Construct Validation of the Louisiana School Analysis Model (SAM) Instructional Staff Questionnaire

Nikki Bray-Clark
Reid Bates
Louisiana State University

The purpose of this study was to validate the Louisiana SAM Instructional Staff Questionnaire, a key component of the Louisiana School Analysis Model. The model was designed as a comprehensive evaluation tool for schools. Principle axis factoring with oblique rotation was used to uncover the underlying structure of the SISQ.

Keywords: Construct Validation, School Accountability, School Improvement

Reform in education has been a major and controversial political topic nationwide for the past two decades, with Louisiana part of the larger movement and cited as having made significant progress in improving student and school performance (Ed Week, 2004). It is accountability that has become the dominant driver in education reform and is one impetus for change.

School Accountability

Accountability is a traditional concept grounded in the belief that when people have a responsibility, there are consequences associated with that responsibility. In the current context, accountability is usually used to describe a system put in place to hold states, local education systems, and students accountable for student achievement. It is primarily a system concept that is ideally designed to provide clear objectives and standards for how states and local education systems will be assessed. There are four main assumptions associated with the current accountability movement:

- Stricter accountability requirements will lead to educational reform and school improvement.
- Meaningful educational improvement is possible through legislated mandates.
- Schools should be the focus of accountability measures.
- School and school district involvement is essential to success (Hansen, 1993).

Accountability and Student Achievement

Routinely, accountability systems provide rewards for improvement and sanctions for failure to improve. Research has shown that schools do respond to incentives in accountability systems, but there is a great deal of variation in the form and strength of school responses based on the type of accountability system examined. Hanushek and Raymond (2002) compared mathematics achievement on the National Assessment of Educational Progress (NAEP) between states with no accountability system and those with differing accountability systems. Their findings indicated a growth of 0.6% percent in states with no formal accountability system in place, a 1.2 percent growth in states with reporting systems only, and a 1.6 percent growth in states with accountability systems that provided for sanctions and rewards as well as public reporting of performance (Hanushek & Raymond, 2002). Additionally, they attribute the highest gains to states with high stakes measures. Typically, accountability systems require state and local education agencies to provide assistance to individual schools that are in need of improvement. The assistance ranges in scope and complexity as school systems move from federal and state mandates to local implementation of school improvement activities.

The investment of state and federal funds in Louisiana’s accountability system has been exceptional, and the stakes for districts, schools, and students are high. Public expectations for improvement in the public schools of Louisiana have supported the accountability movement. Consequently, State and Federal accountability demands are driving financial, administrative, and instructional decisions that have serious repercussions for students.

Copyright © 2005 Nikki Bray-Clark & Reid Bates
School Effectiveness Research

School effectiveness is impacted by several mediating factors including teacher quality, school leadership, and school culture. Central to any serious discussion of school effectiveness is the issue of teacher quality. Although debate may exist on the efficacy of educational innovations, there is a solid research base for the importance and impact of qualified, competent, and caring teachers. Teacher quality refers to several key characteristics of effective teachers including: education, licensure, credentials, commitment, successful experience, retention, and an active interest in self-motivated professional development. There are a number of pivotal issues associated with teacher effectiveness including knowledge, expertise, support, certification/licensure, retention, professional development, and teacher self-efficacy that have been shown to have effects on student achievement. Linda Darling-Hammond, an eminent authority and presently the Charles E. Ducommun Professor of Education at Stanford University and executive director of the National Commission on Teaching and America’s Future, has stated (Darling-Hammond p.7) “The single most important determinant of student achievement is teachers’ qualifications and teachers’ expertise.” She asserts that many if not most teachers currently in the profession are under prepared for the demands being placed upon them. Research has shown that the impact of a high-quality teacher is substantial. Hanushek (1997), found that after controlling for socio-economic status and other confounding variables, a student with a high quality teacher achieves a learning gain of 1.5 grade level equivalents, while a student with a low-quality teacher achieves a gain of only 0.5 grade level equivalents. Defining quality, however, has been problematic for researchers and policy makers because standards of quality differ across states.

Highly talented and motivated people need to be attracted to the teaching profession; however, the induction process those individuals encounter will determine their success and retention. Numerous studies have shown that nearly 50% of teachers leave the profession within the first five years (Colbert & Wolf, 1992; Darling-Hammond, 2000. Research has shown that job satisfaction is a key factor in teacher retention (Weasmer & Woods, 2002). Teacher dissatisfaction comes from multiple sources, but there are commonalities: lack of resources, lack of parental involvement, lack of administrative support, student misbehavior, limited decision making authority, and low pay (Gonzalez, 1995; Jensen, Meyers & Mortorff, 1994; Sewell & Abel, 1999; Shann, 1998). Addressing the issues most often cited for teacher exodus involves leadership and policy issues outside the control of teachers.

School Improvement Measures: The School Analysis Model

The School Analysis Model (SAM) was developed by the Louisiana Department of Education in 1999 as a comprehensive evaluation tool for schools. According to the SAM User’s Guide for District Assistance Teams (LDE, 2000), the model is based on school effectiveness and productivity research. It has its roots in a two-year research study (school years 1996-97 and 1997-98) conducted by the Louisiana Department of Education that provide foundational data needed for the development of a school assistance system that was put into place in 1999. The SAM is based on a conceptual framework that focuses on four primary sources of data: attitudinal, behavioral, cognitive, and contextual. The provision of this data is mandatory for all public schools under the School Accountability System because it provides the front-end analysis and assessment needed for the development of school improvement plans.

The SAM uses a mixed methods approach to collecting, organizing, and reporting this data. The model includes quantitative data collected via questionnaires completed by administrators, teachers, parents, and students, observation checklists, and interview and focus group protocols for the collection of information-rich qualitative data. The complete model is comprised of sixteen data elements and tools including:

- Archival data organizer
- Administrator interview protocol
- Administrator questionnaire
- Classroom observation summary form
- Comprehensive needs assessment-final report
- Contextual observation checklist
- Counselor interview protocol
- Exit summary report
- Faculty needs assessment
- Instructional staff interview protocol
- Instructional staff focus group protocol
• Instructional staff questionnaire
• Parent/community focus group protocol
• Parent questionnaire
• Student focus group protocol
• Student questionnaire

This combination of tools and data collection methodologies was intended to provide the holistic data required for an effective and comprehensive needs assessment capable of identifying gaps, leverage points, catalysts, and barriers to school improvement.

The SAM Teacher Questionnaire

One key element of the SAM model is the SAM Instructional Staff Questionnaire (SISQ). The SISQ is a scannable instrument developed to provide both demographic and school-specific information. Some archival sources (e.g., User’s Guide for District Assistance Teams, 2000) suggest the SISQ is built around a conceptual structure consisting of six “school processes” that are labeled: climate, leadership, culture, curriculum and instruction, parent and school relations, and staff development. However, there is no information providing an underlying theory or rationale for the processes or their inclusion in the model. Nor is there information available describing how these constructs were chosen, the dimensions of the constructs, or how items proposed to measure the constructs were developed or tested. In short, there is no empirical evidence available supporting the validity of the SISQ. Consequently, there were serious questions about exactly what the measures included in the SISQ actually mean. In short, the absence of data supporting the construct validity of the SISQ it is difficult to tell exactly what the instrument is measuring and what value that data should have in terms of providing guidance for the high-stakes process of school improvement. Although the SAM consists of multiple data collection instruments and protocols the focus of this study is on the Sam Instructional Staff Questionnaire. Teacher level data is one of the most essential elements in constructing school improvement plans. If the data used to construct school improvement plans is faulty the plans have no substantive value. It is imperative that the Sam Instructional Staff Questionnaire be psychometrically sound as well as conceptually valid.

Research Question

There is no existing evidence available indicating that the SISQ measures what it purports to measure or that an interpretable factor structure exists. Therefore the goal of this study is to provide evidence relating to the construct validity of the SISQ. Construct validation is an important step in the instrument translation and development process because it establishes the meaning that scores on the instrument have. Construct validation is an ongoing process that is not established through one numerical index or set of data (Benson, 1998). Rather, it is evidence that is accumulated in multiple ways (e.g., substantive, structural, predictive/correlational and convergent/divergent) through the use of different populations, samples and times. In this study, initial evidence of construct validity of the SISQ will be provided through factor analysis. Factor analysis has recently been recognized as a “powerful and indispensable method of construct validation (Kerlinger, 1986, p. 247) that “is at the heart of the measurement of psychological constructs (Nunally & Bernstein, 1994, p. 111). Factor analysis is one method of establishing the structural component of the construct validation process insofar as it provides evidence linking survey items to the structure of the construct. We were guided by the following research question:

Will exploratory factor analysis of the SAM Instructional Staff Questionnaire result in an interpretable factor structure?

Methodology and Research Design

Sample

The respondents in this study were drawn from the population of schools in the State of Louisiana with a middle level grade component (grades 4 through 9) and that submitted SAM Instructional Staff Questionnaire data to the Louisiana Department of Education. School grade configurations vary. There were 297 such schools. A total of 18,585 instructional staff completed the SAM Instructional Staff Questionnaire.

Sample size requirements are a key consideration in factor analysis. Some authors suggest examining item-to-respondent ratios. For example, Hair et al. (1998) suggest a general rule of a minimum of at least 5 times as many
observations as there are variables to be analyzed with a more acceptable size of a ten-to-one ratio. Others offer guidance based on total sample size. Comrey and Lee (1992) indicate the following guidelines: 50 as very poor, 100 as poor, 200 as fair, 300 as very good, and 1000 as excellent. In this study, because the very large sample of SAM responses far exceeded recommended sample size requirements for factor analysis (and to reduce the sheer bulk of data involved in this study) a random sample of 1000 responses was selected from the larger data set. Of these 1,000, about 56.1% had been teaching 10 or more years, 30.3% had been teaching in their present school 10 or more years and 61.3% held at least a Bachelor’s degree.

**Instrumentation**

The SISQ is a survey designed for needs assessment purposes and purports to gather information on six school effectiveness constructs: climate, culture, administrative leadership, curriculum and instruction and staff development. The survey and administration procedures were designed to protect the confidentiality of respondents. The SISQ uses a four-point Likert-type response scale that ranged from 1 = strongly disagree, 2 = disagree, 3 = agree, and 4 = strongly agree. The questionnaire also provided a “don’t know” response option. Demographic questions addressing teaching experience, tenure at present school, education level, absenteeism, and professional development annual leave are included in the instrument.

**Data Analysis**

Exploratory (common) factor analysis (EFA) using principle axis factoring was used to identify the latent construct structure of the SISQ and to provide some evidence of construct validity. The purpose of EFA is to uncover or reveal the underlying conceptual structure of a measurement instrument by examining the correlations between each variable in a data set with every other variable in the set. The choice of EFA over other confirmatory techniques was based on several key factors. First, because there is no existing evidence – either research or conceptual – suggesting an underlying structure to the data it is reasonable to assume that the SAM is in the early stages of instrument development. Exploratory factor analytic techniques are considered more appropriate early in instrument and scale development because they show item cross-loadings on other factors (Kelloway, 1995) that allows for a more complete evaluation of scale content and dimensionality. Second, EFA procedures are seen as more appropriate when the purpose of the research is to identify latent factors rather than for pure prediction (Nunnally & Bernstein, 1994). Third, the absence of a strong theory base in this situation suggests that EFA is more appropriate that other confirmatory techniques. Oblique rotation was employed because of its suitability for latent variable investigation when latent variables may be orthogonal (Hair, Anderson, Tatham, & Black, 1998). Multiple criteria were used to determine the appropriate number of factors to retain including examination of the scree plot, the eigen-value greater than one rule, and evaluation of several trial solutions. The goal was to find the most interpretable latent factor structure that best represented the data.

Effective factor analysis is comprised of several important sequential steps including: data collection and generation of the correlation matrix, factor extraction, decision-making on factor retention and rotating factors to an interpretable, meaningful solution, and construction of scales or factor scores to use in further analyses (Coolidge, 2000; Ferguson & Takane, 1989; Grimm & Yarnold, 1995). For this study, an oblique rotation method was employed, because it is the most appropriate for latent variable investigation when latent variables are expected to have some correlation and it identifies the extent to which each of the factors is correlated and is appropriate for the investigation of latent variables when latent variables may or may not be orthogonal. In addition, if the ultimate goal of a factor analysis is to obtain several theoretically meaningful factors or constructs, an oblique solution is appropriate (Hair, et al, 1998). Several criteria were examined to determine the most appropriate number of latent factors to extract including: a common criterion with latent root (eigenvalue) greater than one (Gravetter & Wallnau, 2000), scree plot, and percent of variance explained. A conservative cutoff of .40 was used to determine the number of items to retain in a specific factor (Nunnally, 1978).

Kaiser’s measure of sampling adequacy (MSA) was also used to determine the factorability of the dataset. MSA is an index with a range of 0 to 1 where 1 indicates perfect prediction of each variable without error by the other variables. For the data set, Hair et al. (1998) suggests an MSA of .80 or above is meritorious. MSA values for individual items were also examined to exclude those falling below .50. None were excluded.

**Results**

The data were screened in several ways to ensure their normality and appropriateness for factor analysis. With respect to normality, visual inspection of the histogram, mean, median, mode and skewness, and kurtosis for each item shows that the data are normally distributed. With regard to the appropriateness of the data for factor analysis,
two statistical tests (overall Measure of Sampling Adequacy (MSA) and the Bartlett Test of Sphericity were conducted. Both of these measures quantify the intercorrelations among variables and provide evidence of the factorability of the data. The results of the MSA (.94) and the Bartlett Test of Sphericity (p < .05) indicated that the data were suitable for factor analysis.

The initial analysis was run without specifying how many factors to retain. This procedure resulted in 10 factors explaining 64.98% of the common variance. However, this factor structure did not appear to be the best representation of the data because of multiple cross-loadings greater than .20 and latent factors with a preponderance of low loadings less than .30 and several latent factors with 2 or less suitable loadings 40 or greater. Additionally, the conceptual fit of the questions loading on several factors was difficult to interpret. In an effort to respond to these issues, the final decision of how many factors to extract was made after comparing solutions extracting 8, 7, 6 and 5 factors respectively.

The 5-factor solution explained 56.63% of the common variance, it represented the most meaningful solution in terms of interpretable latent factors, and was the most consistent with accepted school effectiveness constructs. After using a cut off value of .40, along with deleting cross loadings above .30, 31 of the original 45 items were retained. Results indicated an interpretable five factor structure with average item loading values greater than .50 on major factors. All five of the constructs showed acceptable estimates of internal consistency as measured by Cronbach’s alpha for scales in the early stages of development (.66-.89). The five factors, explaining 53.63% of the common variance, included teacher development and decision-making, parental engagement, quality of instruction, safe learning environment, and expectations for student achievement.

Table 1: SISQ Factor Structure

<table>
<thead>
<tr>
<th>Item #</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td>.80</td>
<td>.00</td>
<td>.11</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>31</td>
<td>.70</td>
<td>.00</td>
<td>.10</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>21</td>
<td>.66</td>
<td>.14</td>
<td>.00</td>
<td>.13</td>
<td>.00</td>
</tr>
<tr>
<td>45</td>
<td>.62</td>
<td>.00</td>
<td>.00</td>
<td>.11</td>
<td>.00</td>
</tr>
<tr>
<td>13</td>
<td>.60</td>
<td>.14</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>37</td>
<td>.58</td>
<td>.00</td>
<td>-.25</td>
<td>-.12</td>
<td>.10</td>
</tr>
<tr>
<td>06</td>
<td>.58</td>
<td>.00</td>
<td>.00</td>
<td>.29</td>
<td>.00</td>
</tr>
<tr>
<td>39</td>
<td>.57</td>
<td>.00</td>
<td>.00</td>
<td>-.11</td>
<td>.10</td>
</tr>
<tr>
<td>28</td>
<td>.51</td>
<td>.00</td>
<td>-.19</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>07</td>
<td>.48</td>
<td>.00</td>
<td>.00</td>
<td>-.11</td>
<td>.00</td>
</tr>
<tr>
<td>14</td>
<td>.00</td>
<td>.85</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>15</td>
<td>.00</td>
<td>.71</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>05</td>
<td>.00</td>
<td>.68</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>30</td>
<td>.00</td>
<td>.65</td>
<td>.00</td>
<td>.00</td>
<td>-.13</td>
</tr>
<tr>
<td>22</td>
<td>.00</td>
<td>.55</td>
<td>.00</td>
<td>.00</td>
<td>-.15</td>
</tr>
<tr>
<td>36</td>
<td>.00</td>
<td>.00</td>
<td>-.74</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>35</td>
<td>.00</td>
<td>.00</td>
<td>-.68</td>
<td>.00</td>
<td>.14</td>
</tr>
<tr>
<td>44</td>
<td>.00</td>
<td>.00</td>
<td>-.65</td>
<td>.00</td>
<td>-.14</td>
</tr>
<tr>
<td>26</td>
<td>.11</td>
<td>.00</td>
<td>-.57</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>20</td>
<td>.00</td>
<td>.00</td>
<td>-.55</td>
<td>.00</td>
<td>.12</td>
</tr>
<tr>
<td>43</td>
<td>.00</td>
<td>.00</td>
<td>-.54</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>27</td>
<td>.00</td>
<td>.00</td>
<td>-.54</td>
<td>.00</td>
<td>.18</td>
</tr>
<tr>
<td>11</td>
<td>.00</td>
<td>.00</td>
<td>-.51</td>
<td>-.19</td>
<td>.00</td>
</tr>
<tr>
<td>12</td>
<td>.00</td>
<td>.13</td>
<td>-.43</td>
<td>-.27</td>
<td>.00</td>
</tr>
<tr>
<td>09</td>
<td>.00</td>
<td>-.11</td>
<td>-.66</td>
<td>-.10</td>
<td></td>
</tr>
<tr>
<td>08</td>
<td>.00</td>
<td>.00</td>
<td>-.13</td>
<td>-.43</td>
<td>.00</td>
</tr>
<tr>
<td>33</td>
<td>.00</td>
<td>.00</td>
<td>.34</td>
<td>.00</td>
<td>-.73</td>
</tr>
<tr>
<td>25</td>
<td>.00</td>
<td>.12</td>
<td>.00</td>
<td>.00</td>
<td>-.71</td>
</tr>
<tr>
<td>42</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>-.70</td>
</tr>
<tr>
<td>18</td>
<td>.00</td>
<td>.00</td>
<td>-.11</td>
<td>.00</td>
<td>-.65</td>
</tr>
<tr>
<td>34</td>
<td>.00</td>
<td>.20</td>
<td>.00</td>
<td>-.24</td>
<td>-.57</td>
</tr>
</tbody>
</table>

| Eigenvalues | 4.8 | 3.4 | 5.2 | 1.9 | 3.36 |
| % Variance Explained | 12 | 3.7 | 1.9 | 1.4 | 1.3 |
Discussion

As noted at the beginning of this paper, it is critical in these days of high stakes accountability that the tools provided to schools for the purposes of improvement are valid, reliable, and grounded in sound research. Unfortunately, this study strongly suggests that the SISQ, an integral part of the Louisiana School Analysis Model measures, does not measure what it purports to measure. The latent factors that emerged in this analysis showed at best only a vague and indefinite resemblance to the original conceptual framework indicated in the SISQ User’s Guide. That framework indicated that the SISQ was built around six “school processes”: climate, leadership, culture, curriculum and instruction, parent and school relations, and staff development. However, the present analysis showed a latent structure with nothing resembling a leadership, climate or culture scale. The remaining scales said to be measured by the SISQ (staff development, curriculum and instruction, and parent and school relations) bore a loose correspondence to three of the latent factors that emerged in this analysis insofar as several items could be interpreted to reflect some dimension of these broad concepts. In short, the results of the study together with the absence of information about how the original scales were developed, data about their relationships with other variables or ability to predict other variables makes it impossible – indeed unwise - to infer that the factors in the original SISQ framework have any appreciable level of validity as a measurement tools.

Although the SISQ may have suffered from oversights in its original development or flaws resulting from its lack of theoretical grounding, it is interesting to note that the scales that emerged in this study do share some overlap with current school effectiveness research. For example, important to school improvement efforts is the link between parental involvement and high student academic achievement (Epstein, 1987; Griffith, 1996; Peterson, 1997). Students whose parents are involved in their schoolwork and who provide adequate structure and support are better prepared to learn. Producing reliable gains in student achievement requires addressing parental and community issues coupled with a movement away from categorical program structures that support a status quo within the school (Goertz, Floden and O’Day, 1995; Hood & Lovette, 2002).

Despite some loose parallels between the research on school effectiveness and the latent constructs that emerged in this study, this does not overcome the fact that some of the items in the SISQ are poorly composed (e.g., double-barreled, ambiguous) and the scales do not appear to address all the critical dimensions of the constructs that should be measured. But it does suggest that there is developmental hope for the instrument. Future research should seek to develop a theory-based model or conceptual framework for school effectiveness, map that against what the SISQ does measure, and identify what constructs the SISQ should measure. Measures for these constructs should be developed using the best scale development practices available. The resulting measures should be subjected to a rigorous program of construct validation to ensure that the critical high stakes school improvement decisions can be made with the aid of valid measurement tools.

Implications for the Field of HRD

This research is important to the field of HRD because it provides evidence critical for the development of one of the cornerstones of our workforce development systems: schools. In Louisiana – a leader in the school accountability movement - some school improvement progress has been made. However, poverty, diversity, wide achievement gaps across ethnic divides, and other factors argue for redoubled efforts to sustain improvement. Indeed, the investment of state and federal funds in Louisiana’s accountability system has been exceptional and public expectations for improvement are high. It is therefore important that valid and useful diagnostic tools are in place and are capable of providing sound information to guide school improvement efforts. Continued research efforts like this one that focus on the development of tools critical to effective school change and improvement are necessary. Research has consistently shown that the reasons for various achievement gaps are complex, and the success of schools in closing those gaps is site specific in nature (e.g., individual identification of deficiencies), and requires a focus on specific areas of weakness (O’Hare & Mather, 2003; Roza & Hill, 2003). It is measurement tools like the SISQ that have become critical in the high-stakes process of school improvement and will be increasingly important to those of us in the field of HRD charged with improving the quality of our schools.

References


