Assessing ISLLC-Based Dispositions of Educational Leadership Candidates*

Dorothy Rea
Cecil F. Carter
Judy R. Wilkerson
Thomas Valesky
William Lang

Abstract

The Council of Chief State School Officers (CCSSO), through the Interstate School Leaders Licensure Consortium (ISLLC), developed standards for the knowledge, skills, and dispositions necessary for effective practice by educational leaders (CCSSO, 1996).) These standards provide a viable content domain from which to assess leader cognitive and affective learning. The Educational Leader Candidate Belief Scale (ELCBS) is used for measuring educational candidates' leadership dispositions. ELCBS was built based on a systematic sampling process of the ISLLC performance expectations dispositions. Initial evidence of validity and reliability are presented, using the Rasch model of item response theory. The interval-level scale being produced, along with additional measures being developed, provides the potential to assess leaders' impact on children's achievement.

NOTE: This manuscript has been peer-reviewed, accepted, and endorsed by the National Council of Professors of Educational Administration (NCPEA) as a significant contribution to the scholarship and practice of education administration. In addition to publication in the Connexions Content Commons, this module is published in the *International Journal of Educational Leadership Preparation*,[1] Volume 6, Number 3 (July - September, 2011), ISSN 2155-9635. Formatted and edited in

*Version 1.3: Sep 12, 2011 10:40 am GMT-5
†http://creativecommons.org/licenses/by/3.0/
[1]http://www.ncpeapublications.org

Connexions by Theodore Creighton and Brad Bizzell, Virginia Tech and Janet Tareilo, Stephen F. Austin State University.

1 About the Authors

Dorothy C. Rea is assistant professor of Educational Leadership at Florida Gulf Coast University. Dr. Rea was previously a teacher and an elementary principal in the Shawnee Mission, Kansas as well as teaching as an adjunct in Educational Leadership at the University of Missouri at Kansas City.

Cecil F. Carter is an associate professor of Educational Leadership at Florida Gulf Coast University where he has served since 1997. Prior to joining the professoriate, he served in three states as a classroom teacher, principals, state education official, district administrator and superintendent.

Judy R. Wilkerson is a Professor of Measurement, Research, and Evaluation at Florida Gulf Coast University where she teaches undergraduate and graduate courses in measurement, assessment, and research. She has co-authored two books on assessing teacher competencies and dispositions and has consulted widely to state departments of education and universities in the areas of accreditation and assessment.

Thomas C. Valesky is currently a professor in the Educational Leadership program at Florida Gull Coast University and the coordinator of post master's degree programs in the College of Education. Dr. Valesky was previously chair of Educational Administration and Supervision at the University of Memphis, member of the Executive Board of NCPEA from 1995-98 and President of the Southern Regional Council of Educational Administration (SRCEA), 2001-2002.

William Steve Lang is a Professor of Educational Measurement and Research at the University of South Florida St Petersburg where he teachers graduate courses in measurement, statistics and research. He has taught in the K-12 system as well as the universities in South Carolina and Georgia and co-authored two books on dispositions– Assessing Teacher Dispositions – Five Standards-Based Steps to Valid Measurement Using the DAATS Model and its teacher competency corollary, using the DAATS model.

2 Sumario en espanol

El Concilio de Oficiales Principales de Escuela de Estado (CCSSO), por la Escuela Interestatal Líderes Licensure Consorcio (ISLLC), estándares desarrollados para el conocimiento, para las habilidades, y para las disposiciones necesarias para la práctica efectiva por líderes educativos (CCSSO, 1996).) Estos estándares proporcionan un dominio contento viable de que valorar a líder aprender cognoscitivo y afectiva. El Líder Educativo Candidato Creencia Escala (ELCBS) es utilizado para la medición disposiciones del liderazgo de candidatos educativas. ELCBS fue construido basado en un proceso sistemático de muestreo de las disposiciones de esperanzas de desempeño de ISLLC. La evidencia inicial de la validez y la certeza es presentada, utilizando el modelo de Rasch de teoría de respuesta de artículo. La escala del intervalo-nivel para ser producida, junto con medidas adicionales ser desarrollada, proporciona el potencial para valorar el impacto de líderes en el logro de niños.

NOTE: Esta es una traducción por computadora de la página web original. Se suministra como información general y no debe considerarse completa ni exacta.

3 Purpose and Justification

As colleges of education are faced with NCATE requirements to assess dispositions in addition to knowledge and skills, preparation programs across the country are looking for ways to assess dispositions through valid and reliable measures. In this article we describe the development of a survey instrument to assess the dispositions of master's degree Educational Leadership candidates. We began by using the dispositions enumerated in the document developed as a companion piece to the 2008 national educational leadership policy standards (Council of Chief State School Officers, 2008a) titled *Performance Expectations and Indicators for Education Leaders* (Council of Chief State School Officers, 2008b).

4 Literature Review

4.1 Dispositions Definition

To assess dispositions effectively, one needs to define the construct. Katz (1993) defined dispositions as patterns of behavior, exhibited frequently and intentionally in the absence of coercion, representing a habit of mind. In 2001, Ritchhart viewed dispositions as a collection of cognitive tendencies that capture one's patterns of thinking, addressing the gap between abilities and actions. Perkins (1995) defined dispositions as the proclivities that lead us in one direction rather than another within the range of freedom possessed. Wilkerson and Lang (2007) defined dispositions as teachers affect—attitudes, values, and beliefs that influence the use of knowledge and skills.

Multiple operational definitions also exist. Wasonga and Murphy (2007) defined eight dispositions for co-creating leadership. Co-creating leadership refers to the process in which the leaders and the led collaborate to maximize human capacity to realize the vision of the organization. The dispositions related to co-creating leadership include collaborating, active listening, cultural anthropology, egalitarianism, patience, humbleness, trust and trustworthiness, and, resilience. Theoharis and Causton-Theoharis (2008) identified three educational leader dispositions–global theoretical perspective, imaginative vision, and sense of agency. Richardson and Onwuegbuzie (2003) measured 11 dispositions, including collaboration, knowledge application, critical thinking, reflective practice, individualized instruction, professionalism, reliability, enthusiasm, high expectations, communication proficiency, and technological proficiency. Brown, King, & Herron (2008) examined the belief that all children can learn, content currency, commitment to research use, and sensitivity to others' views.

The National Council for Accreditation of Teacher Education (2008) defines dispositions as the professional attitudes, values, and beliefs that are demonstrated through verbal and non-verbal behaviors and support student learning and development. NCATE expects institutions to assess fairness and the belief that all students can learn and also suggests use of the Interstate New Teacher and Assessment and Support Consortium (INTASC) Principles as the professional standards for teacher candidates (National Council for Accreditation of Teacher Education, 2008). The INTASC Principles (2011) include critical disposition for each of the 10 standards. Freeman (in Diez & Raths, 2007) referred to the INTASC standards as having "enshrined dispositions in teacher education apparently with considerable permanence" (p. 8).

Dottin (2009) concluded that educators are just beginning to grapple with the definition. He further stated, "Dispositions, therefore, concern not only what professional educators can do (ability), but also what they are actually likely to do (actions)" (p.85). Damon (2007) warned that for certification-related assessment, dispositions "must be based on clearly defined principles rather than the fuzzy intuitions of whoever happens to be in charge of the process at any one time" (p. 368). The plethora of definitions, then, is of concern.

4.2 Disposition Assessments

In an exploratory, qualitative study (Lindahl, 2009) examined if and how dispositions were taught and assessed in principal preparation programs. All respondents who were interviewed considered that dispositions were a key element of principal preparation. In almost all cases the dispositions identified in the ISLLC standards were used. He concluded that if dispositions were to be addressed in educational leadership programs, a valid and reliable instrument should be developed. However, he qualified this conclusion with cautionary questions about the reliability of assessment practices:

1. Is it possible to develop an effective process for assessing dispositions, or are there some idiosyncratic elements that might not conform well to even a well thought-out process?
2. What levels of expectations ("dispositional tolerance") should be set and what levels define a passing score? Who determines this, and how?
3. How can evaluators prevent their personal biases in favor or against specific dispositions from entering into their subjective judgment of candidates?

4. Are dispositions synergistic in nature, where the whole is greater than a sum of the parts?

At present, the assessment of dispositions is largely dependent on the use of Likert scales of self-reported beliefs that are less closely linked to the standards than are their cognitive counterparts. Examples are reported by Richardson and Onwuegbuzie (2003); Brown, King, and Herron (2008); and Schulte, Edwards, and Edick (2008). Scale development is typically based on locally developed construct definitions such as those identified above, rather than the ISLLC standards directly. These studies also rely on classical statistical procedures, including descriptive statistics, factor analysis, and chi square tests.

In the related area of teacher assessment, the Wilkerson and Lang DAATS battery (2008) uses Rasch modeling, a form of item response theory, to scale teachers' degree of commitment with the INTASC Principles. The more quantitative approach reported by Wilkerson and Lang (2011) also responds to the three popular concerns about the measurement of teacher dispositions: (a) disagreement over definitions, (b) measurement difficulties, and (c) insufficient data relating dispositions to K-12 student learning.

5 Research Question

The gap in the literature of leader dispositions assessment research is twofold. First, there is limited attention to building a scale that systematically samples from the content domain needed for accountability and accreditation (i.e., the ISLLC standards). Second, the measurement process is largely reliant on statistics that fail to address the assumptions for their use and/or do not lead to research designs that take advantage of pairing dispositions results with interval level achievement scores. The question explored here is whether leader dispositions can be scientifically measured using a standards-based instrument and modern measurement/statistical techniques, as is being reported for teacher dispositions assessment (Wilkerson & Lang, 2011).

6 Assessing Educational Leader Dispositions: The Foundation

The Educational Leader Candidate Belief Scale (ELCBS) instrument discussed herein made use of the ISLLC standards, affective measurement literature, psychometric standards, and Rasch measurement. Each is presented briefly below.

6.1 The ISLLC Standards

The Interstate School Leaders Licensure Consortium (Council of Chief State School Officers, 2008b) effectively links knowledge, skills, and dispositions, asserting that:

6.1.1

The performance expectations and indicators exemplify fundamental assumptions, values and beliefs about what is expected of current education leaders... In order to maintain this emphasis in the performance expectations, underlying dispositions are listed as a reminder of importance when interpreting and operationalizing indicators. (p.6)

The standards are organized into six Performance Expectations (PEs), each of which contains a list of dispositions, followed by several elements that include a number of indicators.

6.2 Affective Measurement Literature

Wilkerson and Lang (2007, 2011) provide a comprehensive treatment of the affective assessment literature, including discussion of, and references for, all major assessment methods. They recommend the use of multiple measures, including a combination of self-reports (belief scales and constructed response questionnaires), observations, focus groups, and interviews with stakeholders. Thurstone (1928) agreement scales are recommended for belief scales. The Wilkerson and Lang model, Dispositions Assessments Aligned with Teacher Standards (DAATS), too, is useful in framing the assessment process. The steps are:

1. Define purpose, use, propositions, content, and other contextual factors.
2. Develop a valid sampling plan.
3. Create instruments aligned with standards and consistent with the sampling plan.
4. Design and implement data aggregation, tracking, and management systems.
5. Ensure credibility and utility of data.

6.3 Psychometric Standards

The *Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education,*1999) provide the legal and psychometric standards for all testing and assessment procedures. Extensive guidance is provided for validity, reliability, and fairness. Chapter 14, *Testing in Employment and Credentialing*, is dedicated to testing in certification and licensure contexts, centering on the necessity that such assessments measure job-related functions as a matter of validity. In this case, job-related beliefs are the construct, and the ISLLC Standards provide the content domain from which items were sampled.

6.4 Measurement Method

The Rasch (1960) model is the simplest form of item response theory, calling for careful delineation of the construct during the design stage (Wilson, 2005). Conceptually, the idea behind the Rasch model is simple. The ability (or, in this case, commitment) of individuals and the difficulty of items influence each other conjointly. The Rasch model places ability and difficulty on the same interval scale, so predictions about one from the other can be made. One answers questions like, can a child read a passage because the child is a good reader or because the passage is easy? Lexile scores are used to estimate both the reader's ability and the passage's difficulty. In the physical sciences, we can measure objects for different characteristics, such as weight and hardness. Similarly, an affective application of Rasch would measure the child's level of desire to read as another construct that could explain the cause for reading ability.

With a purposive sample and a skewed distribution, inferential statistics are not appropriate. Rasch modeling is sample independent and requires neither a large sample nor a normal distribution (Bond & Fox, 2007). Rasch allows the user to create an interval level scale that can then be used for associational or intervention research designs in subsequent studies. Validity and reliability statistics are also reported (Linacre, 2003). Rasch is extensively used by most modern test publishers, such as Pearson, in the development of major high-stakes tests.

The analysis used in this study is the dichotomous Rasch model (Smith & Smith, 2004) and *Winsteps* software (Linacre, 2003). The logarithmic formula applied follows, where, P is a probability of answering correctly, and the Rasch parameters are Bn (the ability of person n) and Di (the difficulty of item i):

$$\ln\left(\frac{P_{ni}}{(1-P_{ni})}\right) = B_n - D_i$$ [2]

7 Instrument Development

The DAATS model is being followed in developing the ELCBS. Purpose was defined as both remediation of individual candidates and program improvement efforts; the content domain was defined as the ISLLC standards (Step 1). The sampling plan was based on coverage of most ISLLC indicators (Step 2). The first instrument (ELCBS) was developed using the Thurstone technique (Step 3). Technology (Angel) will be used to manage the data (Step 4). The current study provides beginning evidence of validity and reliability (Step 5.)

[2] http://cnx.org/content/m41072/latest/log_formula.png/image

We created a series of 53 statements, eight to ten per Performance Expectation (PE). Each statement was classified based on our expectation of its difficulty, with the goal of ensuring variability and the expectation that the classifications would change with empirical data. Without variability there is no measurement, only confirmation. Existing measures, such as the one proposed by Brown, Kin, & Herron (2008), showing virtually no variability, are less likely to explain differences in performance. To avoid respondents agreeing without thought to all items, a mix of items with expected "agree" and "disagree" responses was included. Table 1 provides the number of items on the instrument by performance indicator, level of expected difficulty, and agree/disagree distribution.

Allocation of items on ELCBS by difficulty and response mix

Performance Indicator	No. of Items	Expected Difficulty			Expected Response	
		Easy	Medium	Hard	Agree	Disagree
1: Vision, Mission, and Goals	10	3	4	3	3	7
2: Teaching and Learning	10	4	3	3	9	1
3: Managing Organizational Systems and Safety	8	2	4	2	5	3
4: Collaborating with Families and Stakeholders	9	4	3	2	5	4
5: Ethics and Integrity	8	3	1	4	5	3
6: The Education System	8	4	1	3	4	4

continued on next page

Totals	53	20	16	17	31	22

Table 1

Table 2 provides sample items, one for each PE. The scaled scores (Rasch measures) are reported for each of these sample items, showing that in some instances the expected difficulty matched the observed difficulty, whereas, for other items it did not. As calibration continues, these values are likely to shift with more respondents. The point here is that the scaling process is working. We have chosen the most difficult item as the example for PE 6, demonstrating the use of an item that pushes commitments to an extreme level.

Sample Items in the ELDS

Item	Performance Expectation	Expected Response	Expected Difficulty	Rasch Measure	Observed Difficulty
7. All review of progress toward attainment of mission, vision, and goals must be based on systematic evidence.	1: Vision, Mission, and Goals	Agree	Easy	37.36	Easy
16. People can manipulate statistics, so data should be taken with a grain of salt.	2: Teaching and Learning	Disagree	Medium	52.37	Medium
23. Higher performing schools should get additional resources based on merit.	3: Managing Organizational Systems & Safety	Disagree	Medium	23.48	Easy

continued on next page

34. Principals should reach out to business and community members to establish school policy.	4: Collaborating with Families and Stakeholders	Agree	Hard	53.72	Medium
38. If a teacher acts unethically, you need to report it to the authorities.	5: Ethics and Integrity	Agree	Easy	37.27	Easy
52. If the principal believes that a state law is wrong and is morally opposed to implementing it, s/he should resign.	6: The Education System	Agree	Hard	84.69	Hard

Table 2

8 Results of the First Validation Study of the ELCBS

8.1 Procedures

A purposive sample of three types of respondents was identified to participate in the first study, based on their levels of experience and researchers' knowledge of them personally. This personal knowledge was necessary for the judgmental portion of the validation analysis.

Respondents included faculty in the Florida Gulf Coast University (FGCU) College of Education with administrative experience (n=4), practicing principals for two local school districts (n=8), and current students in the Master's in Educational Leadership program at FGCU (n=14). Of the 26 respondents, 16 (52%) were females and 11 were males (38%).

Faculty responded first, as a small pilot test. Minor edits were made based on their suggestions. All data were entered into an Excel file and then converted to a Rasch interval scale, using Winsteps software (Linacre, 2003). The psychometric development and statistical reporting were based on guidelines and recommendations from Bond and Fox (2007), Linacre (2003), Smith & Smith (2004), Smith (2003), and Wilson (2005).

Empirical Results

In Rasch measurement, because both people and items are measured conjointly and ordered on an interval scale, the results are graphed and analyzed on a single vertical "construct map." Typically, one looks at a graphic display of data on a horizontal axis, with the lowest score (a zero) on the left moving toward the highest score on the right. The height of the scores indicates the frequency. In a normal distribution the peak is in the middle. Rasch, however, flips the graph in order to place both people and items on the same vertical scale. Bottom is low; top is high and the peaks are on the left and right, still in the middle in a normal distribution.

The Winsteps output map for the ELCBS is provided in figure 1. Educational leaders are on the left, and items are on the right. At the top are the most committed leaders and the most difficult items. At the

bottom are the least committed leaders and the easiest items. So, for ELCBS, the most committed leaders are F02 and F04 (at the top), and the least committed are S10 and S14 (at the bottom). The most difficult items are 10 and 52 (at the top), and the easiest items are 6, 12, 13, and so forth (at the bottom).

The distance between the items and people is not expected to be equal. That is because the difference in difficulty between the items and commitment among the people is not the same. The goals are to have few gaps in the ruler and to have the items and people at roughly the same places on the ruler. The fewer the large gaps between items, the more confident one can be that the construct is well measured–one indication of construct validity.

Note that the distribution here of people is narrow, representing a homogeneous sample. Statistically, this is not good. In terms of program evaluation, however, it is good. It means that there are no identified low dispositions students in this sample! Note the very high score for one person at the top (F02) and no equivalent scores on the bottom of the scale. In future testing of this instrument, it would be useful to obtain scores from students (or practicing leaders) who are known to have lower levels of commitment to the ISLLC Standards to verify the sensitivity of the scale.

The limited gaps between items provide for confidence that the construct is well measured. There is a relatively normal distribution of people (on the left) but a skewed distribution of scores with a large number of easy items at the bottom right. In future testing of this instrument, it would be useful to obtain scores from students (or practicing leaders) who are known to have lower levels of commitment to the ISLLC Standards to verify the sensitivity of the scale.

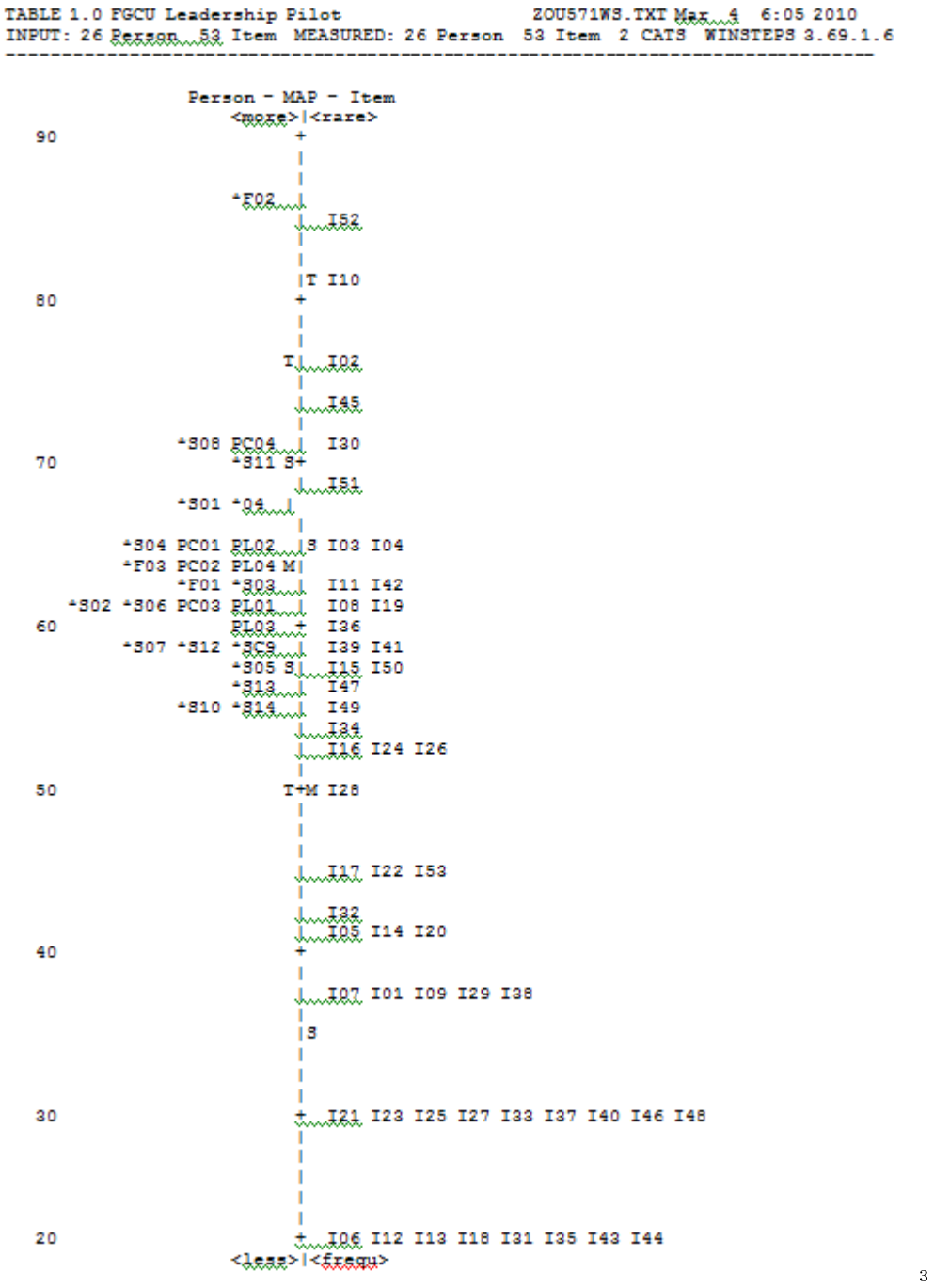

Figure 1. ELCBS Construct Map

Descriptive statistics for the scores, called "measures" in Rasch, are provided in Table 3. Note that the range of scores is greater for items than for leaders, indicating again that there was more homogeneity in persons than in items. All respondents in this sample are at least moderately consistent in their beliefs with the expectations of the ISLLC standards.

[3] http://cnx.org/content/m41072/latest/figure1_valesky.png/image

Descriptive Statistics for ELCBS

Statistic	Leaders	Items
High Measure	54	29.42
Low Measure	86	84.68
Mean Measure	63.31	45.12
Standard Deviation	6.42	18.27

Table 3

The "fit" statistic is critically important in Rasch measurement. While a discussion of the statistic is beyond the scope of this article, we note in passing that this sample of leaders has no extreme misfitting scores; these leaders are well measured, leading to the conclusion that the results are valid. Fit statistics point toward possible refinement of five items, although the fit statistics are not extreme and do not interfere with validity.

The relevant reliability statistic, similar to Cronbach's alpha, is .81, good for this small sample. Because the range of leader scores is limited, the real separation (another Rasch statistic) of 2.04 is low. Fewer easy items, additional difficult items, or more range in the respondents would most likely bring both of these statistics higher. Overall, however, there is nothing in these data to suggest that inferences about leaders and the program would be invalid or unreliable.

8.2 Judgmental Results

The scores of most respondents were logically supportable based on our knowledge of the respondents. At the bottom of the scale are two students about whom we have limited concerns. Several other students expected to be on the lower end are at the lower end, and several expected to be high are at the high end. The rest are appropriately in the middle range. Most of the administrators were located where they were we expected to be, although two principals were lower than anticipated, possibly because of their district and school environments.

9 Conclusions and Recommendations

Responses on this first pilot test of ELCBS meet all of the Rasch statistical requirements for modeling (or measuring) the construct of consistency with the ISLLC dispositions, so there is psychometric reason to have confidence in the measures. The rank order of leaders is as expected, providing judgmental evidence that the scale is working. The construct is well measured; the scores are valid and reliable for these 26 leaders. Reliability is acceptable and should improve with attention to the misfitting items and addition of respondents.

The items on the scale can be related clearly to issues of fairness and the belief that all children can learn, meeting the NCATE requirement for assessing dispositions. For example, leaders who respond appropriately to questions about using data for assessment, equality of resources, diversity, collegiality, and other values tapped provide evidence of their likelihood of making fair decisions. The instrument, then, provides an operational definition of the NCATE requirements as well as the ISLLC standards.

Based on the results of this study, the ELCBS can be implemented in the Educational Leadership program as a valid and reliable measure of candidate dispositions. The items on the scale can be related clearly to issues of fairness and the belief that all children can learn, meeting the NCATE requirement for assessing dispositions. For example, leaders who respond as expected to questions about data use for outcomes assessment, consensus in vision setting, equality of resources, non-exclusive use of test scores, diversity

issues, collegiality, and other values tapped in this instrument have provided evidence that they are likely to make fair decisions benefitting all children in the schools. The instrument, then, provides an operational definition of the NCATE requirements.

Given the psychometric properties of the instrument, combined with raw and scaled scores and distribution, we can conclude that leaders are learning the values associated with the ISLCC standards in this program. The mix of current leaders and future leaders on the scale, with most students slightly lower than practicing leaders confirms that students are making progress toward the acquisition of these learned values.

Because the instrument is technically sound, it also provides useful information to monitor and guide the learning of candidates in the program. Unexpected values surface at an individual item and total score level, providing valuable feedback and opportunities for improving students, who are basically on track.

Similarly, items that were unexpectedly difficult can provide important opportunities influencing the program. For example, one of the most difficult items on the scale was item 2, "Consensus is critical in setting a vision, but sometimes one or two people can get in the way." Lively discussions can result from whether or not those with dissenting opinions are "getting in the way." One of the most difficult items, item 10, asked "With appropriate learning opportunities, most children can learn." This item challenged respondents to choose between most children and all children, again a good topic for discussion. The vast majority of respondents were satisfied with "most" – something we expected but hoped would not happen.

As we progress with the refinement and implementation of the instrument, five items need to be monitored and possibly rewritten to improve the scale. It would also be useful to add some additional difficult items. Leaders with the lowest scores (more than one standard deviation from the mean) should be monitored to determine needs for intervention, especially if their score was different from what was anticipated. Leaders with unexpected incorrect responses on easy items should also be asked about their beliefs on those items.

Most important we also recognize that one instrument is insufficient to measure anything, so we need to continue to explore other instruments to measure dispositions, including the current observation instrument. It may also be useful to review a combination of selected and constructed response items, as recommended in the DAATS model, to provide a more diagnostic picture of future leaders' dispositions toward the nationally accepted standards.

We have addressed the concerns expressed by Lindahl (2009). The care with which we developed and tested the instrument is the result of a well thought-out process. Passing scores are not typically set for dispositional measures; however, using the Rasch model, which produces interval level data it is possible to set cut-scores. We could use traditional standard setting models should there be a programmatic justification to do so. Selected response items avoid the bias of a constructed response or rating form. We are able to produce sub-scale scores for each of the performance indicators as well as a total score from the scale, looking both holistically and individually at the standards.

Most important we also recognize that one instrument is insufficient to measure well, so we will continue developing this and other assessments. The dispositions checklist (an observation instrument) is next. We also hope to develop a scoring system that may be of use to other groups measuring leader dispositions.

9.1 Use of data for program improvement

We will use the results of the data obtained from this instrument to identify candidates who may be having trouble accepting dispositions the profession believes are important for future administrators. A plan to monitor and guide these students in their learning includes a meeting with the advisor to discuss the areas identified as needing improvement. The student will set goals for improvement, and the program faculty meets to ascertain if the candidate shows improvement during classes. It is the advisor's job to be a supportive contact to the student throughout the program.

The strategies will be implemented in various courses in the lesson and activities which focus on the dispositions. The first semester of a required internship will cover ethics and the second semester will focus on social justice. The dispositions will be covered in both these final classes. In addition, the mentor principals will give feedback on each disposition when a candidate has completed the internship experience.

Over time, as the assessment process for leader dispositions is built on this interval scale and the psychometric properties of the instruments have been established, we hope to conduct additional research. This additional research could be focused on the relationships between leaders' dispositions and other variables such as student learning, school climate, teacher satisfaction, and so forth. The potential lessons learned through this type of research appear to be vast.

10 References

American Educational Research Association, American Psychological Association, and National Council of Measurement in Education (1999). *Standards for educational and psychological testing.* Washington, D.C.: American Association of Educational Research.

Bond, T. & Fox, C. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences (2nd).* Mahwah, NJ: Lawrence Erlbaum Associates.

Brown, S., King, G., & Herron, S. (2008). Assessing dispositions of administrative intern candidates. *AASAJournal of Scholarship & Practice,5*(1), 27-34.

Council of Chief State School Officers (1992). *Interstate New Teacher Assessment and Support Consortium standards for teachers.* Washington, D.C.: Author.

Council of Chief State School Officers (1996). *Interstate School Leaders Licensure Consortium standards for school leaders.* Washington, D.C.: Author.

Council of Chief State School Officers (2008a). *Educational leadership policy standards.* Washington, D.C.: Author.

Council of Chief State School Officers (2008b). *Performance expectations and indicators for educational leaders.* Washington D.C.: Author.

Damon, W. (2007). Dispositions and teacher assessment: The need for a more rigorous definition. *Journal of Teacher Education,58*(5), 365-369. doi: 10.1177/0022487107308732

Diez, M. E., & Raths, J. (2007). Dispositions in teacher education. Charlotte, NC: Information Age Publishing.

Dottin, E. (2009). Professional judgment and dispositions in teacher education. *Teaching and Teacher Education: An International Journal of Research and Studies, 25*(1), 83-88.

Katz, L.G. (1993). Dispositions: Definitions and implications for early childhood practice http://ceep.crc.uiuc.edu/eecearch
.

Linacre, J. M. (2003). *A user's guide to WINSTEPS: Rasch-model computer programs.* Chicago: MESA Press.

Lindahl, R. (2009). Teaching and assessing dispositions in principal-preparations programs: A conundrum. In C.M. Achilles, B.J. Irby, B. Alford, & G. Perreault (Eds.), *Remembering our mission: Making education and schools better for students* (pp. 15-38). Lancaster, PA: Pro>activePublications.

National Council for Accreditation of Teacher Education (2008), *Professional standards for the accreditation of teacher preparation institutions.* Washington, D.C.: Author.

Perkins, D. (1995). *Outsmarting IQ: The emerging science of learnable intelligence.* New York: The Free Press.

Richardson, D. & Onwuegbuzie, A. (2003). Attitudes toward dispositions related to the teaching of pre-service teachers, in-service teachers, administrators, and college/university professors. (ERIC Document Reproduction Service ED482689).

Ritchhart,R. (2001).From IQ to IC: A dispositional view of intelligence. Roper Review, *23*(3),143-50.

Schulte, L., Edwards, S., & Edick, N. (2008). The development and validation of the Diversity Dispositions Index. *AASAJournal of Scholarship & Practice,5*(3), 11-19.

Smith, E. & Smith, R., (Eds.) (2004). *Introduction to Rasch measurement.* Maple Grove, MN: Journal of Applied Measurement Press.

Smith, R. M. (2003). *Rasch measurement models: Interpreting WINSTEPS and FACETS output.* Maple Grove, MN: Journal of Applied Measurement Press.

[4] http://ceep.crc.uiuc.edu/eecearchive/books/disposit.html

Theoharis, G. & Causton-Theoharis, J. (2008). Oppressors or emancipators: Critical dispositions for preparing inclusive school leaders. *Equity & Excellence in Education, 41*(2), 230-246.

Thurstone, L.L. (1928). Attitudes can be measured. *American Journal of Sociology,* 33, 529-554.

Wasonga, T. A., & Murphy, J. (2007). Co-creating leadership dispositions. *International Studies in Educational Administration, 35*(2), 20-31.

Wilkerson, J. R. & Lang, W. S. (2011). Standards-based teacher dispostions as a necessary and measurable construct. *The International Journal of Educational and Psychological Assessment 7* (2), 34-55.

Wilkerson, J. R. & Lang, W. S. (2008). Measuring teacher dispositions: An application of the Rasch Model to a complex accreditation requirement. *Resources in Education.* (ERIC Document Reproduction Service No. ED502872).

Wilkerson, J. R. & Lang, W. S. (2007). *Assessing teacher dispositions: Five standards-based steps to valid measurement using the DAATS model.* Thousand Oaks, CA: Corwin Press.

Wilson, M. (2005). *Constructing measures: An item response modeling approach.* Mawah, NJ: Lawrence Ehrlbaum Associates.