A Collaborative Approach to Assessment: The Assessment and Improvement Management System (AIMS)

Maggie Payne & Mimi Miller
California State University, Chico

Over the last 25 years, the approaches to determining quality in teacher education have changed considerably. Historically, there has been a tendency to assume that teacher preparation programs are effective, while actual evaluation of programs has been "spotty, evolutionary and limited in scope" (Hall, Smith & Nowinski, 2005, p. 19). During the teacher education reforms of the 1980s and 1990s, programs were judged primarily by the strength and coherence of their curricula, rather than by evidence of student learning (Cochran-Smith, 2006). While there has been some effort in the past to qualitatively measure the effects of teacher education, the current intense focus on evidence is "a significant departure from far and recent past" (Cochran-Smith, 2006, p. 6). Today, the term *quality* cannot be used without evidence—evidence of learning for teacher candidates, and evidence that these teachers impact their young learners.

In response to this call for evidence, teacher education programs are currently developing assessment systems that emphasize the analysis of data to make educational decisions about students, to give feedback to students, to judge instructional effectiveness and curricular adequacy, and to inform policy (Linn & Miller, 2005). This article presents the assessment system that one school of teacher education developed as

Maggie Payne is the associate dean and Mimi Miller is an associate professor, both with the College of Communication and Education at California State University, Chico, Chico, California.

faculty prepared for state and national accreditation. The discussion provides a window into a culture of evidence by examining how this assessment system was created to collect, analyze, and use evidence for program improvement and how a faculty culture began to develop around the use of this system.

Teacher preparation programs collect considerable amounts of data on student demographics, progress, and performance. A recent study on data use in the California State University (CSU) system revealed, however, that while institutions consistently collect data, much of that data goes unanalyzed, or, if analyzed, the results go unreported (Costa, Bartell, Chin, Jesunathadas, Li, Schlackman, & Wong, 2007). This tendency to amass data that remains both unanalyzed and uninformative is described by DuFour (2004) as "the DRIP syndrome—Data Rich/Information Poor" (p.6).

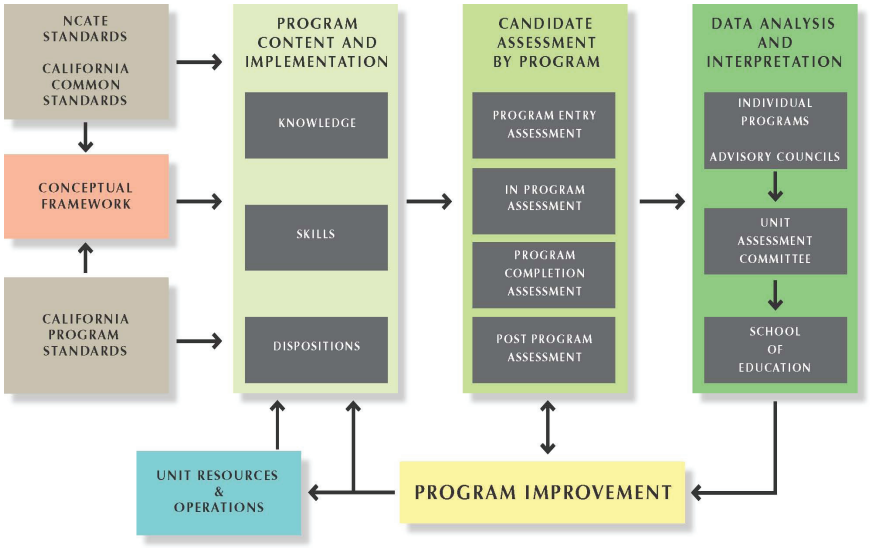
The Carnegie Foundation for the Advancement of Teaching has challenged teacher preparation programs to develop data-driven assessment systems that both inform local program decisions and contribute knowledge to the field (Cochran-Smith, 2003). In recent years, state and national accrediting bodies have increased the pressure on teacher preparation programs to develop a "culture of evidence" in which programs demonstrate effective data collection, management, analysis, and reporting for the purpose of making data-informed program improvements. The National Council for Accreditation of Teacher Education (NCATE) dedicates one of six standards to assessment. In order to earn national accreditation from NCATE, programs must have an "assessment system that collects and analyzes data on applicant qualifications, candidate and graduate performance, and unit operations to evaluate and improve the unit and its programs" (NCATE, 2006, p. 21).

Working with the Evidence

When faced with an impending state and national accreditation visit, the School of Education at a semi-rural, state university looked for a remedy for the DRIP syndrome. The School of Education is comprised of 18 initial and advanced credential and master's programs housed in six departments and three colleges. At the beginning of our process, the School of Education was merely an abstract idea, not fully conceptualized by its own members. Redefining ourselves as a community began with a yearlong, highly iterative process of developing a conceptual framework that would reflect values held by all of our programs.

The Assessment and Improvement Management System (AIMS) (see Figure 1) was developed through a three-year collaborative process

Figure 1
Assessment and Improvement Management System



between School of Education faculty and K-12 partners. This assessment system serves a dual purpose: first, to assess candidates' development of proficiencies based on the unit's conceptual framework and professional and state program standards, and second, to provide data on candidates, resources, and operations to be analyzed for the purpose of program improvement.

As shown in Figure 1, assessment of programs is systematically conducted through the application of a continuous improvement model. The diagram component *Candidate Assessment by Program* shows the data sources utilized to assess program effectiveness in terms of candidate performance. Formative assessment of candidates takes place as they progress through program stages, specifically at program entry and in-program transition points. Summative assessment takes place at program completion. Post-program assessment includes indicators of graduates' professional effectiveness after completing the unit's programs. Each of these stages includes multiple measures to provide an accurate image of candidate learning and performance.

The unit's resources and operations, including faculty and staff, budget, resources, and facilities also have significant impacts on faculty, program delivery, the assessment of candidate performance, and the effective analysis of data. Information on the quality and sufficiency of resources, as well as on the effective use of those resources for unit

operations, is analyzed in the Unit Assessment Report to provide additional feedback for program and unit improvement.

In the AIMS system, coordinators of each program leading to a credential or advanced degree in the School of Education develop and submit a Program Assessment Report to the Unit Assessment Committee (UAC), which is a rotating committee composed of representatives from across the School of Education. Each report includes (1) information about candidates, (2) analysis of key assessments results, and (3) a program improvement plan based on the data analyzed. This improvement plan includes selected area(s) of focus based on the data, an action plan for program improvement in that area of focus, the data to be analyzed to evaluate the action plan after implementation, and the next steps to be taken. The reports are submitted biennially to provide programs with adequate time to implement their action plans and to evaluate the results. The UAC is then responsible for aggregating and analyzing program and unit data and for preparing the Unit Assessment Report. This report combines information from the programs on student learning outcomes and progress on action plans with additional information on demographics, field placements, faculty, equipment, and fiscal resources. The Unit Assessment Report provides a comprehensive picture of the overall vitality of the School of Education in addressing its goals.

When first applying the AIMS system, the UAC quickly realized that the assessment reports, while necessary, were cumbersome; it would be difficult to aggregate and compare data from lengthy program reports. The solution was to create a Program Improvement Summary Chart for each program to complete and submit. The sample Summary Chart shown in Table 1 comes from one of the elementary credential programs in the unit. Column A shows that the program identified "prepare[s] candidates to use technology" as one goal for program improvement. Column B shows data that were used to choose this goal. In this case, the program used data from surveys of program completers at exit, program graduates at the end of their first year of teaching, and the employers/supervisors of those teachers. Data indicated that less than 80% of students were adequately prepared to use technology in their classroom practice. Column C shows this program's actions toward the improvement goal, which included information gathering, curriculum articulation, and the development of technology-based assignments on which candidate performance could be evaluated. Column D explains the results of data after the actions were implemented. In this case, data from several sources indicated that preparation to use technology was still an area of growth. The conclusions, shown in Column E, were to keep technology as a program improvement goal, to implement and

Table 1
Program Improvement for Multiple Subject Program Based on Annual Report, Fall 2006

A. 2005-06 Area(s) of Focus	B. 04–05 Data Sources	C. 2005-2006 Action Plan	D. Data Collected from 05 – 06 Action Plan Implementation	E. Conclusions and Next Steps for 06-07 based on 05-06 Data Analysis
Prepare candidates to use technology	Exit Survey: 73% Adequately to well prepared to use technology for instruction. 78% Adequately to well prepared to use technology for research and record keeping System-wide Evaluation of Teachers 04-05 survey: 70% of first-year teachers and their employers indicate "Adequately Prepared" to "Well Prepared" to use Educational Technology	1. Hold a series of technology meetings to create an action plan and infuse technology-based assignments in coursework. 2. Create a technology matrix that identifies where technology occurs in coursework. 3. Develop technology surveys for candidates' cooperating teachers to find out how frequently they used certain technologies themselves and with their students. 4. Schedule candidate focus groups for additional data.	1. Meetings began 05-06 and continued 06-07. 2. Matrix has been created. 3. Technology Survey: Candidates (n=16) Student teachers most frequently used technology for technical reinforcement of students' learning (44%) and projects/ reports (25%) CTs (n=77) most often use technology for reinforcement (38%) and adapting instruction for special needs students (21%). 4. Data from focus group. *Additional findings*: CSU Exit Survey 2006 (n=78): * How valuable was instruction in "using computer technology for classroom instruction?" 25% answered "Very." 15th of 15 areas of instruction. * How prepared are you "to use computer-based technology for instruction, research, and record keeping?" 21% throught they were well prepared. Lowest rating of 23 areas. * Qualitative data suggests that EDTE 526 could be improved. One student wrote the course "was insufficient in providing concrete and practical applications in a school setting."	Based on our findings, technology should remain a program improvement goal. 1. Analyze technology matrix to identify where technology goals are currently met in coursework. 2. Articulate where technology goals can be met throughout program coursework. Articulation should include technology applications that candidates will most likely use in the classroom (technical reinforcement and supporting special needs students). 3. Change course assignments as needed to reflect this articulation. 4. Continue to collect data. Focus group should specifically consider technology.

evaluate the planned assignments, and to continue the process of articulation across program courses. As next steps, the program planned to continue to collect data and to hold focus groups with students, faculty and community partners.

The decision to use a summary chart was welcomed by all involved, specifically faculty, administrators, and accrediting agencies. Each summary chart gave a snapshot of program progress and, when taken together, gave an overall indicator of the health of the unit. The charts provided program faculty with a simple means of monitoring their program's assessment activity and progress over time, and they allowed the UAC to see patterns in goal choice and to plan how best to support program improvement and continued assessment.

Developing the Culture

Through the processes described above, AIMS has provided a system for dealing with data, the *evidence* in the "culture of evidence." It is equally important, however, to consider the word *culture*. To determine what this culture should look like, we can turn to the work already being done in the K-12 community on the development of Professional Learning Communities (PLCs). Based on this model, the culture of evidence would be characterized by (1) collective responsibility for student learning, (2) systemic thinking, (3) collaborative work with data, and (4) increased reflective practice (DuFour, 2004; Morrissey, 2000; Hord, 1997). Creating such a faculty culture, however, requires overcoming resistance resulting from concerns about the following:

- Lack of time due to intensive workloads and reluctance to be pulled away from teaching and scholarship. (Hall et al., 2005)
- Lack of expertise in conducting data analysis. (Costa et al., 2007)
- Lack of technology resources to support data management and analysis. (Costa et al., 2007)
- Lack of trust in the assessments and the data. (Banta, 2008; Cummings, Maddux, & Richmond, 2008)

Because of the resistance that results from these concerns, it is tempting to assign the assessment work to a single person or to a small committee, particularly when pressure to meet accreditation requirements is high. A key aspect of the culture of evidence, however, is that it can only be fully developed in the context of an engaged faculty community (Dwyer, Millett, & Payne, 2006). Finding ways to work through the resistance, rather than merely avoiding it is, therefore, essential.

Lack of Time

The issue of lack of time is often a symptom of a more deep-seated concern, which is that faculty often fail to see student attainment of program goals and outcomes as a shared responsibility and an important focus for their work. Bringing faculty to consensus regarding a conceptual framework and a set of shared goals was a critical first step in helping them to see assessment as central to their work rather than a distraction from it (Dwyer et al., 2006). By having a clear understanding of program standards and desired learning outcomes, the role of their courses in addressing those standards and outcomes, and the assessments used to measure program and student success, faculty were able to think more systemically and to understand their role in the larger enterprise. The discussion of data and program improvement became a frequent focus of department and program meetings, and the integration of these discussions into regularly scheduled meetings continued the process of engagement without adding to the demands on faculty time. The Program Improvement Summary Charts were particularly helpful in framing conversations around assessment, as they provided a simple, clear picture of the data and the action plan being prompted by or evaluated by that data. In addition, faculty spoke of using data for their scholarly work. Research born of these discussions assisted faculty in meeting professional goals while addressing Cochran-Smith's (2003) call for programs to use evidence to contribute to the field.

The effectiveness of this combination of strategies in overcoming faculty resistance was investigated in the fall of 2007, when a survey was conducted of SOE faculty to determine the extent to which they were familiar with and participated in the AIMS system. Of the 33 full-time faculty, 26 or 79% responded to the survey. Survey items 1-9 addressed responsibility for student learning and systemic thinking, two critical indicators of an emerging culture of evidence (DuFour, 2004; Morrissey, 2000; Hord, 1997). The faculty's responses, shown in Table 2, indicate that most faculty, while less confident of their familiarity with the conceptual framework, were aware of the assessment system, the kinds of assessment data being collected, and the ways in which that data were informing program improvement. In addition, most faculty indicated that they were active participants in the AIMS system process.

Lack of Expertise with Data

While a certain amount of preparation in statistics and data analysis is common across doctoral programs, and while many faculty may engage in scholarship that regularly involves the analysis of quantitative and qualitative data, others may feel less qualified or less inclined

Table 2
Results of Faculty Survey on the Assessment and Improvement Management System (N=26)

Item	Yes	No
Responsibility for Student Learning		
1. I know the student outcomes/expectations for the School of Education as delineated in the unit's conceptual framework.	19 73%	7 27%
2. I know the student outcomes/expectations specific to my program.	25 96%	1 4%
3. I am responsible for contributing grades/scores for one or more of my program's key assessments.	22 85%	4 15%
Systemic Thinking		
4. I know what my program's key assessments are.	24 92%	2 8%
5. I participated in the selection of my program's key assessments.	24 92%	1 4%
6. I know what my program's focus improvement areas are.	25 96%	0 0%
7. I participated in the selection of my program's focus improvement areas	24 92%	2 8%
8. I know what my program's action plan for improvement is.	24 92%	1 4%
9. I have a specific role to play in the action plan for improvement (instructor, assessment scorer, other).	23 88%	3 12%
Collaborative Work with Data and Increased Reflective Practice		
10. I know what data were used to inform the selection of the focus area(s).	24 92%	2 8%
11. I know what data will be used to evaluate the program improvement plan.	18 69%	6 23%
12. In the past year, my program has had one or more discussions about the implications of assessment data.	26 100%	0 0%
13. In the past year, my program has made curricular, policy or other changes based on data.	26 100%	0 0%

to conduct that analysis, particularly when it involves the aggregation of data from multiple sources. What was needed was a way to present analyzed assessment results in a completed report format to allow faculty to use their time and expertise to focus instead on the implications of the results for program evaluation and improvement. The need to manage extensive amounts of in-program assessment data for 18 programs and to provide statistical analysis of those data led to the creation of a staff position of assessment coordinator. Other models being used elsewhere

to address these concerns include providing assigned time to a faculty member or a stipend to a graduate student with expertise in working with data to serve as an assessment coordinator, or identifying a few faculty to serve on an assessment committee.

Lack of Technology

To be accurate, the issue is not the lack of technology but the array of technologies and the variety of locations used by the institution and by faculty to store assessment data. Such storage locations include state and national reports on websites, university information management systems, grade books in online course shells, spreadsheets in individual computer files, and hard copy documents in file drawers. What is often lacking is a coherent, centralized data management system to interface with these varied sources to support the effective use of data. Dealing with this issue on our campus has involved ongoing work with Institutional Technology staff and a data management system developed and managed by students in the College of Business. Many other institutions are making use of the variety of third-party data management systems now available.

Lack of Trust

The lack of trust among faculty for the use of assessment data springs from a couple of concerns. First, faculty may worry that an evidence-based approach to program improvement will lead to a narrowing of the curriculum and a tendency to value only what is measured. This concern underscores the need for broad participation of faculty in early discussions of shared mission, values, and goals, and in decisions regarding key assessments to be used (DuFour, 2006; Dwyer et al., 2006). The more faculty have control over what is assessed and how, the more they "own" the data and the less resistance they will likely feel. The use of multiple measures of outcomes is also critical, as data that are validated through triangulation are far more difficult to dismiss or repudiate.

Faculty may also fear that a focus on assessment data will lead to negative evaluations of their professional practice or a loss of autonomy. It can be helpful to start discussions of evidence with a focus on positive results and on what is working well. Such discussions might be used to identify best practices and to consider ways to replicate them throughout programs. Consideration of the implications of negative data should then focus on systemic concerns and improvements, rather than on individual courses or instructors. Institutions may also want to acculturate new faculty into the culture of evidence by requiring them to describe how

they use assessment data to improve their courses and their teaching as part of the tenure and promotion review process.

Through these strategies, faculty defensiveness and dismissiveness in response to data are being gradually replaced with an eagerness to see results and to translate them into action. For example, when a new state-mandated Teaching Performance Assessment was piloted, a faculty member took on the task of analyzing the data and presenting them at a department meeting within a week of the completion of scoring. While the results were generally encouraging, faculty were quick to point out two areas for improvement and begin brainstorming ideas for addressing the gaps. Clearly, working collaboratively with data is becoming part of our faculty culture, and faculty are becoming increasingly eager to engage in reflection on the implications of those data, to see not just numbers but meaning and possibility. This change is validated by items 10-13 in Table 2 which address the last two characteristics of an emerging culture of evidence: collaborative work with data and increased reflective practice (DuFour, 2004; Morrissey, 2000; Hord, 1997). Survey responses indicated that a considerable majority of faculty knew how focus areas were chosen, although they were less aware of the data that would be used to evaluate the effectiveness of the program improvement plan. All faculty responded that they had been involved in discussions about data during the year and in making decisions based upon this evidence.

The Impact of AIMS

The initial impetus for the development of AIMS was to meet the requirements for national accreditation, which included both the creation of a comprehensive assessment system and the development of a culture of evidence. We can measure our success in three ways. The first measure, and the one we believed would be most important, was the success of AIMS in meeting NCATE's assessment standard. The second measure, and the one that was most unexpected, was the validation that came from the redesign of California's accreditation system to include a biennial program report process that closely mirrors the one developed on our campus. The third measure, and the one that is truly the most important, is the change that has occurred as a result of AIMS in the way faculty are becoming actively engaged in using evidence for program evaluation and improvement. In the fall 2007 survey, the faculty were asked to provide a response to the following question: *Since the implementation of AIMS, what changes have you seen in the way that your program evaluates its effectiveness and engages in program*

development? The following are representative of the responses received from the 26 faculty who participated:

- We are collecting data on student learning.
- Reflective teaching promotes positive learning outcomes.
- Discussions are taking place.
- We have more concrete data.
- More data driven.
- A system is in place to gather data. Regular meetings to monitor program effectiveness.
- *We use data for decision-making.*

This change in faculty culture has resulted in curricular changes to programs designed to more effectively prepare educators to succeed in the K-12 school setting. The gaps revealed by the focus on evidence have led to increased faculty collaboration *within* departments, such as the complete redesign of the approach to teaching technology applications in education. Additionally, there has been increased collaboration *across* departments, such as the co-sponsored daylong Response to Intervention workshop in which SOE faculty learned together about this multi-tiered approach to help struggling learners, including those with disabilities, in both general and special education (Wedl, 2005). In addition, faculty have become more open to inviting K-12 partners into the collaborative process, as when cooperating teachers were asked to help design a rubric, now used by all 18 School of Education programs, for assessing candidates' professional dispositions in the field. Recently, too, several faculty members participated in a daylong workshop on developing Professional Learning Communities (PLCs) based on the work of DuFour, DuFour, Eaker, and Many (2006) and returned inspired to work with K-12 partners to integrate what they learned into the teacher and administrator preparation programs. That they immediately saw the value in the PLC approach is further testament to their newfound appreciation of the power of a culture of evidence. Increasingly, too, faculty have brought candidates into that culture to provide feedback on program effectiveness and to discuss implications of data, giving them a glimpse of the kind of collaboration and analysis that increasingly permeates the field of education.

The development and implementation of AIMS has been a critical factor in the evolution of the School of Education. From a loose collection of programs working in isolation from each other and, in a few cases, relying heavily on faculty perceptions and anecdotal evidence to determine

program quality, the School of Education has become an organization whose members are guided by clearly articulated goals and a system for measuring and improving their effectiveness in meeting those goals.

References

Banta, T. (April 9, 2008) Assessment for improvement vs. assessment for accountability. *Magna Online Seminars.* Retrieved May 9, 2008 from http://www.magnapubs.com/catalog/cds/601221-1.html

Cochran-Smith, M. (2003). Assessing assessment in teacher education. *Journal of Teacher Education, 54*(3), 187-191.

Cochran-Smith, M. (2006). Taking stock in 2006: Evidence, evidence everywhere. *Journal of Teacher Education*, *57*(1).

Costa, V., Bartell, C., Chin, E., Jesunathadas, J., Li, J., Schlackman, J., and Wong, P. (2007). Developing a culture of evidence in teacher education for the California State University. Unpublished report to the Deans of Education, California State University.

Cummings, R., Maddux, C., & Richmond, A. (2008). Curriculum-embedded performance assessment in higher education: maximum efficiency and minimum disruption. *Assessment & Evaluation in Higher Education*, *99999*(1), 1-7. Retrieved September 12, 2008, from http://www.informaworld.com/10.1080/02602930701773067.

DuFour, R. (May 2004). What is a professional learning community? *Educational Leadership, 61*(8), 6-11.

DuFour, R., DuFour, R., Eaker, R., & Many, T. (2006). *Learning by doing: A handbook for professional learning communities at work.* Bloomington, IL: Solution Tree.

Dwyer, C., Millett, C., & Payne, D. (2006). *A culture of evidence: Postsecondary assessment and learning outcomes.* Princeton, NJ: Educational Testing Service.

Hall, G., Smith, C., & Nowinski, M. (Summer 2005). An organizing framework for using evidenced-based assessments to improve teaching and learning in teacher preparation programs. *Teacher Education Quarterly*, *32*(3), 19-33.

Hord, S. (1997). *Professional learning communities: Communities of continuous inquiry and improvement.* Austin, TX: Southwest Educational Development Laboratory.

Linn, R., & Miller, M. (2005). *Measurement and assessment in teaching* (9th Ed.). Upper Saddle River, NJ: Merrill-Prentice Hall.

Morrissey, M. (2000). *Professional learning communities: An ongoing exploration.* Austin, TX: Southwest Educational Development Laboratory.

National Council for the Accreditation of Teacher Education (2006). *Professional standards for the accreditation of schools, colleges and departments of Education*. Washington, DC: NCATE.

Wedl, R. (July 2005) Response to intervention: An alternative to traditional eligibility criteria for students with disabilities. *Education Evolving* (retrieved December 5, 2008 from www.educationevolving.org/pdf/Response_to_Intervention.pdf -).