

Analysis of the Reliability and Validity of a Mentor's Assessment for Principal Internships

Glenn L. Koonce  
Regent University

Michael D. Kelly  
Virginia Tech

August 2014

Abstract

In this study, researchers analyzed the reliability and validity of the mentor's assessment for principal internships at a university in the Southeast region of the United States. The results of the study yielded how trustworthy and dependable the instrument is and the effectiveness of the instrument in the current principal preparation program. Study results were analyzed using reliability requirements for instruments utilized in obtaining national accreditation and to provide suggestions for program improvement. The instrument was reviewed by a panel of experts from areas outside of the research population to establish content and face validity. Internal consistency and reliability were measured using Cronbach's alpha. A total of 229 candidate internship scores were used. The reliability test resulted in an overall alpha of .949. The results obtained in this study indicate the instrument has a very high level of validity as well as reliability.

Keywords: Assessment Reliability, Program Accreditation, CAEP/TEAC, Principal Internship Assessment, Program Improvement, School Leadership Preparation.

*Even after decades of use, designing and implementing worthwhile administrative internships remains a work in progress* (Lehman, 2013).

National accreditation for educational leadership preparation programs has recently been approved by the United States Department of Education (USDOE) and the Council for Higher Education Accreditation (CHEV) for one accreditor, the Council for the Accreditation of Educator Preparation (CAEP). One pathway for accreditation under CAEP is the *Inquiry Brief (IB)*. Although standards and principles of quality are key elements in the *IB*, all assessments adopted must be valid and reliable interpretations of the evidence (Teacher Education Accreditation Council, 2012). One instrument utilized by a university in the Southeast region of the United States to evaluate the principal internship is the Principal Internship Mentor's Assessment (PIMA). This study will address reliability and validity for the PIMA.

Reliability measures are designed to yield confidence that an assessment is trustworthy and dependable indicating to what degree the assessment measures what it is designed to measure. There is a difference in the relationship between reliability and validity. Arcy, Jacobs, Razavieh, and Sorensen (2006) posit that, "A measuring instrument can be reliable without being valid; but it could not be valid unless it is first reliable" (p. 256). Arcy, et. al., provided as an example that intelligence may be studied by determining the circumference of the head. Consistency in the measurements of the head from time to time may be very similar (reliable), but doing this would not yield valid inferences about intelligence because there is no correlation of measurements of the head with any other criteria of intelligence or theory of intelligence. The result is very reliable scores over time but not a sufficient condition for valid interpretation of test scores.

Scores for this study are derived from one internship assessment (PIMA) and aligned with common elements from the six Interstate School Leaders Licensure Consortium (ISLLC) Standards. Williams and Alawiye (2014) report that, "there is a huge increase in adopting the ISLLC Standards

in higher education as a pre-requisite for new leaders in the school system” (p. 2). Consistency in measuring elements of the ISLLC Standards would yield inferences about principal preparation because of the correlation between the ISLLC Standards and the principal internship assessment,  $r=0.872$ ,  $n=59$ ,  $p=0.01$  (Koonce and Causey, 2011).

### **Definition of terms**

#### **Reliability**

Reliability refers to the consistency of a measure. If we attain the same result repeatedly the measure is considered reliable. For example, “if an assessment is designed to measure a trait (such as introversion), then each time the assessment is administered to a subject, the results should be approximately the same” (Cherry, 2013, p. 1). It is not possible to calculate reliability exactly, but it can be estimated in various ways.

#### **Cronbach’s Alpha**

“Cronbach’s alpha is a model of internal consistency reliability based on the average inter-item correlation of an instrument” (Rovai, Baker, and Ponton, 2014, p. 545).

#### **PIMA**

The instrument for measuring student success in the internship for this study is the Principal Internship Mentor’s Assessment (PIMA) (Arroyo, Koonce, & Hanes, 2008). The PIMA is a 24-item Likert-type scale instrument derived from the ISLLC standards. There are four items per standard (total of six standards) with each item rated on a 5-point scale ranging from “fails to address/no evidence of knowledge, understanding, and/or application” to “very specific/convincing evidence of knowledge, understanding, and/or application” (Hessel & Holloway, 2002, p.24). All assessment items are taken from “Components of Professional Practice for School Leaders” (p.27). Each sub-score on the PIMA are directly linked to a correlating ISLLC Standard.

### **Purpose**

This study looks at one important component of a good assessment, reliability. If a measure were not consistent and produced different results every time it would not be considered and certainly would not meet rigor for attaining national accreditation for an educational leadership preparation program. For this study, the reliability for the PIMA is investigated.

### **Rationale and Significance of the Study**

The driving force for conducting this study is continuing national accreditation through TEAC/CAEP for a university in the Southeast region of the country. National accreditation is required by the state in order for program completers to be awarded the administration and supervision endorsement. TEAC/CAEP accredits programs, not an administrative unit of the institution such as the School of Education (SOE). Accreditation for the educational leadership preparation program within the SOE must “affirm their goal to prepare competent, caring, and qualified leaders for the schools” (TEAC, 2012, p. 28). Program faculty members collaborate on the claims made about their graduates and provide evidence to support these claims. Certain components regarding claims must address: professional knowledge; strategic decision-making; caring and effective leadership skills; learning how to learn; multicultural perspectives and accuracy; and technology.

TEAC (2012) requires faculty to provide evidence by assessing what candidates have learned. Examples of this evidence includes: (a) grades and grade point averages; (b) scores on standardized tests, for example, the School Leaders Licensure Assessment (SLLA); (c) ratings of portfolios; (d) case studies; (e) surveys, like an alumni survey; and (f) ratings by cooperating mentors (usually principals) of clinical experiences, like the principal internship. It is Item (f) that this study addressed. The principal internship is the capstone course/activity in the university

leadership preparation program in this study. The assessment for the internship is mostly comprised of an instrument developed in house by the faculty. The PIMA is an example. TEAC/CAEP requires that such assessments “meet the accepted research standards for reliability and validity” (p. 30). Evidence must be sufficient, clear, and consistent. It must be supported with “local evidence about the trustworthiness, reliability, and validity of the assessment (p. 30).

The Missouri Standards for the Preparation of Educators (MoSPE) (Missouri Department of Elementary and Secondary Education, 2013) directed in their program standards that assessments be fair, valid, and reliable. Most educators assume that when they see a principal internship measure that it is fair, valid, and reliable. As the review of literature will reveal, this is not always the case. There is little in the literature on internship assessment instrument design and often, instruments have not been established as valid and reliable by the program faculty. More studies of this nature are needed from the field to inform current and future practice.

The faculty have used the PIMA since 2006 providing mentoring principals in the field with a tool to evaluate how candidates have performed. PIMA was designed based on a review of the literature on the ISLLC Standards. The rubric for the PIMA provides a common language for assessment by the mentoring principal, university supervisor, and the intern. The PIMA is a major component for assigning a final grade for the required internship. The reliability for the PIMA has not been fully addressed.

### **Literature Review**

In response to the standards movement, many principal preparation programs have embedded the ISLLC Standards in the curriculum to provide the necessary knowledge for effective school leadership. Knowledge of the curriculum presented in coursework alone will not suffice. The ability to apply that knowledge comes from candidates in the program successfully completing

the principal internship. (Barton, 2013). This internship has been described as exposure to “situations requiring the application of acquired skills, knowledge, and problem solving strategies within authentic settings” and “higher levels of involvement on the part of interns resulting in higher perceived levels of knowledge” (p. 2). Interns themselves report that “increased involvement in leadership activities resulted in more learning” (p. 2). With national accreditation comes the requirement that programs are accountable for providing real world practice in authentic settings for student learning (Teacher Education Accreditation Council, 2012). The review of literature focused on assessment of interns as well as use of Cronbach’s alpha.

### **Assessment of Interns**

A report from the Southern Regional Education Board (SREB) (ND) stated that although many university leadership preparation programs have recognized the need for, “high-quality” (p. 2), they have not fully addressed “one of the most serious problems-the quality of internships” (p. 2). The report was based on a survey of 61 programs in the 16 state SREB. A key finding was that the evaluation of candidate performance during their internship “lack a high degree of rigor” (p. 6). Of those programs that participated in the survey, only 45% were considered to have internship evaluations that are rigorous, valid, and consistent. No in-depth discussion on a tool to assess the internship or validity and reliability of an assessment was found.

Are universities conducting rigorous evaluations of their candidate performances during the internship? Other than a few books that focus on the principal internship: *School Leader Internship: Developing, Monitoring, and Evaluating Your Leadership Experience* (Martin, Wright, Danzig, Flannery, & Brown, 2012); *The Standards-Based Administrative Internship:*

*Putting the ISLLC Standards into Practice* (Hackman, Schmitt-Oliver, & Tracey, 2005); and *The School Administrator Internship Handbook: Leading, Mentoring, and Participating in the Internship Program* (Capasso & Darish, 2000), the review of literature did not produce any reliable assessments of the internship by a mentoring principal in the field. The books only provide sample assessments, mostly included in a portfolio style, with no validity or reliability measures.

SREB (ND) reported 95% of the evaluations of the principal internships are based upon portfolios. Since no samples of portfolios or their scoring rubrics were found in the literature review, it is assumed that preparation programs keep those assessments in-house, use them for grading of candidates and accrediting purposes, but do not provide copies or their own research related to these tools in principal program preparation literature. It is not possible to make a defensible argument for or against the rigor, validity, and reliability in principal internship assessment since little to no criteria or samples are available for making judgments. National accreditation for preparation programs may require more of these assessments to be reviewed and studied.

The literature produced one sample of an internship assessment that did not use quantifiable data in which to assess results or any measures for reliability. The Illinois Principal Preparation Program (ND) *Sample Worksheet for Candidate Mastery of the SREB Critical Success Factors and Activities* lists 13 SREB Critical Success Factors for the internship and includes a space for interns to list their internship experiences for each critical success factor. In addition, there is a space for each factor that asks the intern to, “Describe Assessment Used to Determine Proficiency”. No assessment models were provided. It is assumed that this document is meant to be a template for all state leadership preparation programs to meet regulations and/or



policy. It appears to leave the data-gathering instrument, which includes validation and reliability up to each program. No instruments from Illinois were found in the literature review.

Only 3 university programs reviewed in the literature indicated a scale for rating an intern by a mentoring principal. Texas' Lamar University's (2011) Principal Internship Supervisor Evaluation (PISE) is made up of 9 items based on the Educational Leadership Constituent Council (ELCC) Standards and rated by the mentoring principal on a scale of 3=Exemplary, 2=Proficient, 1=Minimally Proficient, and 0=Unacceptable. A review of the Lamar University program documents did not reveal any measures for validity or reliability. The University of South Dakota (2014) competencies for the administrative internship were also established using the ELCC standards and includes three assessments (mentor, university supervisor, and intern self-assessment). Although they have a formal document describing the principal internship program for a master's, specialist', and doctorate' degrees, there is no mention of validity or reliability. The instrument used for all three assessments is composed of 27 items on a four point scale: Area Needing Further Improvement; Area of Moderate Strength; Area of Significant Strength; and Not Observed/Applicable. No measures of validity or reliability were found in the Texas A&M University-Kingsville (2012) *Principal Internship Handbook* for the Mentor Summative Evaluation of Intern form. The instrument is composed of 9 items with ratings ranging from: 5=Clearly Outstanding, 4= Exceeds Expectations, 3=Meets Expectation, 2=Below Expectations and 1=Unsatisfactory.

The literature review revealed one TEAC Inquiry Brief related to teacher educator preparation. The *Montana State University* (2010) *TEAC Inquiry Brief for Teacher Education* includes 2 assignments in the internship (student teaching) that requires candidates to submit: 1) a fall evaluation from the school administrator responsible for evaluating interns as a formative

measure and the same evaluation (summative) submitted in the spring at the conclusion of the internship. Based on the Interstate Teacher Assessment and Support consortium (INTASC) standards, the measure used is the *Internship Performance Evaluation Instrument* (pgs. 143-149). The instrument is made up of 13 items scored on a scale of 4:1= Unacceptable, 2=Less Competent, 3=Competent, and 4=Highly Competent. The big difference is that this is an assessment for student teacher interns and not educational leadership interns. Montana State University's Inquiry Brief noted on page 44 that the internship assessment has not been evaluated with respect to validity and reliability, thus providing some credence to the fact that many preparation programs (teacher preparation and leadership preparation) use measures that have not been determined to be valid or reliable.

### **Use of Cronbach's Alpha**

Although a review of the literature did not produce any studies regarding reliability for principal internship assessment, Gaudrea's (2008) research provided evidence leading to a better understanding of how engagement during a principal internship relates to the readiness for school leadership. It incorporated a Cronbach alpha application. Using summated Likert scales from the National Association of Secondary School Principals Assessment Center's Skills and Behavior Indicators as a measure of readiness for school leadership an instrument (questionnaire), with scales, was designed. The questionnaire contained 38 items with ratings for 3 different responses per item. Item 1) "Domains" with ratings from 1-4 included: "Engagement in internship"; "Institutional Support for the Internship"; "Relevance of internship to school-level leadership"; and "Quality of field supervision". Item 2 "Association" with ratings 1-4 included: "very weak"; "weak"; "strong"; "very strong". Item 3) Clarity with ratings 1-3 included: "very unclear, delete"; "somewhat clear, revise"; and "clear, leave as written" (p. 50). Reliability of the

three scales was assessed with Cronbach's alpha: .84, .89, and .90 respectively. Gaudrea found that all variables had an alpha greater than .80, however, "the low *N* contributed to the questionable reliability of the results" (p. 50), *N*=44.

Riggs, Verdi, & Arlin (2009) also used Cronbach alphas in their study of the California Teacher Performance Assessment. The study utilized quantitative and qualitative approaches to assess if the data possess adequate psychometric properties that include internal reliability. Quantitative results included descriptive statistics regarding subject specific pedagogy, designing instruction, assessing learning, and culminating teaching experience. For internal consistency, Cronbach alphas were calculated on each set of items by task. All items were entered across all dimensions as though they were measuring a single construct. Based on this single measure approach Alphas were very high (.95 for subject specific pedagogy, .95 for Designing Instruction, .95 for Assessing Learning, and .96 for the Culminating Teaching Experience. "These values indicate that all items can be used as a measure of a single, global construct" (p. 22).

Another use of Cronbach Alpha was found in a presentation (Murphy, 2011) to the Council of Chief State School Officers (CCSSO) for building a principal evaluation system. Although not linked to internship assessment, principal evaluation assessment has been an area of change over the last few years. Murphy designed a pilot study on the evaluation of practicing principals using the Vanderbilt Assessment of Leadership in Education (VAL-ED). In providing background information regarding the VAL-ED (i.e.: the conceptual model, leadership behavior framework, core components, key processes, and support from a grant from the Wallace Foundation) the psychometric properties were reviewed that included reliability. For the 108

item-form scales, the Cronbach's Alpha was .92. Feedback from the pilot-study provided potential areas of growth for principals and key targets for professional development.

In summary, the literature is terse when addressing quantitative principal internship assessment instruments; even less (almost none) when seeking measures for reliability such as utilizing Cronbach's alpha. This makes the current study valuable to the principal preparation field and to the program addressed in this study.

## **Methodology**

### **Participants**

A total of two hundred and twenty-nine (229) educational leadership internship completers participated in this research over a four year period from one university Educational Leadership Preparation Program in southeast Virginia. The participants were uncompensated and were not interviewed, tested or surveyed beyond the normal program requirements. All participants were licensed and experienced educators (minimum three years teaching experience) prior to participating in the internship. The participant group makes up a purposeful sample consisting of all students that completed the internship between September 2009 and August 2013.

### **Instrument**

The Principal Internship Mentor's Assessment (PIMA) is a 24-item Likert scale assessment derived from the ISLLC Standards that is divided into six distinct categories based on the ISLLC Standards. The instrument was initially developed in 2006 as a way for program faculty to evaluate student success during the principal internship experience.

There are four items per standard with each item being rated on a 5-point Likert-type scale: 0 – fails to address/no evidence of knowledge, understanding, and/or application; 1 –

vague/skeletal evidence of knowledge, understanding, and/or application; 2 – less detail/specific evidence of knowledge, understanding, and/or application; 3 – detailed/specific evidence of knowledge, understanding, and/or application; and 4 – very specific/convincing evidence of knowledge, understanding, and/or application.

The four items per standard are taken from “Components of Professional Practice for School Leaders” (Hessel & Holloway, 2002, p.27). The assessment reflects the 24 Components of Professional Practice for School Leaders derived directly from the ISLLC standards. Each item on the assessment is directly linked to one of the six broad ISLLC Standards. Below are the broad categories as well as the four individual items under each category raters used to evaluate graduate students during the internship experience. The stem provided for each to each item is, “please rate the intern regarding the following standard using the scale below”.

### ISLLC Standard 1: The Vision of Learning

A school administrator is an educational leader who promotes the success of all students by facilitating the development, articulation, implementation, and stewardship of a vision for learning that is shared and supported by the community.

- Component 1a: Developing the Vision
- Component 1b: Communicating the Vision
- Component 1c: Implementing the Vision
- Component 1d: Monitoring and Evaluating the Vision

### ISLLC Standard 2: The Culture of Teaching and Learning

A school administrator is an educational leader who promotes the success of all students by advocating, nurturing, and sustaining a school culture and instructional program conducive to student learning and staff professional growth.

- Component 2a: Valuing Students and Staff
- Component 2b: Developing and Sustaining the Culture
- Component 2c: Ensuring an Inclusive Culture
- Component 2d: Monitoring and Evaluating the Culture

### ISLLC Standard 3: The Management of Learning

A school administrator is an educational leader who promotes the success of all students by ensuring management of the organization, operations, and resources for a safe, efficient, and effective learning environment.

- Component 3a: Making Management Decisions to Ensure Successful Teaching and Learning
- Component 3b: Developing Procedures to Ensure Successful Teaching and Learning
- Component 3c: Allocating Resources to Ensure Successful Teaching and Learning
- Component 3d: Creating a Safe, Healthy Environment to Ensure Successful Teaching and Learning

#### ISLLC Standard 4: Relationships with the Broader Community to Foster Learning

A school administrator is an educational leader who promotes the success of all students by collaborating with families and community members, responding to diverse community interests and needs, mobilizing community resources.

- Component 4a: Understanding Community Needs
- Component 4b: Involving Members of the Community
- Component 4c: Providing Opportunities for the Community and School to Serve Each Other
- Component 4d: Understanding and Valuing Diversity

#### ISLLC Standard 5: Integrity, Fairness and Ethics in Learning

A school administrator is an educational leader who promotes the success of all students by acting with integrity, with fairness, and in an ethical manner.

- Component 5a: Demonstrating a Personal and Professional Code of Ethics
- Component 5b: Understanding One's Impact on the School and Community
- Component 5c: Respecting the Rights and Dignity of All
- Component 5d: Inspiring Integrity and Ethical Behavior in Others

#### ISLLC Standard 6: The Political, Social, Economic, Legal, and Cultural Context of Learning

A school administrator is an educational leader who has the knowledge and skills to promote the success of all students by understanding, responding to, and influencing the larger political, social, economic, legal and cultural contexts.

- Component 6a: Operating Schools on Behalf of Students and Families
- Component 6b: Communicating Changes in Environment to Stakeholders
- Component 6c: Working Within Policies, Laws, and Regulations
- Component 6d: Communicating with Decision-Makers Outside the School Community

## **Data Collection and Statistical Analysis**

### **Data Collection**

Data was collected through the mentoring principals' completion of the Principal Internship Mentor's Assessment (PIMA) for each graduate student completing the principal internship. The PIMA assessment was administered online using SurveyMonkey™. The final date for collection of data was August 31, 2013. Data was then organized into an Excel spreadsheet entered into the *Statistical Package for the Social Sciences* (SPSS).

### **Validity**

After the initial development of the PIMA assessment in 2006, the assessment was then pilot-tested in the 2006-2007 academic year with four mentoring principals and internship students. The feedback from the pilot testing was reviewed, and adjustments to the instrument were made. The items and the instrument were next reviewed for content validity by the program faculty as part of the continuous improvement plan for the program in the fall of 2008. Feedback from the faculty resulted in additional adjustments to the instrument. One adjustment was to use less verbiage and link all items to a common repeated theme. The more time principals spent on the instrument the less time they would devote to completing the assessment. In addition, the whole look of the instrument was streamlined.

During the summer of 2009 an educational focus group was formed in order to provide feedback on the instrument, and help further establish content validity. The Educational Focus Group (Cannizzaro, 2007) provided feedback on the form and confirmed its content validity since participants were practicing experts in the field. To address inter-rater reliability, 4 sets of two raters used the form and discussed the outcomes in the Focus Group and the four teams of practitioners rated the intern similarly (Cannizzaro, 2007). Subjective scoring (Inter-rater-

reliability/consistency between tests) is helped when usable guidelines for scoring are developed such as the scoring rubric for the PIMA.

### Reliability

Rovai, et. al. (2012) recommended measuring internal consistency and reliability using Cronbach’s alpha. Each of the six distinct categories based on the ISLLC Standards were individually tested for reliability based on the responses to the four sub-questions under each standard. Then the entire response set was also tested to determine overall reliability. Reliability tests resulting in an alpha of .7 are generally accepted as having high reliability (Rovai, Baker & Ponton, 2012, p. 385). Cronbach’s alpha reliability coefficient generally ranges between 0 and 1. However, there is actually no lower limit to the coefficient. The closer Cronbach’s alpha coefficient is to 1.0 the greater the internal consistency of the items in the scale. George and Mallery (2003) provide the following rules of thumb: “\_ > .9 – Excellent, \_ > .8 – Good, \_ > .7 – Acceptable, \_ > .6 – Questionable, \_ > .5 – Poor, and \_ < .5 – Unacceptable” (p. 231).

### Findings

#### ISLLC Standard 1: The Vision of Learning

For responses to the intern’s performance related to ISLLC Standard 1, 206 cases of the possible 229 were included in the analysis (Table 1). Students who did not have a complete data set were excluded. Cronbach’s alpha for the 206 of 229 items were .945 (Table 2), which represents an excellent correlation between items. The instrument for ISLLC Standard 1 related to the vision of learning can be deemed highly reliable.

Table 1

Case Processing Summary			
		N	%
Cases	Valid	206	90.0
	Excluded <sup>a</sup>	23	10.0



Total	229	100.0
-------	-----	-------

a. Listwise deletion based on all variables in the procedure.

Table 2

Cronbach's Alpha	N of Items
.945	4

### **ISLLC Standard 2: The Culture of Teaching and Learning**

For responses to the intern's performance related to ISLLC Standard 2, 222 cases of the possible 229 were included in the analysis (Table 3). Students who did not have a complete data set were excluded. Cronbach's alpha for the 222 of 229 items were .880 (Table 4), which represents a good correlation between items. The instrument for ISLLC Standard 2 related to the culture of teaching and learning can be deemed highly reliable.

Table 3

		N	%
Cases	Valid	222	96.9
	Excluded <sup>a</sup>	7	3.1
	Total	229	100.0

a. Listwise deletion based on all variables in the procedure.

Table 4

Cronbach's Alpha	N of Items
.880	4

**ISLLC Standard 3: The Management of Learning**

For responses to the intern’s performance related to ISLLC Standard 3, 201 cases of the possible 229 were included in the analysis (Table 5). Students who did not have a complete data set were excluded. Cronbach's alpha for the 201 of 229 items were .868 (Table 6), which represents a good correlation between items. The instrument for ISLLC Standard 3 related to the management learning can be deemed highly reliable.

Table 5

<b>Case Processing Summary</b>			
		N	%
Cases	Valid	201	87.8
	Excluded <sup>a</sup>	28	12.2
	Total	229	100.0

a. Listwise deletion based on all variables in the procedure.

Table 6

<b>Reliability Statistics</b>	
Cronbach's Alpha	N of Items
.868	4

**ISLLC Standard 4: Relationships with the Broader Community to Foster Learning**

For responses to the intern’s performance related to ISLLC Standard 4, 191 cases of the possible 229 were included in the analysis (Table 7). Students who did not have a complete data set were excluded. Cronbach's alpha for the 191 of 229 items were .840 (Table 8); though this is the lowest reliability score for the instrument, the alpha still represents a good correlation between items. The instrument for ISLLC Standard 4 related to the developing relationships with the broader community can be deemed highly reliable.

Table 7

<b>Case Processing Summary</b>			
		N	%
Cases	Valid	191	83.4
	Excluded <sup>a</sup>	38	16.6
	Total	229	100.0

a. Listwise deletion based on all variables in the procedure.

Table 8

<b>Reliability Statistics</b>	
Cronbach's Alpha	N of Items
.840	4

### **ISLLC Standard 5: Integrity, Fairness and Ethics in Learning**

For responses to the intern’s performance related to ISLLC Standard 5, 224 cases of the possible 229 were included in the analysis (Table 9). Students who did not have a complete data set were excluded. Cronbach's alpha for the 224 of 229 items were .903 (Table 10), which represents an excellent correlation between items. The instrument for ISLLC Standard 5 related to integrity, fairness, and ethics in learning can be deemed highly reliable.

Table 9

<b>Case Processing Summary</b>			
		N	%
Cases	Valid	224	97.8
	Excluded <sup>a</sup>	5	2.2
	Total	229	100.0

a. Listwise deletion based on all variables in the procedure.

Table 10

Cronbach's Alpha	N of Items
.903	4

**ISLLC Standard 6: The Political, Social, Economic, Legal, and Cultural Context of Learning**

For responses to the intern’s performance related to ISLLC Standard 6, 186 cases of the possible 229 were included in the analysis (Table 11). Students who did not have a complete data set were excluded. Cronbach's alpha for the 186 of 229 items were .874 (Table 12), which represents a good correlation between items. The instrument for ISLLC Standard 6 related to the political climate can be deemed highly reliable.

Table 11

		N	%
Cases	Valid	186	81.2
	Excluded <sup>a</sup>	43	18.8
	Total	229	100.0

a. Listwise deletion based on all variables in the procedure.

Table 12

Cronbach's Alpha	N of Items
.874	4

**ISLLC Standard 1-6 Combined**

As a final test of reliability, a combined analysis of all responses on the instrument was run across all six subcategories. For responses to the each intern’s performance related to ISLLC Standard s 1-6, 157 cases of the possible 229 were included in the analysis (Table 13). Students who did not have a complete data set were excluded. Cronbach's alpha for the 157 of 229 items were .949 (Table 14), which represents an excellent correlation between all items across the entire instrument. The instrument for ISLLC Standards 1-6 related to the political climate can be deemed highly reliable.

Table 13

<b>Case Processing Summary</b>			
		N	%
Cases	Valid	157	68.6
	Excluded <sup>a</sup>	72	31.4
	Total	229	100.0

a. Listwise deletion based on all variables in the procedure.

Table 14

<b>Reliability Statistics</b>	
Cronbach's Alpha	N of Items
.949	24

**Conclusions and Recommendations**

Through a thorough analysis of both the validity and reliability of the Principal Internship Mentor’s Assessment (PIMA), the researchers are able to determine that the instrument itself is both valid and reliable for its intended purpose. The validity of the instrument has been verified

using a variety of methods, including investigating face validity, content validity, and making use of an independent panel of experts.

The instrument has also been deemed to be extremely reliable, with an overall alpha value of .949. Each independent sub-group on the instrument also had alpha values that ranged from good (.800) to excellent (.999). The alignment of the instrument to national standards undoubtedly played a significant role in producing the strong results found in this study.

In light of these findings, the researchers make the following three recommendations for future use/study:

1. Other programs that use the ISLLC standards as a foundation for their programs are encouraged to use a similar model for evaluating internship progress in the field. By doing so, programs can also correlate the internship experience to student results on the School Leaders Licensure Assessment (SLLA), and other assessments as well. This could build a broader data base upon which improvement in leadership preparation can emerge and rigor for national accreditation can be met.
2. Finally, the current program should consider conducting a pre- and post- PIMA for students in the educational leadership program. Currently the PIMA is only administered as a post-experience assessment. The program would benefit greatly from data on students that could be collected as they enter the program as well. Student growth in the program could also be more easily measured as well.

## References

- Arcy, D., Jacobs, L., Razavieh, A., and Sorensen, C. (2006). *Introduction to research in education*. Belmont, CA. Thomson Wadsworth.
- Arroyo, A., Koonce, G., & Hanes, J. (2008). *Inquiry brief: A case for TEAC accreditation of the Educational Leadership, Educational Specialist, and K-12 School Leadership programs*. Virginia Beach, VA. Regent University School of Education.
- Barton, L. (2013). Knowledge of effective educational leadership practices. *International Journal of Educational Leadership Preparation*, 8(1).
- Cannizzaro, S. V. (2007). *Executive Summary: Focus Group of Practitioners in Educational Leadership*. Paper presented at Regent University, School of Education, Virginia Beach, VA.
- Capasso, R. & Daresh, J. (2001). *The school administrator internship handbook: Leading, mentoring, and participating in the internship program*. Thousand Oaks, CA: Corwin Press.
- Cherry, K. (2013). What is reliability? The importance of consistency in psychometrics. *About.com: Psychology*, Retrieved from: <http://psychology.about.com/od/researchmethods/f/reliabilitydef.htm>
- Gaudreau, P. (2008). *The effects of engagement in an internship on readiness for school leadership*. Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University.
- George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update (4<sup>th</sup> ed.)*. Boston: Allyn & Bacon.

Running Head: Reliability and Validity of PIMA

Hackman, D., Schmitt-Oliver, D., & Tracey, J. (2005). *The standards-based administrative internship: Putting the ISLLC standards into practice*. Lanham, MD: Rowman & Littlefield.

Hessel, K. and Holloway, J. (2002). *A framework for school leaders: linking the ISLLC standards to practice*. Princeton, NJ. Educational Testing Services.

Illinois Principal Preparation Program. (ND). Retrieved from: [http://www.isbe.net/prep-l/pdf/prin\\_prep\\_app.pdf](http://www.isbe.net/prep-l/pdf/prin_prep_app.pdf)

Lamar University (2011). Principal Internship Supervisor Evaluation. Retrieved from: <http://dept.lamar.edu/education/ncate-ir/data/PISE.pdf>

Lehman, L. (2013). Principal Internships in Indiana: A promising or perilous experience. *International Journal of Educational Leadership Preparation*, 8(1).

Martin, G., Wright, W., Danzig, A., Flannery, D., & Brown, F. (2012). *School leader internship: Developing, monitoring, and evaluating your leadership experience*. New York, NY. Taylor & Francis.

Montana State University. (2010). *TEAC inquiry brief for teacher education*. Retrieved from: [http://www.montana.edu/nptt/accreditation/docs/NPTT\\_TEAC.pdf](http://www.montana.edu/nptt/accreditation/docs/NPTT_TEAC.pdf)

Missouri Department of Elementary and Secondary Education (2013). *The Missouri standards for the preparation of educators*. Retrieved from: <http://dese.mo.gov/sites/default/files/MoSPEStandards.pdf>

Murphy, J. (2011). *Principal evaluation: The national story*. PowerPoint presented at the April 2011 meeting of the Council of Chief State School Officers. Retrieved from: [https://www.google.com/search?sourceid=navclient&aq=&oq=Principal&ie=UTF-8&rlz=1T4ADRA\\_enUS469US472&q=principal+evaluation%3a+the+national+story+cc](https://www.google.com/search?sourceid=navclient&aq=&oq=Principal&ie=UTF-8&rlz=1T4ADRA_enUS469US472&q=principal+evaluation%3a+the+national+story+cc)



[sso+april+2011+joseph+murphy&gs\\_l=hp..5.015j4113.0.0.0.9103.....0.xZHASZpsg9](#)

[A](#)

- Riggs, M., Verdi, M. & Arlin, P. (2009). A local evaluation of the reliability, validity, and procedural adequacy of the Teacher Performance Assessment Exam for teaching credential candidates. *Issues in Teacher Education, 18*(1).
- Rovai, A., Baker, J. and Ponton, M. (2014). *Social science research design and statistics: A practitioner's guide to research methods and IBM SPSS analysis*. (1<sup>st</sup> ed.) Chesapeake, VA. Watertree Press LLC.
- Southern Regional Education Board. (ND). *The principal internship: How can we get it right?* Atlanta, GA. Author.
- Teacher Education Accreditation Council (TEAC). (2012). *Guide to accreditation*. Washington, DC. Author.
- Texas A&M University-Kingsville. (2012). *Principal internship handbook*. Retrieved from: <http://www.tamuk.edu/cehp/edlc/docs/edadmin-handbook.pdf>
- University of South Dakota (2014). *Graduate Internship Program: EDAD 794 (M.A.)-EDAD 894S (Ed.S.)-EDAD 894D (Ed.D.) PK-12 School Principal*. Retrieved from: <http://www.usd.edu/education/educational-administration/upload/Internship-Packet-PK12-Prin10-13.pdf>
- Williams, H & Alawiye, O. (2014). Performance assessment of residency principal interns on Washington standard 5 knowledge and skills. *National Forum of Educational Administration and Supervision Journal, 32*(4).