as the foundation for instructional practice. Implementation of a systematic core reading program is essential to a successful reading program. The PIQ includes questions regarding the quality of core reading program implementation in each school that can be influenced by principal behavior.

Professional Development

Providing effective professional development that is targeted to the actual needs of school staff is one of the responsibilities of any principal and, particularly, *Reading First* principals. The challenges faced by *Reading First* schools are magnified due to especially challenging student populations, adding emphasis to the need for effective professional development. Principals should provide opportunities for professional development to build the professional competency of staff and themselves, improve environmental conditions affecting reading instruction and build rapport and collegiality among faculty. Thus, the PIQ includes questions regarding the existence and quality of professional development efforts in each school.

Leadership/Organizational Practices (Leadership)

Clearly, effective leadership is critical to the implementation of effective reading programs at the school level, as it is in any school improvement effort.

Organizational management, goal setting, progress monitoring and evaluation are some of the essential elements associated with this dimension. In this section of the PIQ, principals are asked to respond to items that are focused on the relevant practices that are associated with effective instructional leadership in reading.

Assessment

Developing an appropriate assessment plan, monitoring student outcome data and using these data to improve overall instructional practices are elements of this dimension. The use

of assessment to inform and drive reading instruction is a keystone of *Reading First* and, indeed, any instructional effort. This section of the PIQ includes items associated with the principal's role in ensuring a powerful assessment component is part of overall program implementation.

Intervention

One of the most compelling aspects of *Reading First* is the requirement that schools implement an Immediate Intensive Intervention program based on credible reading research. These programs are designed to provide focused, explicit reading instruction to struggling readers on their specific instructional level. Principal oversight of this aspect of reading implementation is crucial because effective intervention is, perhaps, the most powerful tool in the overall improvement of school-level reading achievement, while being the least understood from a practical standpoint. Thus, the PIQ includes items associated with the principal's role in intervention.

Method

Participants

Respondents on the PIQ were 585 secondary education school principals in the state of Florida who were currently receiving *Reading First* funding. The age range for the sample was 21 to 67 ($M=51.48$ $SD=7.46$), with approximately 73% of the sample female ($n=257$). Participants were 68% White, 25% African American, 6% Latino, and 1% identified themselves as Multi-racial. The highest level of education completed for 74% of the principals was at the Master's level, followed by 13% Specialists, 12% Doctoral, and 1% at the Bachelor's level. The average level of education experience for principals was 26.51 years ($SD=7.73$), with an associated average level of administrative experience at 12.58 years ($SD=6.86$). Principals reported a mean 5.77 years ($SD=5.01$) administrative experience at their current school, and 1.79 years ($SD=0.41$) implementing the *Reading First* initiative at their current school.

Additionally, a second sample of participants were concurrently administered the PIQ as part of a separate validation study. This cohort contained 53 principals in the state of Florida who were principals of schools supported by the North East Florida Education Consortium (NEFEC) that were implementing strategies set forth by the Florida Reading Initiative (13 school districts). Eighty-seven percent of principals were White ($n=46$), with the remaining participants identifying themselves as Black ($n=7$). The sample was predominantly female (71%), with 42% of the total sample having between 1 and 3 years of experience as a school administrator ($n=22$); 16% having 4–6 years of experience, 16% having 7–9 years of experience, 16% having 25 years or more experience, 7% having 22–24 years experience, and 3% having less than a year of experience. The highest level of experience for most principals was at the Masters level (94%; $n=49$) with only a few individuals completing only a Specialists ($n=2$) or Doctorate ($n=1$) degree. The sample was split ($n=638$) into two stratified groups to facilitate a baseline calibration testing of PIQ score measurement properties, with final model decisions cross-validated on the second sample. Both the calibration ($n=325$) and cross-validation ($n=313$) samples were equitable in their

demographic distributions (i.e., gender, ethnicity, years experience), stratified according to the weighted averages of the two initial sample pools.

Measure

The PIQ was a revised version of the 2003–2004 Principal Implementation Survey that was previously developed and validated using appropriate psychometric methods (Petscher and Nettles, unpublished manuscript). Although the Principal Implementation Survey retained acceptable levels of reliability and validity, the authors decided to revisit the content area of the items, which led to a re-writing and re-validation process. Subsequently, the Principal Implementation Questionnaire emerged from the definitions of each of the five proposed constructs of principal leadership and content from the Principal Implementation Survey.

Items were written by the authors, which were critiqued by a focus group of four content experts from the Florida Center for Reading Research. Out of this development, an initial pool of 105 items was considered for testing. All items were rated on a four-point Likert scale ranging from 1 (never/strongly disagree) to 4 (very frequently/strongly agree). The PIQ was designed to be used on a multidimensional basis, with interpretable mean scores across each scale, corresponding to the Likert range used for ratings. It was established a priori that scales did not necessarily need to be reduced to the same number of items since data reduction work may not lend itself to such a feature.

Procedure

The PIQ was distributed to all principals of *Reading First* schools in Florida by the Florida Center for Reading Research (FCRR) at Florida State University via the Progress Monitoring and Reporting Network (PMRN) maintained by the center. The PMRN is a data collection and reporting system through which *Reading First* schools both report reading data and receive reports on these data. Principals, reading coaches and teachers in *Reading First* schools use this system as the primary source of *Reading First* data regarding the state, districts, schools, classrooms and students.

Principals were prompted via email to log on to the PMRN and complete the survey instrument at the end of the school year. The PIQ Web-page on the PMRN was active for a period of one week. Prior to activation, principals were informed via email of the specific dates of the response period. Data derived from the survey were automatically compiled into a database formatted for empirical analysis.

The PIQ was subsequently given to the sample of NEFEC principals utilizing an online survey tool. Principals were prompted at the end of the school year to complete the survey and were given a 2-week timeframe for submission. Again, principals were informed via email of the specific dates of the response period and data derived from the survey were automatically compiled into a database for analysis.

Methods of Statistical Inquiry

Analysis of the 105 items was imperative to assess the parsimony and non-redundancy of the items, and was tested using SPSS 14. Internal consistency analyses tested the reliability

of the retained items from the data reduction process to ascertain the degree to which the items consistently measured similar constructs.

Confirmatory factor analysis using maximum likelihood estimation was conducted using LISREL (Jöreskog and Sörbom 2004) to test the validity of the item scores indicated by factors. Several iterations of the model were fit to the data in the sample based on theory and LISREL modification indices, fit indices, and standardized loadings.

Indices used to evaluate the goodness of fit for the models include: χ^2/df ratio, the root mean square error of approximation (RMSEA), the root mean residual (RMR), the normed and comparative fit indices (NFI, CFI), and the goodness of fit index (GFI). The NFI, CFI, and GFI are bound by a continuum of 0-to-1 with values ≥ 0.90 indicating acceptable fit (reference). RMSEA estimates below 0.10 often indicate acceptable fit, with values below 0.06 indicating adequate fit (Hu and Bentler 1999). Bollen (2005) provides a useful heuristic for RMSEA cutoffs by using the formula $(1 - \text{RMSEA})$ with values ≥ 0.90 indicating acceptable fit. The RMSEA has been widely utilized in reporting SEM model fit data due to its high sensitivity to misspecification in the model (Bentler 1992; Fan et al. 1999).

Results

Item Reduction—Calibration Sample

Data reduction of the initial 105 item pool indicated that numerous items could be removed while maintaining moderate to high levels of reliability. Multiple criteria were used in determination of retention of items: (1) item should have a corrected item-total correlation estimate ≥ 0.40 , (2) the Cronbach's alpha-if-deleted index for the item should be lower than the scale reliability, (3) the content of the item should be distinguishable from other items. Using these guidelines, a total of 58 items were initially retained for the five subscales: Leadership ($\alpha=0.90$, 19 items), Professional Development ($\alpha=0.85$, 13 items), Assessment ($\alpha=0.90$, 12 items), Intervention ($\alpha=0.79$, 11 items), and Core Reading Program ($\alpha=0.72$, 3 items). However, subsequent testing of this model indicated that several items needed to be re-examined for content. Upon inspection of the item content as well as modification indices provided by LISREL, it was determined that 15 additional items could be removed from the model while still retaining acceptable levels of internal consistency.

Subsequent testing of the reduced 43 item model indicated that further modifications could be made to the model based on standardized loadings, and modification indices. A final 28-item model was tested for both internal consistency and factorial validity. Results indicated that this model improved both parsimony and reliability of the constructs Leadership ($\alpha=0.89$, 8 items), Professional Development ($\alpha=0.81$, 2 items), Assessment ($\alpha=0.90$, 13 items), and Intervention ($\alpha=0.79$, 5 items). Overall scale reliability was ($\alpha=0.92$). Through the process of item reduction, all items pertaining to the Core Reading Program construct were deleted, improving the overall fit of the model.

The standard error of measure (SEM) was calculated to determine variation that would occur across testing for an individual principal. SEM estimates allow one to create a confidence band around the mean to predict the magnitude of expected departure from the mean across

multiple assessments (McDonald 1999). A summary of the final item means, standard deviations, and scale reliability are reported in Table 1.

Measurement Model Testing—Calibration Sample

Testing of the initial five-factor, 58 item model indicated poor fit of the model to the data $\chi^2(1,585, n=352)=5,353.21, p<0.01$, with comparable fit indices (RMSEA, RMR>0.10; NFI, CFI, GFI, <0.90; $\chi^2/df>3.0$) and standardized loadings corroborating misspecification of the model. Review of the modification indices suggested adding several paths from factors to items, which were studied in terms of content and theoretical parsimony. A total of 30 items in successive model iterations performed poorly in model based upon low standardized loadings, nominal amount unique variance explained by the factor, comparatively lower inter-item correlations with the rest of the grouped items within a factor. Additionally, a panel review of the Core Reading Program construct retained more programmatic rather than perceptual qualities, and did not group well with the rest of the latent domains as they indicate principal implementation of reading and was thus removed from the model. The final 4 dimensional, 28 item model was tested, with results indicating good model fit to the data $\chi^2(317, n=325)=998.56, p<0.05$; $\chi^2/df=3.34$, RMSEA=0.08 (1RMSEA=0.92), NFI=0.92, CFI=0.95, RMR=0.03, GFI=0.81. A summary of model iteration indices is presented in Table 2.

Review of the standardized structure coefficients (Table 3) indicated that all items maintained a moderate–high correlation with its proposed latent construct, and a significant level of explained variance of the items (R^2) were explained by the factors. Correlations among the PIQ subscales are presented in Table 4. Intercorrelations were low to moderate ranging from 0.15 between Intervention and Professional Development to 0.63 between Leadership and Assessment.

Cross-validation of Measurement Model

Descriptive statistics for the final measurement model (i.e., 4 factor–28 item) are reported in Table 1. Reported mean scores for this sample were slightly higher than that of the calibration sample; however, given the adjustments in the standard deviations and internal consistency estimates, the standard error of measure approached that of the original sample. This indicated that the range of variability for a principal's score was relatively constant. Additionally, estimates of internal consistency from the calibration sample were observed at a comparable magnitude for the cross-validation group. The most reliable scale was Assessment ($\alpha=0.88$) and the lowest observed estimate was for the Intervention domain ($\alpha=0.78$), again replicating findings from the initial sample.

Discriminant validity of factors was again tested using previously used indices as a measure of overall fit. Results indicated that the model fit the data well $\chi^2(305, n=313)=982.77, p<0.05$; $\chi^2/df=3.2$, RMSEA=0.081 (1-RMSEA=0.92), NFI=0.92, CFI=0.95, RMR=0.05, GFI=0.80. A summary of model fit indices is reported in Table 2. Structure coefficients and proportion of variance indices were sufficiently large with all correlations between items and proposed factors >0.50 ($z>2.0$; Table 3). Correlations among PIQ factors ranged from

0.20 (Professional Development–Intervention) to 0.65 (Leadership–Assessment), replicating findings from the previous samples (Table 4).

Construct Validity

Convergent validity was tested by correlating the PIQ subscales' mean scores with the LISI total scores. Results indicated that a moderate amount of shared variance was observed (0.29–0.61) between PIQ and LISI measured constructs. Gender and ethnic statistics were individually regressed on PIQ subscales to assess the degree of non-relationship (i.e., discriminant validity). The largest observed correlation for gender-PIQ was 0.19 (Professional Development) and 0.23 for ethnic status (Intervention). None of the correlations were large enough in magnitude to make a relational inference between the measured traits. Results are summarized in Table 5.

Discussion

This study was designed to explore a new way of evaluating principal behaviors in the implementation of effective reading programs to gain a better understanding of what practices were effective in the school environment. The development and validation of the PIQ was done in an attempt to provide the field with a formal measure of this important instructional leadership function. The resulting statistics from two independent samples provided appropriate psychometric evidence for the underlying theoretical content of the PIQ. This instrument could therefore serve as a foundation for investigating the relationship between principal implementation practices and student achievement in reading. Thus, the PIQ was designed to facilitate this line of research in an ongoing attempt to establish a platform for identifying effective principal practices, so that these practices could be measured, tracked and developed in school leaders. Relevant issues associated with the development of the PIQ are discussed below.

Removal of the Core Reading Program Construct

Although the use of an effective core reading program (CRP) was one of the five dimensions identified as essential to quality implementation of reading programs, this dimension did not emerge as a valid PIQ construct. The statistical elimination was due to extremely low levels of variability in responses, for this dimension, among principals completing the PIQ. It is suggested that one of the primary reasons for such low variability in the CRP dimension is that this is an area of reading program implementation where there is the least flexibility, and thus, little opportunity for site based differences. Principals had little impact on this dimension. That is, all *Reading First* schools must use core reading programs, so all of the principals responded positively to the items included in this dimension.

This lack of variability suggests that the use of a core reading program is one area of reading implementation where principals are uniform in their practice. This is somewhat intuitive because the implementation of a core reading program is the most straightforward of the dimensions being measured. It is also the easiest requirement to implement purely by executive order, needing little additional understanding, or follow up activity.

Value of the Instrument in Principal Education

The retained constructs of the PIQ represent an effective instrument in measuring principal practices where differences exist, presenting the opportunity to quantify levels of implementation effectiveness. This instrument was designed to measure constructs associated with the implementation of *Reading First* requirements. As this program is used in all of the 50 states, the District of Columbia, the Bureau of Indian Affairs, American Samoa, and the Virgin Islands, there is the potential for widespread relevance of the PIQ. However, additional research is needed to establish the utility of the instrument outside of Florida. Further, since the requirements of *Reading First* are based on a synthesis of modern reading research, the relevance of the PIQ in evaluating principal effectiveness in reading program implementation could potentially be useful in non-*Reading First* schools using these research-based principles.

References

- Bentler PM. On the fit of models to covariances and methodology. *Psychological Bulletin*. 1992; 112:400–404. [PubMed: 1438635]
- Bollen, K. Structural equation model examples; 2005, May; Presented at 2005 LISREL Summer Workshop; Chicago, IL.
- Bossert, ST. School effects. In: Boyan, NJ., editor. *Handbook of research on educational administration*. New York: Longman; 1988. p. 341–352.
- Bossert ST, Dwyer DC, Rowan B, Lee GV. The instructional management role of the principal. *Educational Administration Quarterly*. 1982; 18(3):34–64.
- Connor CM, Morrison FJ, Fishman BJ, Schatschneider C, Underwood P. The early years: Algorithm-guided individualized reading instruction. *Science*. 2007; 315(5811):464–465. [PubMed: 17255498]
- Ediger M. The principal and the reading curriculum. Monograph, ERIC# ED432006. 1999
- Ediger M. The school principal as leader in reading instruction. *Reading Improvement*. 2000; 37(1): 20–29.
- Fan X, Thompson B, Wang L. Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes. *Structural Equation Modeling*. 1999; 6:56–83.
- Florida Department of Education. Florida Reading First grant application: Reading first in Florida. Submitted to the United States Department of Education. 2002
- Hallinger P, Heck R. Reassessing the principal's role in school effectiveness: A review of empirical research, 1980–95. *Educational Administration Quarterly*. 1996; 32(1):5–44.
- Hallinger P, Heck RH. Exploring the principal's contribution to school effectiveness: 1980–1995. *School Effectiveness and School Improvement*. 1998; 9:157–191.
- Heck RH. School context, principal leadership and achievement: The case of secondary schools in Singapore. *The Urban Review*. 1993; 25(2):151–166.
- Hoy WK, Ferguson J. A theoretical framework and exploration of organizational effectiveness of schools. *Educational Administration Quarterly*. 1985 Spring;21(2):117–134.
- Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*. 1999; 6:1–55.
- Jöreskog, K.; Sörbom, D. LISREL 8.7 [computer software]. Chicago: Scientific Software International; 2004.
- Levine, DU.; Lezotte, LW. ERIC Document Reproduction Service No ED330032. Madison, WI: National Center for Effective Schools Research and Development; 1990. Unusually effective schools: A review and analysis of research and practice.
- McDonald, R. Test theory: A unified treatment. Mahwah, NJ: Erlbaum; 1999.

- National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction. Washington, DC: National Institute of Child Health and Human Development; 2000.
- No Child Left Behind Act of 2001. Pub L 107–110, 115 Stat 1425. 2002
- Perie, M.; Grigg, W.; Donahue, P. The nation's report card: Reading 2005 (NCES 2006-451) U S Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office; 2005.
- Pitner, NJ. The study of administrator effects and effectiveness. In: Boyan, NJ., editor. Handbook of research on educational administration. 1st. New York: Longman; 1988. p. 99-122.
- Sammons, P.; Hillman, J.; Mortimore, P. ERIC Document Reproduction Service No ED389926. London, England: Office for Standards in Education; 1995. Key characteristics of effective schools: A review of school effectiveness research.
- Snow, CE.; Barnes, WS.; Chandler, J.; Goodman, IF.; Hemphill, L. Unfulfilled expectations: Home and school influences on literacy. Cambridge, MA: Harvard University Press; 1991.
- Snow, C.; Burns, M.; Griffin, P., editors. Preventing reading difficulties in young children. Washington, DC: National Academy Press; 1998.
- Snyder KJ, Acker-Hocevar M, Snyder KM. Principals speak out on changing school work cultures. Journal of Staff Development. 1996 Winter;17(1):14–19.
- 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.
- U.S. Department of Education. Reading first program funding. 2006. n.dRetrieved from <http://www.ed.gov/programs/readingfirst/funding.html>
- Witziers B, Bosker RJ, Kruger ML. Educational leadership and student achievement: The elusive search for an association. Educational Administration Quarterly. 2003 Aug; 39(3):398–425.

Table 1
Descriptive statistics for final PIQ model

Factor	Mean	SD	SEM	α
Leadership ($n = 8$)	29.28 (34.26)	3.36 (4.13)	1.11 (1.60)	0.89 (0.85)
Professional development ($n = 2$)	7.17 (9.10)	1.12 (1.14)	0.48 (0.48)	0.81 (0.82)
Assessment ($n = 13$)	45.10 (55.19)	5.59 (5.13)	1.77 (1.78)	0.90 (0.88)
Intervention ($n = 5$)	14.79 (20.03)	3.40 (3.34)	1.63 (1.57)	0.77 (0.78)

SEM: standard error of measure

Table 2

Summary of model fit indices for CFA iterations

Model	Items	χ^2/df	RMSEA	90%CI ^a	RMR	NFI	CFI	GFI
Calibration								
1	58	4.6	0.122	(0.099, 0.124)	0.07	0.86	0.89	0.58
2	43	6.0	0.120	(0.119, 0.123)	0.05	0.88	0.90	0.59
3	38	3.5	0.085	(0.082, 0.089)	0.05	0.91	0.94	0.74
4	33	3.2	0.084	(0.080, 0.089)	0.03	0.91	0.92	0.79
5 ^b	28	3.1	0.080	(0.073, 0.082)	0.03	0.92	0.94	0.81
Cross-validation								
1	28	3.2	0.081	(0.075-0.082)	0.05	0.92	0.95	0.80

RMSEA: root mean square error of approximation, RMR: root mean residual, NFI: normed fit index, CFI:comparative fit index, GFI: goodness of fit index

^a Represents the 90% confidence interval for RMSEA

^b Final model retained

Table 3

Estimated structure coefficients for PIQ constructs

Item	Leadership	Professional development	Assessment	Intervention	R ²
13		0.97 (0.94)			0.95 (0.88)
14		0.71 (0.80)			0.50 (0.65)
16	0.74 (0.60)				0.55 (0.36)
17	0.72 (0.65)				0.52 (0.42)
18	0.76 (0.95)				0.57 (0.90)
19	0.74 (0.96)				0.54 (0.92)
22	0.71 (0.70)				0.51 (0.49)
23	0.68 (0.54)				0.46 (0.29)
24	0.66 (0.54)				0.44 (0.29)
29	0.66 (0.54)				0.43 (0.29)
30			0.52 (0.74)		0.27 (0.55)
31			0.52 (0.74)		0.27 (0.55)
32			0.50 (0.69)		0.25 (0.48)
57			0.59 (0.74)		0.35 (0.55)
64			0.66 (0.72)		0.44 (0.52)
65			0.56 (0.76)		0.32 (0.58)
66			0.78 (0.78)		0.61 (0.61)
67			0.74 (0.75)		0.55 (0.56)
68			0.75 (0.75)		0.56 (0.56)
69			0.59 (0.67)		0.35 (0.45)
70			0.71 (0.75)		0.51 (0.56)
71			0.71 (0.76)		0.51 (0.58)
72			0.71 (0.63)		0.50 (0.40)
80				0.63 (0.54)	0.40 (0.29)
81				0.68 (0.67)	0.47 (0.45)
82				0.70 (0.57)	0.49 (0.33)
83				0.61 (0.80)	0.37 (0.64)
86				0.51 (0.56)	0.26 (0.31)

Estimates for constructs are statistically significant at $\alpha > .05$ level

Table 4
Correlations for PIQ dimensions

Factor	Leadership	Professional development	Assessment	Intervention
Leadership	1.00			
Professional development	0.45 (0.46)	1.00		
Assessment	0.63 (0.65)	0.40 (0.43)	1.00	
Intervention	0.31 (0.53)	0.15 (0.20)	0.48 (0.52)	1.00

Table 5
Construct validity results for PIQ—cross validation sample

Scale	Convergent validity	Discriminate validity	
		Gender	Ethnicity
Leadership	0.61	0.18	0.13
Professional development	0.54	0.19	0.04
Assessment	0.68	0.15	0.05
Intervention	0.78	0.12	0.23