Educational Considerations

Standards-Based Leadership Preparation Program Involvement Through the Use of Portfolio Assessments

Donald G. Hackmann
University of Illinois at Urbana-Champaign

Thomas L. Alsbury
Iowa State University

Recommended Citation

Standards-Based Leadership Preparation Program Improvement Through the Use of Portfolio Assessments

Donald G. Hackmann and Thomas L. Alsbury

The school principal's role has changed dramatically in the past few decades, moving away from management issues and into responsibilities related to leading school reform and facilitating student learning. There is an emerging consensus that successful principals not only must be effective instructional leaders but they also must possess the capacity to transform the school culture to promote improved student achievement (Grogan & Andrews, 2002). Recognizing the administrator's changing role expectations, the Interstate School Leaders Licensure Consortium (ISLLC) crafted six standards for leadership in 1996, which maintain a consistent focus on teaching and learning and assert the leader's responsibility to create "powerful learning environments" (Council of Chief State School Officers (CCSSO), 1996, p. 8). A majority of the 50 states have incorporated the ISLLC standards into their licensure requirements for the principalship. Additionally, in all 50 states, many colleges of education are evaluated and accredited through the National Council for Accreditation of Teacher Education (NCATE) that uses the ISLLC standards in their assessment and processes, requiring seven assessment points and multiple measures including portfolio options.

Through accreditation and state licensure requirements, administrator preparation programs have been called upon to restructure their curricula to more fully address the principalship's shifting role expectations and to better prepare aspiring school leaders. Due to ISLLC mandates, many educational leadership programs are adopting standards-based programs, which are designed to prepare aspiring principals with the competencies necessary to lead school reforms and structure schools that promote improved student learning. This article shares one educational leadership program's experiences with the use of student portfolios to assist in assessment of the program's effectiveness in preparing aspiring school principals. We begin with a discussion of market pressures for program reforms, which include the use of student portfolios for student assessment. After describing various types and purposes of portfolios, we provide a brief review of literature related to evaluation of educational leadership programs and note how portfolio assessments can be used not only for individual assessment but also for program assessment. We then share the results of our analysis of student portfolios and describe programmatic changes our faculty has made to our principal preparation program as a result of this summative evaluation activity.

Market Pressures for Program Modifications

In recent years, preparation programs have been subject to intense scrutiny and criticism because they are perceived as being slow to integrate the principal's changing responsibilities into curriculum content and, consequently, continue to prepare aspiring administrators for outdated roles as top-down managers (Grogan & Andrews, 2002). In addition, market pressures are emerging from alternative leadership preparation programs venues, providing incentives for university-based preparation programs to engage in self-evaluation activities (Glasman, Cibulka & Ashby, 2002).

Continued advancements in distance learning delivery mechanisms may eventually drive programs to more substantive self-evaluation in an effort to determine necessary reforms that may increase appeal to potential clients at the expense of program rigor. This is evidenced by the paradoxical calls from educational administration researchers for the increase in rigor and an emphasis on leadership over management in existing training programs against a growing number of potential leaders who are opting for less rigorous alternative preparation programs that focus on using current practitioners to prepare future leaders with applicable and politically potent management tools that will assure they survive their first year on the job. As a result of these and other forces, many educational leadership programs indeed have restructured, incorporating ISLLC standards into their curriculum content and promoting an enhanced focus on issues related to instructional leadership and school improvement. Some models are being touted as "innovative" (Jackson & Kelley, 2002), experimental (Glasman, 1997), and performance-based (Cox, Biance & Herrington, 1999). Course activities are moving away from traditional forms of assessment—such as research papers and in-class examinations—to more authentic assessment measures to assist the student in skills mastery (Hackmann & Walker, 2001).

Assessment Alternatives in Higher Education Programs

The discussion concerning alternative assessment in education has risen as a natural outcome of a paradigm shift from teacher-centered to learner-centered instruction that started in K-12 settings and has moved into higher education (Huba & Freed, 2000). Cross (1996) noted "it is through a lens that focuses on learning that we must ultimately examine and judge our effectiveness as educators" (p. 9). Although learner-centered instruction within the classroom is not within the scope of this paper, Huba and Freed (2000) and other prominent higher education leaders have stated that the paradigm shift to a learner-centered approach to instruction in graduate programs necessitates a similar shift from assessments used to monitor learning to assessments used to promote and diagnose learning.

Learner-centered assessment is a broad concept that can be defined as a process of gathering and discussing information from multiple and diverse sources in order to develop a deep understanding of what students know, understand, and do with their knowledge as a result of their educational experience. Far from simplistic, there are multiple elements to a learner-centered assessment model, including the formulation of statements of intended learning outcomes, the selection or development of assessment measures, the creation of experiences leading to outcomes, and the discussion and use of assessment results

Donald G. Hackmann is Associate Professor and Program Coordinator of Educational Administration at the University of Illinois at Urbana-Champaign. Thomas L. Alsbury is Assistant Professor of Educational Administration at Iowa State University.

to improve learning (Huba & Freed, 2000). More directly speaking to our study's focus, Huba and Freed (2000) indicated that no definition of learner-centered assessment was complete unless, "the process culminates when assessment results are used to improve subsequent learning"(p. 8). Focusing on the final element of learner-centered assessment, this study focused on using student portfolios for program improvement. Plater (1998) may have stated the need to focus on this element most succinctly when he wrote, "What does the degree or certificate that we award mean and how can we prove it?" (p.12)

Although this study focused on portfolios, assessment measures in higher education programs should include both direct and indirect measures of student learning (Palomba & Bates, 1999). Direct assessments include projects, products, paper, exhibitions, performances, case studies, clinical evaluations, interviews, and oral exams as well as portfolios. Indirect assessments of learning can include surveys of students or past graduates that elicit feedback on what the graduate or student knows or can do with their knowledge. Assessment through objectively scored paper and pencil tests can also be used; however, while easy to use and effective in measuring factual knowledge, they have been criticized for assessing knowledge in discrete bits and lacking references to real-world application (Resnick & Resnick, 1992). Assessments for prospective school administrator are needed that allow the measurement or demonstration of complex abilities such as reasoning, using information to solve complex problems, and the simultaneous use, application, and integration of knowledge in situations where there is often no one correct answer. Huba and Freed (2000) discuss and defend the use of assessment like projects, papers, performances, and exhibitions as well as portfolios in higher education courses. Indeed all of the abovementioned assessment measures are currently used in individual courses within the administrator preparation program in this study and aligned to provide a comprehensive coverage of the ISSLC standards. However, in the administrator preparation program in this study, portfolios were selected as the preferred summative assessment because they allow the inclusion of multiple authentic assessment forms. Black (1993) supported this contention stating, "Perhaps more than any other assessment technique, portfolios provide a detailed mosaic of student learning as it develops over time" (p.146).

Portfolio Use in Administrator Preparation Programs

An increasing number of educational leadership faculties require students to create portfolios during their preparation programs, and the literature base contains an array of diverse programmatic perspectives related to their use. There is general agreement that this compilation permits students to demonstrate theory-to-practice connections (Cornett & Hill, 1992; McCabe, Ricciardi & Jamison, 2000; Wilmore & Erlandson, 1995) or their theories-in-use (Barnett, 1991). In addition, documentation of reflective practice and personal growth is an integral component through the inclusion of reflective writings developed in course activities, daily internship reflection journals, and explanation of portfolio entries (Cornett & Hill, 1992; Edmonson & Fisher 2002; Harris & Arnold, 2001; McCabe et al., 2000; Meadows, Dyal & Wright, 1998; Stader & Neely, 2001).

The support for the use of a portfolio as an appropriate summative alternative assessment is dependent on the format used within the portfolio. Student reflection summaries and self-examination allow for students and instructors to evaluate their work in a systematic way. The inclusion of significant and relevant field experiences in the portfolio along with classroom papers, activities, and presentations place the

emphasis on the demonstration of what students can do rather than simply on whether knowledge has been acquired. However, a portfolio that is a collection of student work is not an assessment tool--it is just a folder. Huba and Freed (2000) noted that in order for a portfolio to be an assessment, "someone must reflect and make judgments about its contents" (p. 234).

Portfolios Defined

An administrative portfolio can be defined as "a collection of thoughtfully selected exhibits or artifacts and reflections indicative of an individual's experiences and ability to lead and of the individual's progress toward and/or the attainment of established goals or criteria" (Brown & Irby, 2001, p. 2). Because it contains the learner's careful and deliberate self-selection of documents that are illustrative of her/his competence and growth, the portfolio—by definition—is unique to the individual.

Two types of evidence are appropriate for inclusion in the portfolio: artifacts and attestations (Barnett, 1995). Artifacts represent tangible products created through the individual's participation in various assignments or work-related responsibilities. For example, an educational leadership student's artifacts may include such course assignments as research papers, an educational philosophy statement, a leadership platform, the student's resume, and a variety of performance-based assessments, such as: student's materials from a clinical supervision activity conducted with a teacher; action research project; case study analysis; data dissagregation and analysis of a school's achievement test scores; creation of a three-year parent involvement plan for a school; or a school cultural analysis. Work-related artifacts may include products developed during the student's clinical or internship placement, such as: a completed school master schedule; school budget; analysis of a school's comprehensive school improvement plan; school crisis management plan; student orientation materials; teacher handbooks; student handbooks; and internship reflective journals. Attestations represent documents created by someone other than the student which verify her/his competencies or accomplishments. Among these artifacts could be college transcripts; letters of recommendation; professional licenses; personal notes from parents or students; and honors and awards.

Types and Purposes of Portfolios

Several portfolio formats are possible, depending on the intended function, which may "vary from enhancing the quality of the learning process to that of standardized reporting by districts or states" (Gredler, 1995, p. 432). An effective portfolio contains three components: biographies of student work; a variety of work; and student reflections (Wolf, 1989). The biography of work illustrates the student's depth of effort within the discipline, noting the development of thought and understanding of content. In contrast, the variety of work documents breadth of effort within the discipline as the learner selects an array of artifacts in various formats across the content area standards. Finally, student reflection is essential for the student to describe each artifact in context; to explain how it documents content knowledge and skills mastery and illustrates personal growth; and to explain what the student learned through the process of creating the artifact (Barnett, 1995; Wolf, 1989).

Portfolio Structure versus Individuality

Portfolios may be accessed to promote self-assessment, program assessment, and external assessment, and different types of evidence will be collected to accomplish each purpose (Barnett, 1995). When used as a self-assessment mechanism, there may be minimal institutional concerns related to standardization of format because the aim is to develop self-directed learners. The student maintains a high degree of control over the contents, selecting artifacts and other entries that demonstrate strengths and weaknesses, while capturing growth over time. Self-reflection is an important element as the student develops the capacity to evaluate her/his academic progress and develop personal goals for continuing learning. A showcase portfolio, in which the learner selects his/her best or favorite works, provides one example of this type of portfolio (Gredler, 1996; Valencia & Calfee, 1991).

When used for program assessment purposes, there likely would be increased institutional requirements for structural consistency, which will restrict the student's freedom in artifact selection. Entries are used as a formative assessment mechanism as the student progresses through the program, with instructors working closely with the student to assess current levels of performance, to note areas in which the student has mastered content standards, and to recommend areas in which additional growth is needed. When the student completes the program, the portfolio becomes a summative assessment tool, with entries scored through the use of predetermined evaluation criteria and rubrics (Gredler, 1996). An evaluation portfolio, containing largely standardized student work collections to report student achievement, provides an example of a portfolio developed for program assessment (Gredler, 1996; Valencia & Calfee, 1991). Portfolios become an external assessment tool when they are shared with others outside the institution to describe the student's skills and abilities (Barnett, 1995). The structure and format of this dossier will vary depending on the intended audience. Aspiring administrators may submit this type of portfolio when interviewing for an administrative position or when applying for their initial administrative licensure.

Portfolios created by practicing administrators are used for three purposes: professional development; performance evaluation; and career advancement (Brown & Irby, 2001). The evaluation portfolio developed while the aspiring principal is enrolled in an educational leadership preparation program could seamlessly evolve into a professional development portfolio once the student has successfully gained an administrative post (Guaglianone & Yerkes, 1998).

Academic Freedom versus Program Continuity

Many leadership programs employ the portfolio as a both a formative and summative assessment tool for the learner, designing it to satisfy the university's comprehensive examination requirements and/or state licensure conditions (Barnett, 1991; Bradshaw, Perreault McDowelle, & Bell, 1997; Edmonson & Fisher, 2002; Harris & Arnold, 2001; Meadows et al., 1998). Because of the relatively high-stakes nature of the summative evaluation component, program faculties tend to standardize the format, defining those categories in which artifacts can be positioned and identifying specific assignments that must be included. Several programs have elected to use leadership standards to frame this portfolio structure, initially using the National Policy Board for Educational Administration performance domains (Wilmore & Erlandson, 1995), state leadership standards (Bradshaw et al., 1997), and more recently the six ISLLC standards (Hackmann & Walker, 2001; Harris & Arnold, 2001; McCabe et al., 2000; Stader

& Neely, 2001). While the use of the ISLLC standards has become the popular measure for school leadership, the standards are being questioned by some researchers for their narrow focus, and some preparation programs are attempting to assess student performance through a broader lens such as social justice issues (Murphy, 2005; Owings, Kaplan & Nunnery, 2005).

The literature base contains few references to concerted faculty efforts to align course content, instruction, and performance assessments in an effort to enrich the quality of authentic assessment activities that could be included in student portfolios. Barnett (1991) noted that assessment measures "must be integrated into the overall curriculum and course delivery" (p. 6), requiring instructors to "infuse new ideas into their teaching" (p. 7). Hackmann and Walker (2001) explained that their program faculty are engaged in identifying authentic class assignments that could be effective portfolio artifacts. Cox et al. (1999) reported that their program's competency-based approach to leadership includes an aligned curriculum, multiple assessments, and a performance portfolio that students develop over the course of their entire program of studies. Although Meadows et al. (1998) noted that "a positive result of the implementation of portfolio assessment has been the resulting improvement of instructional practices and course design throughout the educational leadership preparation program" (p. 97), they acknowledged that this outcome was unanticipated. That many reports concentrate on the creation of the portfolio itself (the product) and do not discuss the interrelationships of curriculum and instruction to the design of performance assessments (the process), however, does not necessarily provide evidence that pedagogical discussions did not occur among the faculty.

Program Evaluation in Administrator Preparation Programs

Educational administration faculty members should engage in continuous self-assessments of the effectiveness of their administrator preparation programs so that they can identify areas in which their students could be more effectively prepared to assume leadership roles. However, preparation programs traditionally have not actively engaged in program evaluation. Glasman, Cibulka, and Ashby (2002) point out that leadership programs actually have had numerous disincentives for program improvement, including a lack of universal agreement on standards for leadership, a lack of pressure from the policy community to reform leadership programs, resistance from within the university community, and market restraints that historically have discouraged academic rigor.

When self-evaluations have been reported by leadership faculty, they typically include the compilation of perceptual data, such as surveys to assess graduates' perceptions of the quality of their preparation (Krueger & Milstein, 1995; Slater, McGhee & Capt, 2001) and feedback from supervisors and hiring officials related to novice administrators' preparation (Krueger & Milstein, 1995). These data are limited in that they relate to only individuals' perceptions, rather than addressing a program's efficacy in ensuring that students have attained program goals and have internalized essential content knowledge and skills.

The literature base related to portfolio analysis for program evaluation purposes is virtually nonexistent (Glasman et al., 2002), and there is a lack of agreement on the appropriate usage of portfolios for evaluation purposes. For example, Gredler (1995) and Lindle (1997) caution against their use as an evaluation tool while Harris and Arnold (2001) actively promote this purpose. Although McCabe et al. (2000) reported that graduates believed their portfolios assisted them

in demonstrating attainment of administrative knowledge and skills, this information, once again, relied on surveys to assess graduates' perceptions. An analysis of authentic artifacts contained in student portfolios could be helpful in evaluating a program's effectiveness in aligning curriculum, instruction, and assessments to the program goals and curriculum standards.

Standards-Based Portfolios: Iowa State University's Experience

At the beginning of the Fall 1999 semester, the Iowa State University educational administration faculty implemented a restructured principal preparation program that was aligned to the ISLLC standards. A new assessment requirement was the inclusion of portfolios to document content mastery upon program completion. Students were to self-select a minimum of two authentic artifacts within each standard that they had developed in their course activities and through their 400-hour internship placements. Reflective writings were included within each standard in which the student explained why each artifact was selected and described how the artifacts *in toto* documented proficiency under the standard. A portfolio defense became the foundation of each student's two-hour oral examination with her/his committee of professors.

The first students to complete the restructured program graduated in Fall 2001, and formative data generated through informal analysis of the portfolios and faculty questioning of students during the oral examinations immediately began to disclose both strengths and limitations of the standards-based curriculum. Faculty observed that quality varied tremendously among the submitted artifacts; yet students generally were able to verbalize sufficient content knowledge and skills during the oral examination. In addition, portfolio entries frequently did not fully demonstrate authentic theory-to-practice connections because students tended to include artifacts that contained few references to the educational administration literature.

The faculty accumulated the portfolios of graduating students over a two-year timeframe, providing sufficient numbers to engage in a summative evaluation of the program as evidenced in the content of these documents. Results of this analysis would enable faculty to draw conclusions related to the effectiveness of the restructured program in adequately preparing aspiring school leaders, illuminating weaknesses in student mastery for individual ISLLC standards and to permitting cogent recommendations for modifications in curriculum content, instruction, assessment, or portfolio design directives for staff and students at Iowa State University. The remainder of this paper explains the methods used to analyze the portfolios, explains the results, and discusses programmatic reforms implemented as a result of this inquiry.

Methods

During the Fall 2003 semester, two faculty members conducted a summative portfolio analysis, closely examining all available portfolios (n = 26) from principal licensure students who had graduated between the Fall 2001 and Summer 2003 semesters. These 26 students represented 9 females and 17 males who were experienced teachers when entering the program. At the time of their oral examinations, nine of these individuals had attained an administrative position, either as principal or assistant principal, and 8 of the 9 were males.

A qualitative research method was used in conducting a content analysis, generally categorized as a deductive qualitative analysis where the data were analyzed according to an existing framework (Patton,

2002). In this study the pre-existing set of typologies or rubrics was the six ISLLC standards and descriptors as well as portfolio quality measures including: organization; critical and reflective thinking; grammar; spelling and mechanics; overall presentation; and use of references. A scoring scale was developed to translate the content analysis into a numerical rating for level of overall demonstration of each of the ISLLC standards as well as each of the quality measures noted above. The following category headings and descriptions were used:

1. Advanced (4 points) - All reflections and artifacts clearly and effectively demonstrate the knowledge, dispositions, and complex performance related to the standards.
2. Basic (3 points) - Most reflections and artifacts clearly and effectively demonstrate the knowledge, dispositions, and complex performance related to the standards.
3. Emerging (2 points) – Some reflections and artifacts clearly and effectively demonstrate the knowledge, dispositions, and complex performance related to the standards.
4. Unacceptable (1 point) – Few reflections or artifacts clearly and effectively demonstrate the knowledge, dispositions, and complex performance related to the standards.

To provide some measurement reliability and validity, several methods were employed including inter-rater reliability and a content analysis protocol. Researchers independently evaluated and scored the portfolios using the same ISLLC-based rubric and scoring scale. The protocol called for the rater to review and use a list of the ISLLC standards delineated into its 44 knowledge, 44 dispositions, and 97 performance descriptors. The raters were instructed to checkmark one or more of the 185 ISLLC descriptors as they reviewed the content of the six portfolio reflections (one for each ISLLC standard), the student's overall reflection of their learning over the entire preparation program, and the 12 artifacts (two for each standard). The rater then scored the portfolio contents on the scoring scale (one to four) described above for each of the ISLLC standards as well as the quality measures noted.

Researchers then compared, discussed, and agreed on the proper valuation for the level of standard attainment demonstrated by the students through their selected artifacts. This technique, called consensual validation (Patton, 2002, p. 467), provides a substantive significance that otherwise is not possible in studies of qualitative data. The method also tends to negate personal bias that might be brought by a single scorer and thus provides a measure of inter-rater reliability (Creswell, 2002). During the course of the analysis, patterns emerged that led to a modification of the original rubric scale, changing the methodology from what appeared would be a straight deductive approach to a combination of inductive and deductive analyses.

Additionally, the portfolio raters noted whether each portfolio artifact was developed within a specific course or created during their clinical activities or other job-embedded activities. Also, in an attempt to determine if artifacts demonstrated theory-to-practice connections, we noted whether artifacts represented authentic activities that would be completed by school leaders or were more theoretical in nature.

Quantifying ISSLC Attainment

Because of the use of rubric rankings, it was possible to procure numerical values as an outcome of the content analysis, moving the analysis methods into a type of quantitative approach. Although this archival content analysis strategy has received mixed support among educational methodologists (Krippendorff, 1980; Weber, 1985), this

mixed methodology was appropriate because it permitted us to view the information from multiple vantage points, leading to a more comprehensive analysis of data for program evaluation purposes. Consequently, correlation data, while providing a measure of inter-rater reliability and instrument validity, is of secondary importance compared to patterns emerging from the descriptive data analyzed through traditional qualitative comparative analyses.

Results and Analysis

Qualitative Content Analysis

Careful analysis of the source of each artifact determined that nearly all items were generated from in-class activities (such as problem-based learning activities and group projects) or course assignments (such as research papers, administrator interviews, and book summations). When assessed through a lens of authenticity, the researchers noted that many artifacts were more theoretical in nature, demonstrating limited connections to administrative practice. This finding primarily was due to the fact that only a few artifacts were presented emanating from students' field-experience placements even though 400 hours of clinical activities were required throughout the program. Artifacts that were closer to the theory side of the theory-practice continuum included such documents as research papers, PowerPoint presentations related to reviews of leadership books, interview summaries, and administrative platforms. In addition, faculty noted that the majority of the submitted artifacts typically did not require students to access the literature base related to educational leadership. When examined by gender, there appeared to be little difference related to artifact origin: both females and males tended to primarily include class-based assignments.

The content analysis disclosed both unnecessary content overlap and the absence of essential curriculum content. Redundancy was noted, in that students had completed essentially similar assignments in multiple classes; for example, students engaged in duplicative group activities dedicated to designing "schools of the future" and conducted numerous interviews of practicing administrators, counselors, and board members. Conspicuously absent were artifacts related to administrative uses of technology, knowledge of effective instructional practices in promoting student learning, effective assessment practices, diversity, transformational leadership, social justice, and school reform.

Some confusion apparently existed related to students' understanding of the type of portfolio that was to be developed. Some presented this document as a learning portfolio that displayed their growth throughout the program; these students tended to include their original class assignments that contained their instructors' grades and corrections. Others chose to include artifacts that were a source of pride even though they had developed other products that could have been more effective in demonstrating mastery of the standards. It was possible that students excluded authentic artifacts generated in the field because they had not previously submitted them to their instructors for review or because they may have found it difficult to fully document and explain their levels of involvement with artifacts jointly developed with their mentor principals. Analysis of the students' reflective writings, however, disclosed that they displayed an understanding of the content knowledge and skills contained within each standard and that they generally were effective in assessing their personal mastery of each standard.

Descriptive Statistics

Group means disclosed that the rubric scores on the 26 student portfolios on average clustered around the basic level on every standard. The numerical ratings followed the values: Advanced = 4, Basic = 3, Emerging = 2, and Unacceptable = 1. As shown on Table 1, students approached the Basic level on Standard 1 (vision of learning) and Standard 5 (integrity, fairness, ethics). They exceeded the Basic level on Standard 2 (school culture and instructional programs); Standard 3 (management of the organization, operations, resources); Standard 4 (collaboration with families and community); and Standard 6 (political, social, economic, legal, and cultural context). Mean ratings were highest overall on Standard 3, which addresses management of the organization. Additionally, proficiency means were achieved under the "quality areas" of organization, critical/reflective writing, writing mechanics, and overall presentation, but the mean was below the Basic level for students' use of references.

Score Variation Based on Gender

Data disclosed a consistent pattern between male and female performance on the portfolio, with females scoring higher on every ISLLC Standard and on the additional quality standards measured in this analysis. The most pronounced difference between male and female scores was observed in Standard 3 (management of the organization, operations, resources), with a difference of 0.49, and Standard 6 (political, social, economic, legal, and cultural context), with a difference of 0.45. Within the criteria for portfolio quality, females showed the highest difference scores in organization and overall presentation, each with a difference of 0.60.

Analysis of Variance (ANOVA) tests, shown in Table 2, disclosed that the score differences between males and females were statistically significant for Standard 3 ($\rho = 0.011$), Standard 6 ($\rho = 0.035$), Total Standards ($\rho = 0.014$), organization ($\rho = 0.10$), and overall presentation ($\rho = 0.019$). The alpha level set for the two-tailed ANOVA test was 0.05. Additional ordinal nonparametric correlation tests included a Mann-Whitney U and Wilcoxon, which yielded confirmation of the results established by the parametric tests. Establishing homogeneity of variances is necessary when conducting analyses of variance, particularly when the population size is small as in the current study. Homogeneity of variance tests indicated that the populations from which the two groups (male and female) were drawn were equally variable. A varimax-rotated principal components factor analysis indicated that scores from Standards 1, 3, and 4 were closely related on one factor while scores from Standard 2, 5, and 6 were closely connected on a second factor. Although this variability in the clustering of the Standards is difficult to explain, it may indicate the need to design portfolios that require a composite and integrative approach rather than our current practice of delineating reflections and artifacts for each independent standard.

This rubric analysis suggest that although Iowa State University's principal preparation program was conceived to focus on leadership principles over management, portfolio artifacts show that student mastery is most highly developed in the area of school management and least developed in demonstrating a vision of learning and engaging in transformational leadership. The lower score on the ethics standard may point to a difficulty in developing high-quality course assignments and field requirements related to students' experiences with professional ethics.

Table I
Mean Scores for ISLLC Standards for the Iowa State University
Principal Leadership Program Culminating Portfolios, 2001 to 2003

Content Standard	Gender	Mean Rating	Standard Deviation	Standard Error
Standard 1	Female	3.22	.441	.147
	Male	2.82	.529	.128
	Total	2.96	.528	.103
Standard 2	Female	3.33	.707	.236
	Male	3.29	.588	.143
	Total	3.31	.618	.121
Standard 3	Female	3.67	.500	.167
	Male	3.18	.393	.095
	Total	3.35	.485	.095
Stabdard 4	Female	3.22	.667	.222
	Male	2.88	.485	.118
	Total	3.00	.566	.111
Standard 5	Female	3.11	.333	.111
	Male	2.82	.529	.128
	Total	2.92	.484	.095
Standard 6	Female	3.33	.500	.167
	Male	2.88	.485	.118
	Total	3.04	.528	.103
Quality Standard	**Gender**	**Mean Rating**	**Standard Deviation**	**Standard Error**
Organization	Female	3.89	.333	.111
	Male	3.29	.588	.143
	Total	3.50	.583	.114
Reflection Quality	Female	3.67	.500	.167
	Male	3.29	.588	.143
	Total	3.50	.578	.113
Writing Mechanics	Female	4.00	.000	.000
	Male	3.76	.437	.106
	Total	3.85	.368	.072
Use of References	Female	2.56	.726	.242
	Male	2.41	.618	.150
	Total	2.46	.647	.127
Overall Presentation	Female	3.78	.441	.147
	Male	3.18	.636	.154
	Total	3.38	.637	.125

n = 26 (Females = 9, Males = 17).

Discussion

This program evaluation activity provided an interesting array of data, which has been helpful in guiding faculty discussions and assisting in the identification of needed improvements to the principal preparation program. This section focuses on the quality of student artifacts, curriculum alignment issues, intended portfolio type, and feedback related to scores on the standards and gender differences.

Quality of Student Artifacts

The artifact analysis disclosed that the quality of portfolios varied greatly, ranging from dossiers that primarily contained theory-based classroom assignments to those consisting mainly of job-embedded products with no theoretical underpinnings. Meadows, Dyal, and Wright (1998) explain that "a major focus of the portfolio should be to address theoretical knowledge gained in courses as well as competencies attained through practical experiences" (p. 96). Certainly, the majority of these students effectively demonstrated the theory-to-practice linkages within their overall portfolio framework, but some students clearly were unsuccessful in establishing this important connection between theoretical knowledge and administrative practice.

A more in-depth analysis of artifacts uncovered the fact that, with appropriate modifications to course assignments, the products could

Table 2
Results of ANOVA Test Comparing Results of Male and Female Students for ISLLC Standards and Portfolio Quality Standards for the Iowa State University Principal Leadership Program Culminating Portfolios, 2001 to 2003

		Sum of Squares	df	Mean Square	F ratio	F probability
Standard 1	Between Groups	.935	1	.935	3.725	.065
	Within Groups	6.026	24	.251		
	Total	6.962	25			
Standard 2	Between Groups	.009	1	.009	.023	.881
	Within Groups	9.529	24	.397		
	Total	9.538	25			
Standard 3	Between Groups	1.414	1	1.414	7.591*	.011
	Within Groups	4.471	24	.186		
	Total	5.885	25			
Standard 4	Between Groups	.680	1	.680	2.229	.149
	Within Groups	7.320	24	.305		
	Total	8.000	25			
Standard 5	Between Groups	.487	1	.487	2.179	.153
	Within Groups	5.359	24	.223		
	Total	5.846	25			
Standard 6	Between Groups	1.197	1	1.197	4.983*	.035
	Within Groups	5.765	24	.240		
	Total	6.962	25			
Total Standards	Between Groups	23.693	1	23.693	7.050	.014
	Within Groups	80.654	24	3.361		
	Total	104.346	25			
Organization	Between Groups	2.082	1	2.082	7.784*	.010
	Within Groups	6.418	24	.267		
	Total	8.500	25			
Reflection Quality	Between Groups	.817	1	.817	2.603	.120
	Within Groups	7.529	24	.314		
	Total	8.346	25			
Writing Mechanics	Between Groups	.326	1	.326	2.556	.123
	Within Groups	3.059	24	.127		
	Total	3.385	25			
Use of References	Between Groups	.122	1	.122	2.82	.600
	Within Groups	10.340	24	.431		
	Total	10.462	25			
Overall Presentation	Between Groups	2.128	1	2.128	6.362*	.019
	Within Groups	8.026	24	.334		
	Total	10.154	25			

*$p < .05$ (two-tailed).
n = 26 (Females = 9, Males = 17).

have been more effective in facilitating theory-practice connections for students. For example, many assignments containing reflective writings or journal entries did not require students to reference the literature in their reflections. Simply incorporating the requirement that students were to cite the literature base within their reflection could be an effective mechanism on promoting these connections to practice. Also, the relative paucity of products from internship experiences may be related to the relative autonomy that our students and mentors have enjoyed during the internship placement. Providing more definition and structure to the clinical experience would enhance the probability of students creating high quality field-based artifacts.

From our knowledge of the types of activities contained in our educational administration course syllabi, we were aware that students frequently chose artifacts that were of lesser quality or were less effective in documenting their content knowledge and skills even though they had completed more authentic activities in their courses. The self-selection feature, while permitting students to embrace their showcase portfolios as personal learning tools, did not provide sufficient structure for the faculty to use the portfolio as evaluation tools for the ISSLC standards.

Curriculum Alignment Issues

The content analysis confirmed our informal formative observations from the students' oral examinations: There was a certain amount of content overlap within the courses, as evidenced by duplicated assignments, and there also were gaps in the curriculum. When developing our restructured principal preparation program in 1999, the faculty had created a curriculum matrix that cross-referenced the ISLLC standards and indicators within the 10-course structure in an effort to ensure curriculum content coverage. However, we had not fully analyzed the three elements of the curriculum alignment triangle—the formal, taught, and assessed curriculum. We also had not taken the subsequent steps of reaching agreement on our instructional methods and assessment practices. Consequently, these concerns were not unexpected, and the students' artifacts (and lack thereof) were very effective in illuminating both areas of content redundancy and potential omission of important content.

Intended Portfolio Type

In reviewing the overall format of most student portfolios, it became apparent that the faculty had not provided clarity that the purpose of the portfolio was for program assessment, as opposed to self-assessment. Consequently, the majority of students were presenting showcase portfolios although the faculty had intended for these dossiers to be evaluation portfolios (Gredler, 1996; Valencia & Calfee, 1991). More structure was needed to the portfolio, which would necessarily limit students' freedom to self-select from their array of work products. Because high quality artifacts were desired, students would need to be informed that they would be required to make necessary revisions to graded assignments to ensure that they were error-free.

Although each of our students received a handbook at the start of their program that explained the portfolio development process, one limitation of our current program was that the faculty did not assist students in continuous self-assessments of their artifacts. Their only opportunity to review and select their artifacts came at the end of the program if they chose to share this information with their faculty advisor a few weeks prior to the oral examination. Barnett (1995) explains that some students can become uncomfortable with a lack of direction regarding types of evidence to include in their portfolios. Clearly, time must be built into the curriculum structure for students to review their portfolio contents as a mechanism to assess their continued growth in the program and as an opportunity to guide students' self-selection of high quality artifacts.

ISSLC Standards and Gender Differences

Group means from the rubric scores related the six ISSLC standards disclosed that the students, as a group, scored below the basic level, the intended proficiency level for our students, on Standard 1 (vision of learning) and Standard 5 (acting with integrity, fairness, and in an ethical manner). An additional and unanticipated finding was that

males' scores averaged below females on every rubric, and the male mean scores were below the basic level on Standards 1, 4, and 6, and for the use of references. Additionally, females averaged above the basic level of proficiency on every measure, with the exception of the "use of references" category. Because of this finding, we also examined the cumulative grade point averages (GPA) of males and females and determined that there was no significant difference in GPAs.

The literature is relatively silent on the issue of gender differences and portfolio quality; however, McCabe et al. (2000) reported that females were more likely to report that the portfolio was useful when applying for administrative positions, and they also viewed their internship experiences more favorably than males. This seems to agree with research that has found an ever-growing majority of women in higher education with higher achievement than men in certain fields, such as the social sciences (Jacob, 2002). Jacob (2002) attributes these findings to poor "non-cognitive" skills among boys, including the inability to pay attention in class, to work with others, to organize and keep track of homework or class materials, and to seek help from others. It is possible that females found more value in both their classroom and internship experiences which may have resulted in the selection of more appropriate portfolio artifacts. Because the preponderance of artifacts were written documents, another possibility may be that our female administrator preparation candidates are more skilled at these written exercises. In addition, females scored higher on the quality domains of organization, reflection quality, writing mechanics, and overall presentation, which may have subtly influenced the researchers' scores of their artifacts within each of the six standards. To the extent that the use of more authentic assessments in coursework and summative evaluations play a factor in the gender gap we discovered is beyond the scope of this study, but warrants further investigation considering the findings on gender gap achievement in higher education (Mortenson, 1999; Sommers, 2001).

In addition to the gender differences, a more significant finding emerged from the analysis of the artifacts but which did not become immediately apparent until we reviewed the rubric scores for each standard. We were attempting to assess students' competence by viewing the ISSLC standards as six separate and distinct entities, but our content analysis and rubrics disclosed the inherent difficulties in determining the most effective positioning of a given artifact within the appropriate standard. Consequently, the student's reflective explanation was critical so that the artifact could be placed in its appropriate context. In developing the ISSLC standards, the task force adopted as one of its principles the belief that "[s]tandards should be integrated and coherent" (CCSSO, 1996, p. 7). Instead of promoting an integrated approach to leadership, our faculty was inadvertently forcing our students to compartmentalize their learning activities into these six distinct areas. Noting the difficulties in developing an effective portfolio assessment process, Milstein (1996) asserts that many programs have struggled with this issue.

Principal Preparation Program Changes

Over the past two semesters, the portfolio review, as well as our informal observations regarding students' oral examination experiences, provided feedback that our graduates, although generally demonstrating content knowledge and skills mastery, could be more effectively prepared. Programmatic changes that we have already or plan to implement as a result of this program evaluation include: (a) grounding our program in a conceptual framework that promotes effective principals

as reflective leaders who support high quality schools that result in high levels of learning for every child; (b) working toward consensus on instructional practices and authentic assessments in each course; (c) standardizing clinical experiences; (d) imposing more structure on the evaluation portfolio; and (e) providing students with both formative and summative feedback on their portfolios through their program.

Conclusion

An important goal of portfolio assessment is to "alter the teaching and learning processes in the classroom" (Gredler, 1995, p. 436). Our faculty has utilized the program self-evaluation process to reach consensus on our curriculum, instructional activities, and assessments. The discussions that have occurred as a result of the portfolio analysis have helped us to more fully understand the interrelationships of our courses and their importance in assisting students' development of content and skills mastery. We are taking significant steps toward the development of a culture of collaboration, which is a departure from "the prevailing culture of individual autonomy of university faculty" (Bradshaw et al., 1997, p. 12). We have become more skilled in achieving curriculum alignment within our courses, and we also have assured that our students' clinical experiences are fully structured to address our curriculum content. Faculty discussions have provided us with an opportunity to share our pedagogical beliefs regarding teaching and learning and to more closely align our beliefs with our classroom practices.

The importance of self-evaluation for continuous improvement cannot be overstated. We are now using student portfolios for the dual purposes of documenting students' competence as individuals and for assessing the effectiveness of our preparation program. In our experience, portfolios have been invaluable tools to assist us improving program quality.

References

Barnett, B. G. (1991, October). *Incorporating alternative assessment measures in an educational leadership preparation program: Portfolios and educational platforms.* Paper presented at the annual convention of the University Council for Educational Administration, Baltimore, MA. (ERIC Document Reproduction Service No. ED342080)

Barnett, B. G. (1995). Portfolio use in educational leadership preparation programs: From theory to practice. *Innovative Higher Education*, 19, 197-206.

Black, L. (1993). Portfolio assessment. In T. W. Banta & Associates (Eds.), *Making a difference: Outcomes of a decade of assessment in higher education* (pp. 139-150). San Francisco: Jossey-Bass.

Bradshaw, L. K., Perreault, G., McDowelle, J. O., & Bell, E. W. (1997, November). *Evaluating the results of innovative practices in educational leadership programs.* Paper presented at the annual meeting of the Southern Regional Council on Educational Administration, Charleston, SC. (ERIC Document Reproduction Service No. ED416200)

Brown, G., & Irby, B. J. (2001). *The principal portfolio* (2nd ed.). Thousand Oaks, CA: Corwin Press.

Cornett, J. W., & Hill, M. (1992). Enhancing administrator theorizing through portfolio construction. *Planning and Changing*, 23, 46-53.

Cox, H. S., Biance, M. C., & Herrington, C. D. (1999, October). *Escaping from the service delivery model in educational leadership preparation: Focusing instruction on performance.* Paper presented at the annual convention of the University Council for Educational Administration. (ERIC Document Reproduction Service No. ED440460)

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

Creswell, J. W. (2002). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research.* Upper Saddle River, NJ: Merrill Prentice-Hall.

Cross, K. P. (1996, March-April). New lenses on learning. *About Campus*, 4-9.

Edmonson, S., & Fisher, A. (2002, March). *Aspiring administrators: Promoting reflective leadership practices.* Paper presented at the annual conference of the Association for Supervision and Curriculum Development, San Antonio, TX. (ERIC Document Reproduction Service No. ED470754)

Glasman, N. S. (1997). An experimental program in leadership preparation. *Peabody Journal of Education*, 72(2), 42-65.

Glasman, N., Cibulka, J., & Ashby, D. (2002). Program self-evaluation for continuous improvement. *Educational Administration Quarterly*, 38, 257-288.

Gredler, M. E. (1995). Implications of portfolio assessment for program evaluation. *Studies in Educational Evaluation*, 21, 431-437.

Grogan, M., & Andrews, R. (2002). Defining preparation and professional development for the future. *Educational Administration Quarterly,* 38, 233-256.

Guaglianone, C. L., & Yerkes, D. M. (1998, April). *Administrative portfolios: Current uses and development by prospective and practicing administrators.* Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA. (ERIC Document Reproduction Service No. ED420079)

Hackmann, D. G., & Walker, J. D. (2001). Standards-based preparation for the principalship: Iowa State University's approach to ISLLC. *Education Leadership Review*, 2(2), 28-34.

Harris, S., & Arnold, M. (2001). Portfolio use in two principal preparation programs. *Rural Educator*, 32(1), 19-23.

Huba, M. E. & Freed, J. E. (2000). *Learner-centered assessment on college campuses: Shifting the focus from teaching to learning.* Needham Heights, MA: Allyn and Bacon.

Jacob, B. A. (2002). Where the boys aren't: Non-cognitive skills, returns to school, and the gender gap in higher education. *Economics of Education Review*, 21(6), 589-598.

Jackson, B. L., & Kelley, C. (2002). Exceptional and innovative programs in educational leadership. *Educational Administration Quarterly*, 38, 192-212.

Krippendorff, K. (1980). *Content analysis: An introduction.* Beverly Hills, CA: Sage.

Krueger, J. A., & Milstein, M. (1995). Promoting excellence in educational leadership: What really matters? *Planning and Changing*, 26, 148-167.

Lindle, J. C. (1997, November). *Whoever heard of standardized art? The complexity of using portfolios for licensing principals*. Paper presented at the annual convention of the University Council for Educational Administration, Orlando, FL. (ERIC Document Reproduction Service No. ED420109)

McCabe, D. H., Ricciardi, D., & Jamison, M. G. (2000). Listening to principals as customers: Administrators evaluate practice-based preparation. *Planning and Changing*, 31, 206-225.

Meadows, R. B., Dyal, A. B., & Wright, J. V. (1998). Preparing educational leaders through the use of portfolio assessment: An alternative comprehensive evaluation. *Journal of Instructional Psychology*, 25(2), 94-99.

Milstein, M. M. (1996, March). *Portfolios: potentials and problems for preparation of educational leaders.* Paper presented at the conference-within-a-conference session on performance assessment of future school leaders at the American Association of School Administrators annual conference, San Diego, CA. (ERIC Document Reproduction Service No. ED409602)

Mortenson, T. G. (1999). Where are the boys? The growing gender gap in higher education. *The College Board Review*, August (188), 8–17.

Murphy, J. (2005). Unpacking the foundations of ISLLC standards and addressing concerns in the academic community. *Educational Administration Quarterly*, 41(1), 154-191.

Owings, W. A., Kaplan, L. S., & Nunnery, J. (2005). Principal quality, ISLLC standards, and student achievement: A Virginia study. *Journal of School Leadership*, 15(1), 101-130.

Palomba, C. A., & Banta, T. W. (1999). *Assessment essentials*. San Francisco: Jossey-Bass.

Patton, M. Q. (2002). *Qualitative research & evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.

Plater, W. M. (1998, November-December). So…Why aren't we taking learning seriously? *About Campus*, 9-14.

Resnick, L., & Resnick, D. (1992). Assessing the thinking curriculum: New tolls for educational reform. In B. R. Gifford & M .C. O'Conner (Eds.), *Changing assessments: Alternative views of aptitude, achievement and instruction* (pp. 37-75). Boston: Kluwer Academic.

Sommers, C. H. (2000). The war against boys. *The Atlantic Monthly*, (May), 59–74.

Stader, D. L., & Neely, R. O. (2001, October). *Portfolios for professional growth and documenting ISSLC standards*. Paper presented at the annual convention of the University Council for Educational Administration, Cincinnati, OH. (ERIC Document Reproduction Service No. ED031702)

Slater, C. L., McGhee, M. W., & Capt, R. L. (2001). Review and renewal of an educational administration program. *Planning and Changing,* 32, 2-23.

Valencia, S., & Calfee, R. (1991). The development and use of literacy portfolios for students. *Applied Measurement in Education*, 4, 333-346.

Wilmore, E. L., & Erlandson, D. A. (1995, January). *Portfolio assessment in the preparation of school administrators*. Paper presented at the annual meeting of the Southwest Educational Research Association, Dallas, TX. (ERIC Document Reproduction Service No. ED394906)

Weber, R. P. (1985). *Basic content analysis.* Beverly Hills, CA: Sage.

Wolf, D. (1989). Portfolio assessment: Sampling student work. *Educational Leadership*, 46(7), 35-39.