DOCUMENT RESUME

ED 363 227 HE 026 837

TITLE	Student Outcomes Assessment: What Makes it Work? Assessment Practices & Experiences in the California State University.
INSTITUTION	California State Univ. and Colleges, Long Beach. Inst. for Teaching and Learning.
SPONS AGENCY	Fund for the Improvement of Postsecondary Education (ED), Washington, DC.
PUB DATE	92
NOTE	95p.; This document is part of a collection produced under the auspices of the California State University Institute for Teaching and Learning. The CSSU/ITL, created in 1988, facilitates a 20-campus systemwide network of faculty affiliates in response to the demand for improved teaching and learning in the college classroom.
AVAILABLE FROM	Institute for Teaching and Learning, Office of the Chancellor, The California State University, 400 Golden Shore, Long Beach, CA 90802-4275 (book: $15; video: $20; both: $30).
PUB TYPE	Reports - Evaluative/Feasibility (142) -- Collected Works - General (020)
EDRS PRICE	MF01/PC04 Plus Postage.
DESCRIPTORS	Administrator Attitudes; Case Studies; *College Outcomes Assessment; Equal Education; *Evaluation Research; Higher Education; Methods; Outcomes of Education; Program Effectiveness; *Program Evaluation; State Universities; Student Evaluation
IDENTIFIERS	California State University; California State University Inst for Teach Learn; *College Teaching and Learning Collection; Faculty Attitudes

ABSTRACT

This monograph is a collection of papers that emerged from a project evaluating the implementation of college outcomes assessment at the campuses of the California State University (CSU) system. Fifteen pilot projects integrated their outcomes assessment in the academic majors and in general education from 1986 to 1990 and the projects were then evaluated to validate project results for dissemination to other institutions. Chapters 1 and 2, "Evaluation of Student Outcomes Assessment Pilot Projects in the California State University" (Matt L. Riggs and Joanna S. Worthley) and "Lessons from Pilot Projects" (Worthley and Riggs), present the findings of that 2-year evaluation including description of the study design and data and lessons gleaned from the study. Chapter 3, "The Evolution of Student Outcomes Assessment: Politics and Collegiality" by Bernard Goldstein and Frank Young, documents the policy environment for CSU's faculty-driven, campus-based approach to assessment. Chapters 4 and 5 provide case studies of the transformation that occurred at CSU Fresno and Sonoma State University; the chapters are titled: "Portfolios in the Major: A 'Success' Story" (Lu Mattson) and "Integrated Program Review: Lessons from Assessment Research at California State University, Fresno" (Priscilla Chaffe-Stengel). Chapter 6, "Assessment and Equity in the California State University" (Ming Lee) describes SCU systemwide assessment practices and provides examples of how assessment programs may be adapted to meet various needs of a diverse student population. The last chapter is by Ralph A. Wolff and is titled: "CSU and Assessment--Second Down and Eight Yards to Go: A View from the Scrimmage Line." This chapter describes patterns that have emerged from the cycle of campus visits and offers concluding recommendations for campus assessment programs. Appendixes list members of the project steering committee and campus pilot directors. (Author/JB)

Student Outcomes Assessment: What Makes It Work?

Assessment Practices & Experiences in the California State University

CSU Institute for Teaching and Learning and ERIC Clearinghouse on Higher Education

The California State University Institute for Teaching and Learning (CSU/ITL) facilitates a 20-campus network of teaching and learning programs in the CSU system. ERIC/HE has entered into an agreement with CSU/ITL to process documents produced by the system and create a mini-collection within the ERIC database.

Major objectives of this initiative are as follows:

- increase awareness of the work of the CSU Institute for Teaching and Learning;
- increase access to the work of CSU/ITL affiliates;
- begin to build a subset of information on teaching and learning that supports *The National Teaching and Learning Forum (NTLF)*, ERIC/HE's newsletter;
- encourage use of the ERIC system by CSU/ITL member affiliates and the *NTLF* readership; and
- test a model for collaboration between ERIC/HE and a major higher education system.

All CSU/ITL ERIC RIE citations are tagged with the following identifiers appearing in the IDEN:Field:

- College Teaching and Learning Collection; and
- California State University for Teaching and Learning.

All CSU/ITL citations carry the following statement in the Note Field:

> This document is part of a collection produced under the auspices of the California State University Institute for Teaching and Learning. The CSU/ITL, created in 1988, facilitates a 20-campus systemwide network of faculty affiliates in response to the demand for improved teaching and learning in the college classroom.

Student Outcomes Assessment: What Makes It Work?

Assessment Practices and Experiences in the California State University

Published by the California State University
Institute for Teaching and Learning

This publication was supported by a grant from the
United States Department of Education
Fund for the Improvement of Postsecondary Education

California State University
Institute for Teaching and Learning
400 Golden Shore, Suite 132
Long Beach, California 90802-4275

Table of Contents

Foreword

Since about 1985, by most accounts, outcomes assessment has been a part of the national landscape of higher education. What difference has it made? How can we design and implement it to make a difference? Certainly over the past seven years there has been no lack of anecdotal answers to these important questions from both researchers and practitioners. But there have been relatively few attempts to systematically study assessment's implementation on campus to substantiate these answers. I am happy to introduce this volume as one of the few such efforts. Its lessons ring true in the light of substantial field experience, and the volume should have much to offer those embarking on the tricky business of moving assessment from intriguing theory to campus practice.

Why is the California State University (CSU) system's experience in assessment worth looking at? The most compelling answer is because it is so typical. The CSU began its efforts in assessment in 1986, closely following the rise of the topic to national prominence. Over five years, moreover, its approach to assessment has mirrored the national picture in both tone and technique. Looking at the CSU experience systematically, as this study does, provides us with a valuable and authentic cross-section of national experience. There are four reasons, I think, why this is so.

First, the institutions that constitute the CSU, and thus the "study group" of this monograph, broadly represent the setting where most assessment activity took place in the late 1980s. In the period 1985-89—illustrated by the efforts of such institutions as Northeast Missouri State University, Kean College of New Jersey, SUNY Plattsburgh, and West Georgia College—the center of gravity for assessment was located in public comprehensive universities. Like these national trailbreakers, the CSU consists of institutions whose primary assigned mission is teaching, but that are, nevertheless, complex and multifunctional. At the same time, the student body served by the CSU is unusually diverse—both in terms of social background and tested ability levels. These are "mainstream" universities in virtually every respect.

Second, the manner in which the CSU as a system approached assessment was broadly typical of late 1980s practice. As a state, California has no legislative or board mandate for assessment. But like many governing and coordinating boards throughout the nation, CSU's central administration chose to take preemptive action on an assessment policy rather than waiting for the contents of such a policy to be politically determined. The resulting CSU system policy allowed broad campus discretion with respect to the choice of instruments and assessment approaches. And its clear rhetoric of instructional improvement was representative of the majority of state-level assessment approaches in the period, as exemplified by such states as Virginia, Colorado, and Missouri. As in many of these states, the CSU system first wrestled with a set of assessment principles that allowed each campus to

develop a process best suited to its culture, curriculum and clientele, but that also required each to provide a plan and to show results. Like many states in this period, moreover, the CSU system provided additional resources to directly support assessment and played a convening role for its campuses on assessment through a series of well-planned annual conferences.

Third, like states and systems across the nation, the CSU found an important partner in its regional accrediting organization, the Western Association of Schools and Colleges (WASC). With its adoption of an explicit criterion on institutional effectiveness in 1988 and its call for the construction of a "culture of evidence" on its member campuses, WASC approached assessment in the same decentralized spirit as did the CSU system. Like the earlier partnership between the Virginia Council on Higher Education and the Southern Association of Colleges and Schools (SACS), moreover, WASC and the CSU could play off one another to help induce local action. Particularly important—and exemplified by this monograph's final chapter—has been WASC's role as "friendly critic" of CSU's assessment efforts.

Fourth, like most state systems, the CSU is currently facing hard financial times and, as in leading states such as Virginia and New Jersey, hard times hit assessment in the midst of its implementation. In the short run, this forced CSU campuses to search for efficiency as well as elegance in data collection—for instance, in Fresno's use of low-cost surveys to reliably supplement more direct but costly assessments of the development of student writing proficiency. In the longer run, hard times have forced CSU institutions to think about assessment as more than an "add-on" to existing practices. Like their counterparts elsewhere, these institutions have increasingly been forced to recognize that if assessment is to make a difference it must become a part of instruction rather than simply being applied to instruction.

If the setting for assessment provided by CSU institutions is nationally representative, what of the lessons learned? Here, the contents of this monograph provide broad empirical confirmation for much of what has been written about the factors responsible for successful campus implementation—plus a few notable surprises. CSU experience, for instance, highlights the importance of faculty participation as a necessary condition for success. But findings also reinforce suspicions that assessment must be led, and that it is futile to wait for complete faculty "buy-in" before proceeding. The two campus case studies included in this volume nicely illustrate the point: Both involved a careful balance between keeping faculty informed and keeping the project moving. Complementing this finding is a confirmation of the overwhelming importance of administrative support. While faculty involvement can eventually be built, visible administrative commitment represents a crucial, and largely nonsubstitutable, precondition for success.

Findings from the CSU study also underline a simple but often overlooked axiom about assessment: Because it is founded on evidence, the technical adequacy of the information provided is critical to success. Not only is this true at the operational level—bad information, after all, will often

result in mistaken action—but it is true at the political and organizational level as well. Above all, if they are to result in positive change, assessment results must be believable—and believable to a faculty audience particularly schooled in the art of picking holes in evidence. As the monograph's authors point out, this finding has several implications. The most obvious is the need for time and careful planning in developing appropriate instruments and approaches. But another is the need for adequate, ongoing technical training and support for campus practitioners, most of whom do not begin their involvement in assessment as professionals in fields such as evaluation or educational measurement.

A particularly heartening finding from CSU experience is that shortages of resources did not preclude successful implementation. Both case studies provide ample support for the proposition that, in the words of one participant, "good assessment doesn't need to be expensive." Far more important appears to be the establishment of good channels of communication—particularly with respect to the dissemination of results. CSU experience reinforces the need for such dissemination to be continuous, indeed repetitive. As practitioners in many other states have often found to their dismay, new people are constantly entering the assessment conversation and must be appropriately reoriented to what is happening. Because paying attention to this requirement is frustratingly tedious, many of us are tempted to leave it out. But however unexciting, it appears a necessary element of a successful campus program.

In sum, the nineties appear far different for assessment than were the eighties. While the last decade involved innovation, experiment, and often generous support for "add-on" activities such as assessment, the next will increasingly require careful integration of assessment into existing teaching and learning activities at the departmental and classroom levels. CSU's experiences, as documented in this volume, provide practitioners with a visible bridge between these two quite different eras. We are grateful to both the institutions and to the Fund for the Improvement of Postsecondary Education (FIPSE) for the opportunity to share them.

Peter T. Ewell
Senior Associate
National Center for Higher Education
Management Systems

Introduction

During the 1980s, while the national educational reform movement was gaining momentum, campuses of the California State University (CSU) began to experiment with new methods for assessing the outcomes of college education and using the results to improve teaching and learning. Because of the size and diversity of the twenty-campus California State University system, it provided an ideal testing ground for new concepts of connecting learning outcomes to curriculum improvement and involving faculty in studying the teaching and learning process. This volume is a collection of knowledge that has emerged from implementation of outcomes assessment in the CSU.

From 1986 to 1990, the CSU Chancellor's Academic Program Improvement Grants (API) provided funding for fifteen campus pilot projects integrating outcomes assessment in the academic majors and in general education. There were projects involving theatre, economics, biology, mathematics, history, English, nursing, gerontology, anthropology, psychology, political science, sociology, liberal studies, and interdisciplinary general education. Several projects examined student outcomes on a campuswide level, and many involved collaboration with faculty at other universities or community colleges. API also sponsored two statewide conferences to familiarize faculty with developments in assessment at the national level and to disseminate successful campus program models.

In 1989, with assistance from the U.S. Department of Education, Fund for the Improvement of Postsecondary Education (FIPSE), the CSU Institute for Teaching and Learning undertook an evaluation of the fifteen campus pilot projects. We particularly wanted to find out what factors contributed to successful implementation of outcomes assessment on campus, and to validate the project results for dissemination to other institutions. Matt Riggs and Joanna Worthley of CSU San Bernardino conducted that two-year study. Their findings are contained in Chapters 1 and 2, "Evaluation of Student Outcomes Assessment Pilot Projects in the California State University" and "Lessons from Pilot Projects."

While individual campuses, schools, departments, and faculty experimented with assessment, a Chancellor's Advisory Committee on Student Outcomes Assessment was examining emerging state and university policy issues and developing principles and guidelines for implementing assessment in the CSU. Chapter 3, "The Evolution of Student Outcomes Assessment: Politics and Collegiality" by Bernard Goldstein and Frank Young, documents the policy environment for CSU's faculty-driven, campus-based approach to assessment.

Our most important challenges lie in making assessment work on a day-to-day basis at the campus. The practical considerations in integrating outcomes assessment into the academic culture must be addressed, particularly in times of budgetary constraints. Chapters 4 and 5 provide inspiring case

studies of the transformations that have occurred at two CSU campuses: "Portfolios in the Major: A 'Success' Story?" by Lu Mattson, and "Integrated Program Review: Lessons from Assessment Research at California State University, Fresno" by Priscilla Chaffe-Stengel.

Because of the rich cultural diversity represented by CSU students, we are particularly aware of the need for multicultural approaches which ensure that all students benefit from the assessment process. Chapter 6 by Ming Lee, entitled "Assessment and Equity in the CSU," describes CSU systemwide assessment practices and provides examples of how assessment programs may be adapted to meet various needs of a diverse student population.

While we continue to work hard on assessment issues in the CSU, there are many areas we would like to improve. For example, we must address new institutional accreditation requirements regarding assessment. Ralph Wolff's perspectives in Chapter 7, "CSU and Assessment—Second Down and Eight Yards to Go: A View from the Scrimmage Line," provide a forward look at the challenges faced by our university campuses in bringing outcomes assessment to full implementation.

The CSU has been fortunate to have an abundance of cooperation and support for our efforts in implementing student outcomes assessment. We are especially grateful to the Fund for the Improvement for Postsecondary Education for sponsoring our evaluation study and the dissemination of the findings. The FIPSE Project Steering Committee provided expert guidance on the goals and the scope of the evaluation, and the interpretation of the results. Other state systems of higher education were consulted during our study, and their insights and suggestions were very helpful. We also owe a great deal to the faculty pilot project directors, who gave freely of their time in providing information for the evaluation and in preparing their projects' results for national dissemination.

We publish this book with the hope that sharing our experience may help other institutions and faculties to implement their own assessment initiatives.

Helen Roberts
State University Dean for Instructional Improvement
California State University
March 1992

Chapter 1

Evaluation of Student Outcomes Assessment Pilot Projects in the California State University

Matt L. Riggs
Assistant Professor of Psychology
California State University, San Bernardino
and
Joanna S. Worthley
Assistant Professor of Psychology
California State University, San Bernardino

Introduction

The current assessment movement in higher education is driven by the wary partnership of reform and accountability, a partnership yielding a complex and diverse collection of assessment activities in university settings (Ewell, 1991). Over the last several years, the California State University (CSU) system has moved to construct an assessment agenda which responds to both reform and accountability in ways that will preserve the commitment of the CSU to intellectual and programmatic diversity. Beginning in 1986, Academic Program Improvement grant funds have supported a series of student outcomes assessment projects initiated by faculty in a variety of disciplines on 11 campuses of the CSU. These pilot projects, under the aegis of the CSU Institute for Teaching and Learning, were aimed at demonstrating how student outcomes of various kinds can be used to assess the effectiveness of General Education and baccalaureate degree programs.

The present study, sponsored by the Fund for the Improvement of Postsecondary Education (FIPSE), provides a "meta-assessment" of the 15 pilot projects. These projects, briefly detailed in Table 1, have developed and field-tested a variety of assessment measures/instruments, including portfolios, interviews, senior/capstone projects, surveys, and examinations. (A further description of these projects is provided in Chapter 2)

Table 1

California State University Pilot Projects in Student Outcomes Assessment

Campus	Project Title & Focus
CSU Hayward, San Francisco SU, San Jose SU	Assessment of Majors: A Three-Campus, Three-Discipline Model Focus: Development of comprehensive examinations for seniors in biology, economics, and mathematics
CSU Fresno	Assessment of Undergraduate Reading Competence Focus: Assessment of student reading strategies and competence related to course assignments and library skills
CSU Fresno	Assessment of Undergraduate Writing Competence Focus: Assessment of student performance on the Upper Division Writing Exam as a function of course exposure and language proficiency
CSU Bakersfield	Knowledge and Attitudes in General Education: A CSU-Community College Joint Assessment Focus: Assessment of impact of GE course in Western civilization on students' knowledge and values
CSU Fullerton	Student Outcomes Related to Curricular Variety in Gerontology Focus: Development of a model for cross-campus assessment of outcomes for interdisciplinary programs in gerontology
CSU Hayward	Assessment of Student Outcomes: A Basic Writer's Writing Program Focus: Development of a model for assessing outcomes in the Intensive Learning Experience writing course sequence
CSU Northridge	Student Outcomes Assessment in Academic Program Improvement in Theater Focus: Development of a performance-based mastery test for summative and formative assessment of student achievement in theater
San Francisco SU	Assessing Outcomes for English Teacher Candidates Focus: Development of an "assessment course" to evaluate the subject-matter competency of teacher credential candidates in English language arts
CSU Chico	Outcomes Assessment of Four Classes of Nursing Graduates Focus: Development of a multimeasure assessment of nursing program graduates to identify trends in program effectiveness from 1983 to present
Sonoma SU	Integrating Student Outcomes Assessment into the Curriculum Focus: Development of a portfolio system to assess formative and summative outcomes for students in an interdisciplinary liberal studies program
CSPU Pomona	Enhancing Quality by Assessment: A General Education Project Focus: Development of a comprehensive assessment program for an interdisciplinary General Education program
CSU Bakersfield	An Empirical Evaluation of Five Baccalaureate Social Science Programs Focus: Development of a model to conduct longitudinal assessments of student performance and perceptions of degree programs in anthropology, economics, political science, psychology, and sociology
San Diego SU	Student Outcomes Assessment: Liberal Studies Major Focus: Development of a multimeasure assessment program for student outcomes in liberal studies

This "meta-assessment" of the projects was designed to identify critical features across the 15 projects which were consistently associated with effective assessment. Once defined, these features might be expected to facilitate wider implementation of the assessment strategies field-tested experimentally. Because they represented a broad range of disciplinary perspectives, the pilot projects offered an opportunity to define factors which predicted assessment effectiveness across curricular boundaries.

Method

The study used a multi-site method to evaluate factors that might have contributed to the outcomes achieved. A three-part framework was used to specify variables and to organize the data. Relevant variables were conceptualized within the categories of assessment environment, assessment methods, and assessment outcomes.

Assessment Environment

The assessment environment was defined as the social and organizational setting in which each project occurred. Variables relevant to the determination of this construct included:

(1) general faculty participation in the planning of the project
(2) faculty participation in the implementation of the project
(3) the faculty's perceived "ownership" of the project (i.e., self-determination)
(4) faculty consensus with the project plan
(5) faculty workload required by the project
(6) the project director's experience in assessment activities
(7) the faculty's experience in assessment activities
(8) the project director's general academic experience
(9) adequacy of budget, supplies and resources
(10) administrative support
(11) student support
(12) previous experience with outcomes assessment
(13) the content domain (e.g., physical science)
(14) the nature of the intended audience for the project's results

Assessment Methods

Assessment methods were defined by variables that described the strategic aspects of each project. General procedures with relevance across curriculums were selected. These variables included:

(15) goal definition
(16) selection of appropriate outcomes
(17) the psychometric adequacy of measures used or developed
(18) the success of data collection and reporting
(19) the appropriate use of statistical analyses

(20) sensitivity to multicultural issues
(21) the use/development of multiple outcome measures
(22) the comprehensiveness of reports describing the results of the project
(23) the utility/economy of procedures used

Assessment Outcomes

Outcome variables were used to capture the systematic growth and change that could be attributed to the assessment project. The following "direct" and "indirect" outcomes of the assessment projects were evaluated:

DIRECT

(24) the survival of the project
(25) attainment of additional funding
(26) gains in student achievement
(27) curricular development
(28) improved teaching
(29) better student feedback
(30) the use of new methods of assessment
(31) improved use of existing databases
(32) the dissemination of results

INDIRECT

(33) improved student recruitment/retention
(34) better general attitudes toward assessment
(35) new sources of money for assessment
(36) higher visibility of assessment
(37) external adoption of measures or methods developed

Conceptual Model

The conceptual model guiding the analysis is based on the assumption that key environmental and methodological variables will determine the nature of outcome variables. This model is illustrated in Figure 1. Included in each of the categorical "bins" are the specific variables that were addressed.

Data Collection

Information relevant to the assessment of the variables described above was obtained from multiple sources. First, copies of all project reports were obtained. All reports were studied and assessed by both researchers conducting this analysis. Independent conclusions concerning the "presence/absence" or "adequacy/inadequacy" of key variables were compared and evaluated. Where consensus on variable indicators could not be reached, that source of information was dropped from consideration. This process was especially important to the assessment of methodological variables (variables #15 to #23).

Figure 1
Conceptual Model

ENVIRONMENTAL FACTORS

FACULTY INVOLVEMENT
- Planning Participation
- Faculty Participation in Project
- Faculty Ownership
- Consensus with Plan
- Perceived Faculty Workload

TRAINING/EXPERIENCE
- Director's Training
- Faculty Training
- Director's Academic Experience

SUPPORT VARIABLES
- Budget and Supplies
- Administrative Support
- Student Support

EXISTING PROCEDURES
- Previous Assessment

PROJECT FOCUS
- Content Domain
- Breadth of Audience

METHOD FACTORS

GENERAL PROCEDURAL ADEQUACY
- Goal Definition
- Selection of Outcomes
- Measures Developed
- Data Collection/Reporting
- Measurement Properties
- Statistical Analysis

PROJECT COMPREHENSIVENESS
- Multicultural Issues
- Development of Multiple Measures
- Report Comprehensiveness

COST EFFECTIVENESS
- Utility/Economy

OUTCOMES

DIRECT OUTCOMES
- Project Continuing
- Additional Funding
- Gains in Student Achievement
- Curricular Impact
- Better Teaching
- Feedback to Students
- New Measures Developed
- Better Data Use
- Dissemination of Results

INDIRECT OUTCOMES
- Recruitment/Retention
- Attitudes Toward Assessment
- New Moneys for Assessment
- Visibility of Assessment
- External Adoption

The second, and perhaps most utilized, source of data was the project directors themselves. A survey was developed and administered to each of the project directors. Results of this survey were especially critical in determining the nature of environmental variables for each project. These results were also primary determinants of the project outcome variables.

Project directors were further utilized throughout the data collection period to fill in missing information and to provide updates on developments not available in the final project reports. A meeting with directors held in the third month of the data collection process yielded additional information on the assessment environment at individual campuses.

Telephone interviews with project directors continued into the final weeks of the data collection process. A special effort was made to obtain additional focus concerning multicultural issues relevant to outcomes assessment.

The final source of information utilized was the reports submitted by the external evaluators of the pilot projects. These were read last in an effort to maintain the objectivity of the researchers through the initial stages of the data collection. These reports were especially useful in supplementing observations concerning the methodological variables.

Analysis

Data organization was achieved in a manner consistent with the recommendations for qualitative data analysis prescribed by Miles and Huberman (1984). All observations, comments, and survey scores were condensed and entered in a highly abbreviated form onto a "meta-matrix." This master chart contained information relevant to all 37 variables for all projects considered complete enough for inclusion in the final assessment. In early versions of the chart, the basic principle was inclusion of all relevant data.

For the final analysis, projects which lacked comprehensive data were excluded. One project assessed three different disciplines on three different campuses. Because the five sources providing information for this analysis were describing discrete events from what were apparently very different experiences, these data sources were treated as five different "sites" for this analysis. The total "n" of sites ultimately used was sixteen.

On the basis of information represented in this matrix, a qualitative categorization of all projects on all variables was completed. For each project, all variables were classified as:

4 - strongly present/achieved,
3 - partly present/achieved,
2 - weakly present/achieved,
1 - absent/not achieved.

The results of this classification are presented in Table 2. Projects are presented in random order. This display is referred to by Miles and Huberman (1984) as a "Site-Ordered Descriptive Meta-Matrix." Though the "values" are

Table 2
Project Raw Scores on All Variables

	PROJECTS							
VARIABLES	1	2	3	4	5	6	7	8
Planning Participation	**1**	**2**	**2**	**1**	**2**	**3**	**1**	**2**
Fac. Part. in Project	**2**	**4**	**2**	**1**	**2**	**2**	**4**	**2**
Faculty Ownership	**1**	**4**	**2**	**1**	**1**	**2**	**3**	**1**
Consensus with Plan	**3**	**4**	**2**	**1**	**3**	**2**	**2**	**3**
Perceived Workload	**3**	**4**	**3**	**1**	**1**	**2**	**2**	**2**
Director's Training	3	3	3	1	2	3	3	3
Faculty Training	2	4	1	1	1	3	2	1
Director's Acad. Exp.	4	4	3	4	1	3	4	3
Budget and Supplies	**3**	**3**	**3**	**3**	**2**	**3**	**1**	**1**
Administrative Support	**2**	**3**	**4**	**1**	**4**	**3**	**4**	**3**
Student Support	**3**	**3**	**4**	**1**	**4**	**3**	**3**	**2**
Previous Assessment	1	4	2	4	1	1	1	3
Mesh w. Existing Prgs.		3	3					1
Content Domain	**1**	**1**	**1**	**1**	**2**	**3**	**3**	**1**
Breadth of Audience	**2**	**2**	**2**	**2**	**2**	**3**	**2**	**3**
Goal Definition	2	2	2	2	2	3	3	2
Selection of Outcomes	2	2	2	2	2	3	3	2
Measures Develped	2	2	2	2	2	1	3	3
Data Colletn/Reporting	2	2	2	2	2	1	3	2
Measurment Properties	1	1	1	1	1	1	1	1
Stat. Analysis	1	1	1	1	1	1	1	2
Multicultural Issues	**1**	**1**	**1**	**1**	**1**	**1**	**2**	**1**
Develop. of Mult. Meas.	**3**	**4**	**3**	**3**	**3**	**1**	**4**	**4**
Report Comprehensvns	**2**	**2**	**2**	**2**	**2**	**2**	**2**	**3**
Utility/Economy	2	2	2	2	2	1	2	2
Project Continuing	**3**	**1**	**3**	**1**	**1**	**1**	**4**	**4**
Additional Funding	**3**	**1**	**1**	**2**	**1**	**1**	**4**	**1**
Gains in Studnet Ach.	**1**	**2**	**1**	**1**	**4**	**1**	**3**	**2**
Curricular Impact	**2**	**3**	**1**	**2**	**2**	**3**	**4**	**3**
Better Teaching	**2**	**2**	**2**	**1**	**4**	**3**	**3**	**3**
Feedback to Students	**3**	**4**	**1**	**3**	**4**	**1**	**3**	**3**
New Measures Developd	**3**	**3**	**3**	**2**	**3**	**1**	**3**	**3**
Better Data Use	**1**	**2**	**1**	**1**	**3**	**1**	**3**	**1**
Disseminatn of Results	**1**	**4**	**1**	**1**	**1**	**3**	**4**	**3**
Recruitment/Retention	1	2	1	1	1	2	3	2
Attitudes towd Assmnt	1	3	2	1	2	2	3	2
Moneys for Assessment	1	2	1	1	1	2	1	2
Visibility of Assmnt	1	3	3	1	1	2	4	4
External Adoption	1	2	1	1	1	2	1	1
Avg. Direct Outcomes	**2.1**	**2.4**	**1.6**	**1.6**	**2.6**	**1.7**	**3.4**	**2.6**
Avg. Indir. Outcomes	**1.0**	**2.4**	**1.6**	**1.0**	**1.2**	**2.0**	**2.4**	**2.2**
Overall Avg. Outcomes	**1.6**	**2.4**	**1.6**	**1.3**	**1.9**	**1.8**	**2.9**	**2.4**

Table 2 (cont'd)
Project Raw Scores on All Variables

VARIABLES	PROJECTS 9	10	11	12	13	14	15	16
Planning Participation	3	1	4	2	1	1	4	4
Fac. Part. in Project	3	2	4	3	2	4	1	4
Faculty Ownership	3	1	4	2	1	2	1	4
Consensus with Plan	3	3	4	3	1	3	1	3
Perceived Workload	3	4	2	4	4	4	1	1
Director's Training	4	4	4	3	4	2	3	3
Faculty Training	3	1	4	2	1	1	1	3
Director's Acad. Exp.	3	3	2	4	2	2	3	3
Budget and Supplies	3	4	1	4	4	4	2	3
Administrative Support	3	4	4	3	4	4	3	4
Student Support	4	2	3	2	3	4	2	4
Previous Assessment	4	2	4	4	4	1	1	1
Mesh w. Existing Prgs.	3	4	2	2	3			
Content Domain	1	2	1	3	2	2	3	3
Breadth of Audience	2	4	2	3	4	2	1	4
Goal Definition	4	4	4	2	4	2	4	4
Selection of Outcomes	4	4	4	2	4	4	4	4
Measures Develped	3	4	3	2	4	4	3	3
Data Colletn/Reporting		4	3	2	4	4	3	3
Measurment Properties		3	2	1	3	3	2	4
Stat. Analysis		3	1	2	3	3	2	3
Multicultural Issues		4	3	4	4	4	3	4
Develop. of Mult. Meas.	4	4	4	4	4	4	4	3
Report Comprehensvns		4	3	3	4	3	3	3
Utility/Economy	3	2	2	2	2	2	3	2
Project Continuing	3	4	4	3	3	1	1	3
Additional Funding	1	3	4	1	1	1	1	3
Gains in Studnet Ach.	1	1	2	1	1	4	1	2
Curricular Impact	3	3	1	3	3	4	1	3
Better Teaching	3	3	3	3	3	4	1	
Feedback to Students	4		4	3	3	4	1	
New Measures Developd	4	4	4	3	4	4	1	3
Better Data Use	4	4	1	4	4	4	1	1
Disseminatn of Results	4	4	3	3	4	4	1	4
Recruitment/Retention	1		2	1		1	1	2
Attitudes towd Assmnt	3	3	3	3	3	3	1	
Moneys for Assessment	1	2	1	1		1	1	
Visibility of Assmnt	3	4	4	3	4	4	1	
External Adoption	1		3	3	4	1	1	
Avg. Direct Outcomes	3.0	3.3	2.9	2.7	2.9	3.3	1.0	3.2
Avg. Indir. Outcomes	1.8	3.0	2.6	2.2	3.7	2.0	1.0	2.0
Overall Avg. Outcomes	2.4	3.1	2.7	2.4	3.3	2.7	1.0	2.6

the results of qualitative assessment and therefore include a certain unavoidable level of judgment, almost all are based on multiple sources of information. The agreement between sources was generally very compelling.

Using the proposed research conceptualization, project outcomes (direct and indirect) would logically represent dependent measures. Environmental and methodological variables would be the multiple predictors. If the data were interval, and the n of projects greater, the analysis of choice would be multiple regression. In a multiple regression paradigm, critical factors that should be attended to in any replication would be those that accounted for the most variance in the dependent measure (program outcomes).

However, the data presented in Table 2 is qualitative, and the n of projects is sixteen. Nevertheless, the problem is conceptually the same. One must identify the variables that were consistently associated with project effectiveness. The task is to determine which predictor variables most consistently co-vary with outcomes. Good predictors are those whose qualitative categorization is most consistent with the qualitative categorization of project outcomes.

Because a project's relative success should not be judged by a single outcome variable, and because there should be some resistance to the temptation to overanalyze qualitative data, outcomes were combined. Subsequent analyses were conducted using: (1) the average of all outcomes; (2) the average of outcomes categorized as direct outcomes; and (3) the average of indirect outcomes.

The next step necessary to enable comparison across variables is a standardization of scores. Though the "scale" used for classification of all predictor and outcome variables was the same (1 through 4), the mean and standard deviation of the resulting distributions for each variable were different. This was resolved by transforming all scores to z-scores. These scores represent the project's deviation above or below each given variable's mean in standard deviation units.

"Consistency" between predictors and outcomes was assessed by computing the squared deviations between each standardized score and the average standardized "outcome" (overall, direct, and indirect) for the project associated with that score.

Finally, these squared deviations were summed across project sites, and divided by the number of projects to produce a value conceptually similar to variance (the difference being that the source of deviation was between each predictor score and its project's outcome score rather than between each predictor score and the average of that predictor score across sites). The resulting "variance with outcomes" across sites will be smallest for those variables that were most consistently related to outcomes. Larger values indicate that the classifications assigned to that variable were not as predictive of the project outcomes.

Results of this analysis are presented in Tables 3, 4, and 5. Because of the limited nature of the data, there can be no assessment of statistical probability or significant differences. The predictor variables are simply ranked in the order of their consistency with outcomes: overall, direct, and indirect. Top-ranked variables are those whose classifications were most consistent with project outcomes.

These results are summarized in Table 6. This table enables comparison of the relative importance of predictors across different combinations of project outcomes. Predictors are also organized into groups by similarity.

Results and Discussion

ENVIRONMENTAL PREDICTORS

Faculty Involvement Variables

Two faculty variables were consistently toward the top of the respective rank orderings. "Faculty Participation" and "Perceived Faculty Workload" appear to be important indicators of assessment success. The first is most closely associated with those outcomes identified as "direct" outcomes. The second is more closely associated with "indirect" outcomes. "Faculty Ownership" and "Faculty Consensus with Plan" were also relatively good indicators, especially for direct outcomes. "Faculty Participation in Planning" appears to have little to do with project outcomes.

It was expected that faculty involvement would be critical to project outcomes. Consequently, it was no surprise that faculty participation in the project was ranked number one in its association with the direct, intended outcomes of each project. Faculty consensus with the project plan, ownership of the project, and perceived workload were also ranked in the top ten in reference to direct outcomes.

Perceived workload appeared especially critical to indirect outcomes. Perhaps as workload increases, it becomes less likely that indirect outcomes will be realized. Workload would also seem especially salient to the development of attitudes toward outcomes assessment (one of the indirect outcomes).

Somewhat surprising was the apparent unimportance of involving faculty in the early stages of project planning. In the general management literature, it is often suggested that participative planning of projects is a good strategy for insuring project participation. Apparently in these academic settings, it was only critical that you do an adequate job of "selling" your plan once it was developed.

Training/Experience with Student Outcomes Assessment

Project outcomes were associated with the project director's efficacy in outcomes assessment, especially indirect outcomes. Faculty training and the project director's general academic background and experience were not as closely tied to project results.

Table 3

Predictor Variables Rank Ordered by Consistency with Overall Outcomes

Predictor Variables	Rank Order	Variance with Overall Outcomes
Adequacy of Measures Development	1	.73
Intended Breadth of Audience	2	.75
Administrative Support	3	.76
Faculty Participation in Project	4	.77
Director's Training in SOA	5	.91
Perceived Faculty Workload	6	.95
Sensitivity to Multicultural Issues	7	1.06
Development of Multiple Measures	8	1.10
Selection of Outcomes	9	1.12
Measurement Properties	10	1.22
Statistical Analysis	11	1.25
Faculty Consensus with Plan	12	1.27
Project Goal Definition	13	1.28
Faculty Ownership of Project	14	1.31
Data Collection/Reporting	15	1.39
Existence of Previous SOA	16	1.50
Faculty Training in SOA	17	1.60
Student Support of Project	18	1.60
Adequate Budget/Supplies	19	1.71
Content Domain	20	1.79
Report Comprehensiveness	21	1.80
Utility/Economy of Project	22	2.32
Director's Academic Experience	23	2.41
Faculty Participation in Planning	24	2.43

Table 4

Predictor Variables Rank Ordered by Consistency with Direct Outcomes

Predictor Variables	Rank Order	Variance with Direct Outcomes
Faculty Participation in Project	1	.56
Adequacy of Measures Development	2	.80
Administrative Support	3	.89
Faculty Consensus with Plan	4	1.00
Intended Breadth of Audience	5	1.09
Development of Multiple Measures	6	1.09
Faculty Ownership of Project	7	1.23
Selection of Outcomes	8	1.23
Perceived Faculty Workload	9	1.25
Sensitivity to Multicultural Issues	10	1.27
Student Support of Project	11	1.31
Director's Training in SOA	12	1.32
Measurement Properties	13	1.37
Statistical Analysis	14	1.48
Project Goal Definition	15	1.54
Faculty Training in SOA	16	1.63
Data Collection/Reporting	17	1.64
Content Domain	18	1.83
Existence of Previous SOA	19	1.85
Adequate Budget/Supplies	20	1.86
Utility/Economy of Project	21	2.07
Report Comprehensiveness	22	2.38
Director's Academic Experience	23	2.41
Faculty Participation in Planning	24	2.45

Table 5

Predictor Variables Rank Ordered by Consistency with Indirect Outcomes

Predictor Variables	Rank Order	Variance with Indirect Outcomes
Intended Breadth of Audience	1	.66
Director's Training in SOA	2	.72
Perceived Faculty Workload	3	.86
Administrative Support	4	.89
Adequacy of Measures Development	5	.92
Sensitivity to Multicultural Issues	6	1.06
Project Goal Definition	7	1.17
Selection of Outcomes	8	1.17
Statistical Analysis	9	1.20
Faculty Participation in Project	10	1.22
Measurement Properties	11	1.24
Existence of Previous SOA	12	1.25
Development of Multiple Measures	13	1.29
Data Collection/Reporting	14	1.36
Report Comprehensiveness	15	1.44
Faculty Ownership of Project	16	1.53
Adequate Budget/Supplies	17	1.62
Faculty Training in SOA	18	1.66
Faculty Consensus with Plan	19	1.69
Content Domain	20	1.80
Student Support of Project	21	1.96
Faculty Participation in Planning	22	2.32
Director's Academic Experience	23	2.33
Utility/Economy of Project	24	2.49

Table 6
Variance between Predictors and Outcomes with Predictor Variable Rank Orders*

Predictor Variables	Variance with All Outcomes	Variance with Direct Outcomes	Variance with Indirect Outcomes
Faculty Participation in Planning	2.43 (24)	2.45 (24)	2.32 (22)
Faculty Participation in Project	0.77 (4)	0.56 (1)	1.22 (10)
Faculty Ownership of Project	1.31 (14)	1.23 (8)	1.53 (16)
Faculty Consensus with Plan	1.27 (12)	1.00 (4)	1.69 (19)
Perceived Faculty Workload	0.95 (6)	1.25 (9)	0.86 (3)
Director's Training in SOA	0.91 (5)	1.32 (12)	0.72 (2)
Faculty Training in SOA	1.60 (17)	1.63 (16)	1.66 (18)
Director's Academic Experience	2.41 (23)	2.41 (23)	2.33 (23)
Adequate Budget/Supplies	1.71 (19)	1.86 (20)	1.62 (17)
Administrative Support	0.76 (3)	0.89 (3)	0.89 (4)
Student Support of Project	1.60 (17)	1.31 (11)	1.96 (21)
Existence of Previous SOA	1.50 (16)	1.85 (19)	1.25 (12)
Content Domain	1.79 (20)	1.83 (18)	1.80 (20)
Intended Breadth of Audience	0.75 (2)	1.09 (5)	0.66 (1)
Project Goal Definition	1.28 (13)	1.54 (15)	1.17 (8)
Selection of Outcomes	1.12 (9)	1.23 (7)	1.17 (7)
Adequacy of Measures Development	0.73 (1)	0.80 (2)	0.92 (5)
Data Collection/Reporting	1.39 (15)	1.64 (17)	1.36 (14)
Measurement Properties	1.22 (10)	1.37 (13)	1.24 (11)
Statistical Analysis	1.25 (11)	1.48 (14)	1.20 (9)
Sensitivity to Multicultural Issues	1.06 (7)	1.27 (10)	1.06 (6)
Development of Multiple Measures	1.10 (8)	1.09 (6)	1.29 (13)
Report Comprehensiveness	1.80 (21)	2.38 (22)	1.44 (15)
Utility/Economy of Project	2.32 (22)	2.07 (21)	2.49 (24)

* Rank Orders in ( )

Of the three factors describing the training and background of project participants, only the project director's training and experience specific to student outcomes assessment ranked very high in association with outcomes. This factor was especially critical to the attainment of indirect outcomes. This result is logical given that the director was generally responsible for the adequacy of project goals and methods.

Support Variables

Administrative support was closely tied to project outcomes, both direct and indirect. Adequacy of budget, supplies, and other institutional resources was not closely tied to project results. Student support appears only moderately associated with direct outcomes.

Nothing happens without administrative support. In all hierarchies, policy tends to come from the top down. Though a goal of the CSU system is to establish faculty-initiated outcomes assessment, these initiatives will most likely fail without the support of academic administrators. These individuals are in the position to set the tone in reference to the academic legitimacy of such efforts. Unless outcomes assessment efforts are rewarded, it will be difficult to maintain motivation to assess student outcomes.

Student support did not rank high. Though this source of support was assessed through the eyes of the project directors rather than from the students themselves, this result does not appear unreasonable.

What does seem somewhat surprising is the apparent unimportance of adequate budget and supplies. Some insight into this result was obtained in discussions with project directors about how much of their project was coming "out of their hide." The efforts of some directors clearly went beyond what might have been expected given the modest budgets they were receiving for their administration of each project. This phenomenon was more likely to occur if the director was working in a content area for which outcomes assessment research could be considered legitimate professional development or for senior faculty members who were no longer struggling to achieve tenure or promotion.

Another possible explanation for the inconsistency between the adequacy of budget/supplies and project success would be the inherent differences in requirements for adequate assessment procedures across different content areas. It may be possible that adequate assessment can be achieved for relatively small costs in some disciplines, but will tend to be very expensive in others. Consequently, adequate results might be obtained for some even when resources are tight while others will find assessment prohibitive without adequate budgetary support.

Existing Student Outcomes Assessment Procedures

Whether or not the project was initiated in an environment that currently included some type of formal outcomes assessment did not appear to have a

critical impact on project outcomes. The good news of this result is that projects breaking new ground do not necessarily have to anticipate damaging levels of resistance. The bad news may be that previous experience with outcomes assessment apparently may not guarantee that new initiatives will be welcomed with open arms.

Project Focus

The content area in which the project occurred had little to do with the relative success of the project. However, the intended breadth of the audience (who and how many individuals/organizations I expect to learn about the results of my project) was very closely associated with outcomes, especially indirect outcomes.

There was some expectation that perhaps the content area in which the project occurred might have considerable impact on the relative success of the project. This may disappoint some educational or behavioral science researchers who might have liked to assume they had a corner on this part of the research market. This is a good result for a system that hopes to initiate outcomes assessment across a universe of content domains.

The tight association between the intended breadth of the audience for the project's results and project outcomes is somewhat surprising. In general, this association is logical in that the excellence of project outcomes should be related to how many individuals or organizations with which the director intends to share the results. This result may also reflect the director's experience and enthusiasm for student outcomes assessment. As previously observed, project director efficacy in outcomes assessment was closely associated with project outcomes.

METHODOLOGICAL PREDICTORS

General Procedural Adequacy

The single most critical variable from this category was the development of good measures of student outcomes. The creation of good measures was predictive of project results, direct and indirect. Appropriate selection of outcomes to measure was also closely associated with project outcomes.

"Project Goal Definition" and the adequacy of "Statistical Analyses" conducted were moderately tied to indirect outcomes. These variables along with "Data Collection/Reporting" and the reporting of "Measurement Properties" did not appear to co-vary tightly with direct project outcomes.

In reference to general procedural adequacy, the most important aspect of the assessment projects appeared to be the adequacy of the measures they developed or adopted. It is highly logical to expect that project success would hinge on the ability of the measures used to reliably and validly measure student outcomes. This process begins with the selection of appropriate outcomes to measure.

Other variables in this category, though not highly ranked in their association with direct outcomes, were toward the top of the list in relation to indirect outcomes. Much of this relationship appeared dependent upon the association with the indirect outcome of external adoption. Those projects bound for adoption were generally the most precise in the definition of project goals and most ambitious and accurate in the production of statistical analyses.

Project Comprehensiveness

The "Development of Multiple Measures" of student outcomes and "Sensitivity to Multicultural Issues" appeared moderately associated with project outcomes. The comprehensiveness of the reports made available were not closely related to project outcomes.

The relatively high ranking of sensitivity to multicultural issues and the development of multiple measures may again reflect the sophistication of the project director in reference to good outcomes assessment. It is also reasonable to expect projects that developed or used more than one form of assessment to be more successful.

Sensitivity to multicultural issues would appear especially relevant to the potential for recruitment and retention of under-represented students. This would partially explain this variable's relatively high ranking in reference to indirect outcomes.

The nature of final project reports varied widely. The low association between the comprehensiveness of project reports and project outcomes might indicate that much occurred that was not completely reported. Many directors indicated that deadlines fell before they had time to adequately process the project results. Some compensated by disseminating results via other channels (regional and national presentations, journal publications). Though the present study followed up and received many of these reports, it appears that much information from successful, comprehensive projects was lost to the CSU system due to the fact that final reporting was required before assessment procedures could be adequately evaluated.

Cost Effectiveness of Project

The "Utility/Economy" of the projects' assessment procedures was near the bottom of the rank ordering. Apparently expensive projects in terms of their dollars spent to students assessed were not always the richest in results.

This result is a perfectly reasonable result given the experimental nature of these projects. Directors "tried out" a wide variety of assessment procedures. There will also always be some cost differences across disciplines necessitated by the different nature of assessment processes that must occur. Some content areas may be able to conduct excellent and inexpensive assessment, while other areas may only achieve moderate results despite a high price tag.

Nevertheless, there is some logic to the assumption that efficient assessment will produce more desirable results. As more data is collected, this relationship should be reassessed within content areas.

Though qualitative analysis does not enable the same level of precision as might be obtained from more quantitative data, the observations made in this study were systematic, relatively objective, and almost always based on multiple sources. Additional testing and replications of this study's results are suggested, and could occur within the CSU's continuing program to develop and monitor programs of student outcomes assessment.

Major Conclusions

The major conclusions drawn from an overview of the study's results are organized in terms of the conceptual model's environmental and methodological variables.

Environmental Variables

The environmental var 'es critical to project outcomes were human resources. First, recruiting and maintaining faculty support was a key variable, but one which showed high variability across projects. The factors underlying this variability are difficult to detect in the quantitative measures, but they appear in the interview data from project directors. One concern in nearly all faculty groups is "the intended primary use of the outcomes data," particularly where data suggest evidence of teaching/program effectiveness. Another might be described as the worry over the human capital costs of department-level assessment activities. This turned up as a particular concern for junior faculty. Many junior faculty perceive that research on teaching and learning is regarded as "second-tier" research which may not be counted in the tenure/promotion process. As several respondents indicated, assessment activities favor two faculty groups: (1) those in social/behavioral disciplines; and (2) those whose professional research activities "fit" with assessment research.

Drawing from the experiences of the project directors, there appear to be several general guidelines for establishing and maintaining faculty participation:

- **Educate participants about the value of assessing student outcomes.** The motivation required to commit to outcomes assessment is dependent upon a general perception that this effort will provide payoffs The general descriptions of outcomes attained (refer to survey results) provide more than adequate examples of the benefits achieved by effective programs. On an individual basis, participation will be enhanced if assessment activities are an integral part of the faculty performance criteria.
- **Maintain local control of the project.** Support was never achieved in one project partly because the program was perceived as someone else's grand plan that got shoved down our throats." Support for

another project was in jeopardy when the local faculty senate returned the project's proposal as an external mandate.

- **Overcome the "threat" associated by many faculty with assessing outcomes.** Many directors reported the suspicion expressed by their faculty concerning the intended use of the data collected. Faculty often saw the potential for information obtained to be used as a club to punish programs or individuals rather than as a constructive tool for self-development.

The importance of the development of administrative support was consistently reported across projects. Data from the majority of respondents suggested that in-kind resources, publicity, campus-level coordination, and establishing a climate receptive to assessment initiatives were important positive contributions of administrative offices. Even in this positive environment, however, data indicate that more concrete evidence of integration of the assessment agenda in campus-level policy and in concrete recognition of assessment activities for professional development is needed.

Two general prescriptions can be made on the basis of director responses. The first deals with the establishment of administrative support; the second partially defines the nature of the support sought:

- **Educate administrators about the value of assessing student outcomes.** Just as executives of business organizations value economic indicators for their companies, educational administrators need to understand the potential value of performance feedback inherent in the assessment of student outcomes. They must also be sensitized to the need for discretionary, constructive use of such data. The validity of departmental reports will quickly erode if messengers delivering "bad news" are shot.

- **Administrators should recognize and reward the development of outcomes assessment in one's field as legitimate professional development.** Many directors reported that contribution of time to outcomes assessment was unrealistic for all but senior, tenured faculty who could afford to "waste some time." The term "waste" was used only in reference to the potential for recognition of these efforts by chairs, deans, and department/school/university evaluation committees in control of the retention, promotion, and tenure process.

Finally, project directors' training/experience in measurement and analysis was key to project effectiveness, and here there were important differences. Some project directors reported dismay over the difficulties in learning assessment procedures as the project progressed. A number of respondents echoed the sentiments of one director who felt that the project lost momentum "just as experience and proficiency began to develop."

These observations clearly suggest the importance of training in outcomes assessment for those administrating assessment development. Two general prescriptions flow from the comments made by project leaders:

- **Assess the training/experience of project directors specific to assessing student outcomes.** General knowledge and experience as an educator is not enough. This study's "assessment" of relevant knowledge was dependent upon self-report. This approach, however, appeared satisfactory, producing large variance in the levels of described outcomes assessment expertise.
- **Provide continued opportunities for training in and exposure to outcomes assessment.** Many of the directors reported that the systemwide conferences and reports on outcomes assessment had been responsible for kindling their interest in the process. If faculty-based outcomes assessment is to be institutionalized within the CSU, we must continue to share knowledge.

Methodological Variables

The key methodological variable was the development or adoption of adequate measures of student outcomes. Adequacy of measurement implied several aspects. The first is the simple psychometric properties of the assessment. This involves the reliability and validity of the assessment procedures. As an example of awareness of measurement adequacy, several directors did an excellent job of assessing the inter-rater reliability of judges producing qualitative assessments of student projects or papers. Others spent considerable time and consulted widely with their peers to evaluate the content validity of their assessments. This process often had positive, retroactive impact on curriculum and teaching strategies.

Multicultural sensitivity also contributed to the adequacy of measurement, especially in reference to the inferences drawn from assessment scores. A critical question for directors to ask was, "What assessment procedures will provide all students with an equitable opportunity to demonstrate their competence?" This consideration should result in the production of multiple, more creative indices that would provide a more comprehensive picture of student achievement.

Multiple types of assessment also enabled directors to obtain feedback on more than one type of outcomes. Rather than focusing solely on content or cognitively-based outcomes, additional measures of affective and attitudinal variables resulted in a much richer, more complex basis for judgments of program adequacy.

The bottom line is that measures cannot be haphazardly developed or selected. Future faculty initiators of student outcomes assessment should not be told if, when, and what to assess, but most may benefit from technical support in reference to how to assess.

The following summary prescriptions are derived from observations related to procedural and measurement adequacy:

- **Clearly define educational objectives.** The nature of the assessment tools cannot be determined until the desired outcomes are described. Several directors reported that healthy reevaluations of their curriculum and program goals were a necessary precursor to the development of the actual assessment instruments.
- **Assess multiple outcomes.** Directors indicated that since educational objectives were seldom unidimensional, it made little sense to attempt to assess educational criteria with a single measure. The richest data sources enabling the clearest assessment of program outcomes involved combinations of content tests, performance-based demonstrations, attitude assessments, affective measures, etc.
- **Sensitivity to test fairness across constituent groups.** Not all directors dealt with this issue, but those who did provided ample evidence of differential performance across groups. Though some differences may validly reflect the results of disadvantaged preparation for higher education, assessment procedures must minimize performance deficits related to native language differences and/or content that contains cultural/socioeconomic bias.
- **Assessment of measurement reliability/validity.** The importance of this recommendation cannot be over-emphasized. If an assessment instrument does not possess adequate psychometric properties, it provides no basis for meaningful inference concerning the relative performance of the student or the success of the academic program. No amount of faculty and administrative support will save a project based on faulty measurement.

Outcomes

The adequacy of project outcomes to provide direction for future assessment initiatives is a critical dimension of project effectiveness. Evidence for many of the direct and indirect outcomes anticipated in the model for the study proved difficult to evaluate, in part because some of the target outcomes are long-term effects which would require longitudinal measurement. In addition, some data on outcome variables were still being collected as project final reports fell due and so did not get included as outcomes.

The outcome which may be the best index of project effectiveness appeared to be dissemination of results. Project directors' verbal reports consistently underscored the importance of information sharing as an outcome of project development. Moreover, the projects that included explicit plans for dissemination to an identified audience in their early goal definitions continue to be active in publishing results. Further study of these projects and their results should provide additional direction for future strategies in faculty-initiated assessment programs.

References

Ewell, P.J. (1991). "To capture the ineffable: New forms of assessment in higher education." In G. Grant (ed.), *Review of Research in Education,* 17, 75-125. Washington, D.C.: American Education Research Association.

Miles & Huberman (1984). *Qualitative Data Analysis: A Sourcebook of New Methods.* Beverly Hills: Sage.

Chapter 2

Lessons from Pilot Projects

Joanna S. Worthley
Assistant Professor of Psychology
California State University, San Bernardino
and
Matt L. Riggs
Assistant Professor of Psychology
California State University, San Bernardino

The goal of the evaluation study was to identify features across the fifteen CSU pilot projects which were consistently associated with effective assessment programs. Effectiveness was defined by a number of outcomes, including project longevity, impact on curriculum, cooperative attitudes toward assessment, and visibility of assessment efforts. At the outset, we identified 24 **environmental** and **methodological** factors expected to influence project success. Among those, eight variables emerged as the strongest predictors of both direct and indirect outcomes of the project initiatives. In brief, the eight are:

- **Measurement issues:** Measurement adequacy, especially reliability and validity
- **Audience:** Intended breadth of project dissemination
- **Administrative support:** Helpful administrative practice/policy
- **Faculty involvement:** Faculty "ownership"/investment
- **Assessment expertise:** Project directors' assessment background/experience
- **Faculty workload:** Perceived "burden" of assessment activities
- **Multicultural issues:** Attention to diversity in measurement content/strategies
- **Multiple measures:** Extent of "cross-checking" through use of multi-measures in a single domain

Our study showed that in the pilot projects overall, careful attention to these eight factors was consistently associated with several measures of effectiveness. The identification of a "common list" confirms our expectation that characteristics of successful projects could be defined across curricular

boundaries, and provides a baseline prescription for implementing assessment. Thus, for the "big picture" the projects' curricular diversity is an advantage—it allows us to say that specific features of assessment initiatives are critical to outcomes—whatever the discipline.

Nonetheless, the projects' diversity shouldn't be ignored; their differences may be a source of information. The assessments we surveyed differed on several dimensions, including **setting, focus,** and **strategies.** For example, among the project **settings** were a theatre department, a gerontology program, and an economics department; project **focus** included the assessment of an experimental theatre ensemble curriculum, a cross-campus description of curricular goals and content for CSU-based gerontology programs, and the creation of a "local" manual for computer-assisted teaching in economics; and assessment **strategies** involved evaluations conducted by hired consultants, evaluations based on a systemwide survey of gerontology initiatives, and student evaluations of the laboratory component of introductory economics courses. In addition, the projects differed on **level** of assessment—from the level of development of individual measurement strategies, through program-, campus-, and course-level assessments.

As we've noted, differences on these dimensions do not prevent general inferences based on project outcomes overall; in fact, they provide essential guidelines for "big picture" planning, certainly for planning at the campus or system level. However, these same differences are likely to assume more importance for planning individual projects. Given the range of these projects, faculty just embarking on assessment will surely be interested to know whether the specific factors we've identified take on differential importance for different kinds of assessment attempts.

To provide some help with this, we looked for ways to group the projects on the basis of shared features. On most dimensions (i.e., setting, strategy, focus, and level), subgroups can be discerned. For example, on the **strategy** dimension, we can identify projects using similar measurement techniques (e.g., "homegrown" v. "off-the-shelf" measurement), or, within **focus,** locate projects which have assessed comparable programs. In fact, the dimensions above form a continuum of shared features. Beginning with setting, we find little basis for grouping the projects, but as we move across focus, strategy and level, the number of shared features increases. What would seem most useful for assessment planning, then, is to look at the projects by **level,** in part because **level** strongly influences **strategy** and **focus,** but, more importantly, because grouped by level, the projects appear similar in "magnitude" and audience.

Arranged by **level,** the projects form four groups: (1) measurement-level; (2) program-level; (3) campus-level; and (4) course-level evaluations of student outcomes. Within each level we describe projects by **focus** and **strategy,** then present the predictor variables from which inferences for the project **type** may be drawn. What we focus on in this second analysis is the order of the eight predictor variables which showed the strongest association with project

success in our original analysis. The order for each project type was established through a reanalysis of the original data with the projects now grouped by level.

Project Type

Measurement-Level Assessment Strategies

Project	Focus and Strategy
Theatre Program Outcomes	Sought to document both formative and summative achievement through development of a mastery test in theatre knowledge, plus evaluation of theatre ensemble teaching method.
Evaluation of Teacher Credential Candidates' Competency in English Language Arts	Designed an "assessment course" to test subject mastery.
Exit Assessment for Majors in Mathematics/Biology/ Economics	Three "sister" CSUs designed major "comp" exams in collaboration with faculty from relevant departments. The assessment includes an innovative plan to administer exams using faculty from cooperating departments as "outside" consultants.

Rank Order of Variables

- Multiple measures
- Measurement issues
- Faculty involvement
- Multicultural issues
- Audience
- Assessment expertise
- Faculty workload
- Administrative support

The rank order of predictor variables for measurement-level projects is consistent with their focus: Each of these projects developed a comprehensive measure of end-point outcomes where none existed before; not surprisingly, measurement issues appear at the top of the list. Bear in mind that in most cases, each of the eight indicators contributes importantly to project success; what the **relative** importance of these variables can tell us is which types of projects will put particular pressures on specific resources for assessment activities. For example, in the list above, faculty workload and administrative support assume the least importance with respect to the other predictors; this was not the case for the projects overall, where both indicators ranked higher on the list. In the measurement level group, we can speculate

that faculty workload is a somewhat less important indicator because comprehensive exams are typically developed by faculty in their own major departments, where certification of majors is an expected (not added) part of workload. Similarly, administrative support is relatively less important at this level, probably because "comps" are part of the normal departmental agenda and require more measurement expertise than administrative backing.

Program-Level Assessment Strategies

Project	Focus and Strategy
Assessment of Systemwide Gerontology Programs	Measuring effects of curricular variety and program type/organization on student outcomes, particularly for post-graduates.
Portfolio Assessment in Liberal Studies	Design and implementation of a portfolio assessment system for outcomes in an interdisciplinary liberal studies program to make assessment activities an integral part of the curriculum.
Nursing School Program Outcomes	Evaluating competency test outcomes, workplace competency, and the relation of program performance indicators (e.g., GPA) to job performance.
Review of General Education	Program evaluation in an Interdisciplinary General Education program.
Program Review in 5 Areas	Exit and post-grad assessment of majors' performance and student/employer perceptions of program effectiveness in 5 areas: anthropology, economics, psychology, political science and sociology.
Evaluation of a Liberal Studies Program	Multimeasure assessment of student outcomes in a liberal studies program.

Rank Order of Variables

- Faculty involvement
- Administrative support
- Faculty workload
- Audience
- Assessment expertise
- Measurement issues
- Multicultural issues
- Multiple measures

Here, where the reach of the projects is greater, faculty, audience, and administrative features take on greater importance. Most of the projects in

this category include both current students and alumni in their assessments, most are intended to survey large numbers of students, and several are designed to continue data collection well beyond the year term of the pilot projects. Moreover, the directors of these projects expect to share information fairly widely, both with analogous CSU programs and, in some cases, with non-academic constituencies, particularly potential employers for the discipline area. Consistent with this wider scope, administrative backing is more closely tied to project outcomes. No longer "in house," the program-level projects need both the resources and the legitimacy that administrative sponsorship can provide. Project scope is also one likely reason that faculty workload appears high on the list; project directors' self-report indicates that assessments beyond the course or department level are likely to be conducted as "overload" activities. A second reason for the significance of workload in this group of projects may be that faculty with measurement skills are likely to be recruited to the "ambitious" projects, with little guarantee that their discipline-based quantitative skills (e.g., in the physical or behavioral sciences) applied to assessment will count toward promotion and tenure.

Once more, the indicators which are ranked lower in terms of project outcomes for this group aren't unimportant; their link to outcomes is established. What seems to be true for this set of projects is that directors who have undertaken program assessments already have some assessment expertise, so the measurement aspects are perceived to be "under control."

Campus-Level Assessment Strategies

Project	Focus and Strategy
Assessment of Undergraduate Writing Competence	Analysis of student outcomes on upper division (UD) writing exam assesses student performance as a function of enrollment in UD writing course and language proficiency.
Assessment of Undergraduate Reading Competence	Analysis of students' reading strategies/skills related to differential course assignments and library skills/use.

Rank Order of Variables

- Measurement issues
- Assessment expertise
- Multicultural issues
- Faculty workload
- Audience
- Administrative support
- Multiple measures
- Faculty involvement

The pilot projects in the group we defined as campus-level operate in a particular niche: These are projects which assess an aspect of the shared cur-

riculum in the CSU. One example in this category is the assessment of campus strategies for certifying students' writing proficiency at the level required for graduation. The focus of such projects is both sweeping and narrow: Upper division writing competence is required of everyone, so the samples are very large, but the assessment itself looks at a limited but critical sample of student performance. The indicators most closely associated with outcomes for this type of project, measurement issues, assessment expertise, and multicultural issues, make sense: This is an assessment initiative with high visibility, with important implications for large groups of students, and one in which multicultural issues play a crucial role through the relationship of such culturally related aspects of performance as proficiency in first and second languages. Of course, measurement issues (especially reliability and validity) are key factors at every level of assessment; what probably accounts for their primacy in the campus-level initiatives reported here is the level of methodological sophistication demanded by the goals of this project type, which include: (1) the assessment of undergraduate competence in a basic skill area, (2) the statistical comparison of proficiency levels by group-based characteristics such as course experience, primary language, transfer status, etc., and (3) the statistical demonstration of relationships between competency in basic skill areas and performance outcomes on a variety of other academic measures.

Course-Level Assessment Strategies

Project	Focus and Strategy
Western Civilization Course Review	Assessment of pre- and post-course knowledge and attitudes in a Western civilization class.
Student Outcomes in Remedial Writing	Assessment of an Intensive Learning Experience (ILE) writing course sequence.

Rank Order of Variables

- Multiple measures
- Administrative support
- Faculty involvement
- Faculty workload
- Multicultural issues
- Audience
- Measurement issues
- Assessment expertise

Course level assessments in our study have focused on the measurement of post-course changes in attitudes and performance for limited student samples. This project type shares several problems related to short-term assessment, includinxg: (1) time limitations, (2) limited control over non-course influences which may produce changes attributed to in-course experience, and (3) problems with conclusions resulting from the gap between

attitude/perception change and **behavioral** change. It's no surprise then that the indicator most related to project outcomes is multiple measures: One way to increase confidence in pre/post-test results is to collect performance evidence using more than one method. An example of this strategy from this group is the multiple measurement of outcomes in the writing project: Student gains were indexed through portfolio evaluations, a questionnaire aimed at affective and cognitive variables associated with growth in writing, and a statistical comparison of "pass" rates for students in different course sequences. In light of the relatively lower ranking of administrative support in measurement-level projects, its higher ranking here seems anomalous, since several features of these project types overlap. A likely explanation is that the significance of this variable for the two projects in this group is project-specific rather than a general phenomenon. On the other hand, the other top indicators, faculty involvement and faculty workload, could be anticipated for course-level assessment. There are at least three features of course-level assessments that make special demands on faculty. First, by definition these assessments are undertaken by individual faculty, who may be adopting the pre/post-design to replace post-only testing. Pre/post-testing suggests an expanded commitment to assessment; tracking changes in student outcomes over time means increased demands on individual faculty's resources. Second, course-level initiatives differ from grade-driven testing in content areas in that the former develop explicit mechanisms to compare outcomes as a function of course content and approach; that is, they establish the link between assessments and course modifications to improve teaching. Finally, in our study, course level assessments piloted the inclusion of affective as well as cognitive variables in looking at student outcomes. While there is considerable interest in the assessment of affective variables, e.g., commitment to lifelong learning, altruism, and social consciousness (Virginia Commonwealth University, 1988), affective outcomes are difficult to capture, and require specific faculty expertise for adequate measurement.

Summary and Conclusions

Our efforts to define the baseline characteristics of successful assessment projects in the CSU provide two crucial lessons for future assessment initiatives. First, the results of our study suggest that attention to eight factors: measurement, audience, administrative support, faculty involvement, assessment experience, faculty workload, multicultural issues, and multiple measures is critical to effective assessment practice. Moreover, our reanalysis of the key indicators shows that in each project group, the ordering of predictor variables changes, suggesting that a project's level influences the **relative** contribution of these factors to project outcomes. Grouping the assessments allows planners to locate their project type among the levels we've identified, and to use the **order** of the critical variables presented as guidelines in project development. Thus a second lesson for future assessments emerges from the projects' diversity: At least within systems comparable to the CSU, settings (e.g., department or program affiliation) have less impact on assessment

effectiveness than focus (e.g., curriculum or program revision), strategy (e.g., "local" versus "off-the-shelf" measures), or level (e.g., measurement construction or departmental self-study). In particular, establishing the level of proposed projects can help anticipate the differential impact of the factors we've identified as critical to project success. Attention to these baseline characteristics, and to their relative importance for specific project types, should prove helpful in future assessment initiatives in the CSU.

References

Virginia Commonwealth University (1988). *Values.* Unpublished report of the VCU Ad Hoc Value Committee, Office of Vice President for Academic Affairs.

Chapter 3

The Evolution of Student Outcomes Assessment: Politics and Collegiality

Bernard Goldstein
CSU Faculty Trustee and Professor of Biology
San Francisco State University
and
Frank Young
Director
California Academic Partnership Program
California State University

The California State University approach to outcomes assessment illustrates how a large, multicampus system of higher education can deal with the serious public concern of accountability and do so without political mandate.

Collegiality in the CSU

Academic governance in the CSU, as in most institutions of higher learning, is a complex process of shared decision making which translates academic values and goals into university policy. Authority derives from three often conflicting sources. One is vested by law and administrative code in the governing boards and administration. Another obtains from the knowledge of the subject matter and from the pedagogic expertise of the faculty. A third source is the public policy and budgetary authority exercised by the state legislature and the Governor's Office.

The potential for political intrusion into a process generally thought to be central to the academic world is very great, particularly for universities whose budgets are subject to line item review and approval. One strategy institutions can employ to minimize such intrusion is to engage academic leaders in discussion and analysis of emerging policy issues affecting higher education that are likely to attract political attention. Prior consensus of faculty and administrators enables the university, or, as in this case, an entire system of higher education, to respond with a strong, unified voice to external pressures.

The CSU's collegial process is protected by law and by policy adopted by the Trustees—a very unusual situation in any system of higher education. This law, the Higher Education Employer/Employee Relations Act (HEERA), adopted in 1978, establishes the faculty union. It also includes language that protects the academic senates and maintains a collegial decision making process in all areas of vital concern to faculty. Particularly important in this law are the provisions for joint decision making regarding the criteria and standards for promotion, hiring and retention. We know of no other state that has enacted laws protecting both academic senates and unions at the same time.

In 1986, the CSU Board of Trustees adopted a policy on collegiality and responsibility. The policy provides that "the faculty shall have the primary responsibility for the educational mission of the institution." In cases where the views of the administration differ from those of the faculty, the policy requires that a faculty recommendation may be rejected only for "serious and compelling reasons."

This collegial process—slow, critical and carefully considered—has been instrumental in establishing academic policies affecting all 20 CSU campuses and, by virtue of the close linkage between public two- and four-year institutions established in the California Master Plan for Higher Education, with the more than 100 community colleges in the state. The CSU response to legislative interest in student outcomes assessment provides an interesting example of how this complex process works in practice.

Initial Interest in Student Outcomes Assessment

CSU involvement in the student outcomes assessment is in large part a response to pressures from outside the academy. It is important to note, however, that concerns about declining student performance in the 1970s had already led to the adoption of systemwide requirements for assessment of students' writing skills at entry and exit, and their entry level mathematical competencies.

Renewed CSU interest in assessment evolved directly out of the educational reform movement of the 1980s. The "rising tide of mediocrity" documented in *A Nation at Risk* provided a powerful impetus to efforts already under way in California to reverse declines in student achievement. Over one hundred changes were incorporated in the School Reform Act of 1983. Among them was a requirement that all prospective teachers pass the California Basis Educational Skills Test (CBEST). As did many other universities around the country, the CSU raised the admission requirements to the university and imposed new standards for acceptance of students into teacher credential programs.

The surprisingly high initial failure rate on the CBEST reinforced the perception of some critics of California higher education that universities were not, in fact, attending to student learning, and that award of a baccalaureate

degree no longer guaranteed even basic competencies on the part of the recipient. These views were nourished by the appearance of two reports in 1984: *Involvement in Learning* and *Integrity in the College Curriculum.* Both reports were critical of the pressures within higher education that tended to weaken faculty commitment to teaching.

As the largest four-year university system in the nation, and as a system priding itself on the quality of instruction, the CSU that fall hosted the western regional dissemination conference of *Involvement in Learning.* In response to the report and what was said at the conference, the CSU statewide faculty senate undertook a major self-study of undergraduate education in the system. The aim of the study was to evaluate these criticisms and see to what extent they applied to the CSU.

Just as this study was getting under way, in January 1985, student outcomes assessment was introduced into the national debate on undergraduate education. The notion that information on student performance could be used to evaluate the effectiveness of programs, and ultimately of whole institutions, struck a very responsive chord with those within the institution interested in educational reform and with legislators concerned about accountability and cost effectiveness.

In the CSU, the implications of outcomes assessment for the reexamination of the missions of the three public sectors of California higher education then in progress, and for the discussion of how institutional performance might be measured in terms of those missions, were immediately recognized. The chancellor determined that it was in the interest of the CSU to get involved with assessment early, to find out as much as we could about it, and to be in a position to influence the debate when it came to California.

By March of the following year, 1986, the "assessment movement," as it quickly came to be known, reached California. Assemblyman Tom Hayden published a stinging critique of California higher education, and he introduced four bills calling for state mandated assessment. That summer, the National Governors Association published *Time for Results;* a few days thereafter, the Educational Commission of the States released *Transforming the State Role in Education.* Both reports call on state governments to require institutions of higher education to provide evidence of institutional effectiveness based on the outcomes of students' attendance.

The First CSU Assessment Conference

With the support and advice of academic senate leaders, staff in the CSU Chancellor's Office convened a systemwide conference in October 1986 to enable faculty and administrators from all of the campuses to meet with persons from the states where assessment programs had been implemented. Teams of administrators and faculty from the University of Tennessee, Knoxville, Miami-Dade Community College, Northeast Missouri State University, and Trenton State College joined with Alexander Astin, Peter Ewell, Daniel

Resnick, and other leaders in assessment to talk about the assessment programs that had captured national attention in recent months.

From these discussions it became clear that there were significant differences between the situations and institutional needs addressed by assessment models in other states and those in California. For example, one of the indicators of progress in the University of Tennessee, Knoxville's performance funding plan—the accreditation of degree programs by COPA-recognized, specialized accreditation associations—had been standard policy in the CSU for over a decade.

In the months following the conference a great deal of activity related to assessment occurred. The system's instructional improvement program funded pilot projects to enable faculty to experiment with outcomes assessment. The consequence of these activities was the gradual evolution of communities of faculty and administrators on the individual campuses who were interested in assessment. This was reflected in growing attendance by campus teams at national and regional assessment conferences.

The CSU Advisory Committee on Student Outcomes Assessment

In the fall of 1987, to help coordinate all of these efforts and to assist in developing a policy framework, the chancellor established a systemwide advisory committee on student outcomes assessment. Its chair was a former academic senate chair and faculty member. Members included a former president of the CSU Alumni Association, two students from the California State Student Association, three faculty members appointed by the Statewide Academic Senate, a campus president, a campus vice president of academic affairs, and two administrators from the chancellor's staff.

Committee members studied the burgeoning assessment literature, surveyed campus constituencies, and advised CSU representatives working with a task group established by the California Postsecondary Education Commission to respond to one of Assemblyman Hayden's assessment bills. Drawing upon the knowledge base that had been developed through CSU's involvement with assessment, advisory committee members were able to make clear why the models adopted in other states, and in Tennessee in particular, were not necessarily appropriate for California. CSU representatives argued against the imposition of state-mandated testing programs for university students in favor of the incorporation of student outcomes assessment in institutional accreditation.

Reports on advisory committee activities were regularly given to the Academic Affairs Committee of the Academic Senate. Draft statements developed by the advisory committee were routinely reviewed and critiqued by the senate committee. Advisory committee members were informed, at each important step in their work, of the Academic Senate's reactions and suggestions.

The Guiding Principles

The following set of principles was derived from responses to a survey of campus views about assessment. They served to guide the committee in advising the chancellor and in formulating assessment policy. The principles address three key questions:

For What Purposes Should Assessment Be Done?

- The only legitimate purpose of assessing student outcomes is to improve teaching, learning, and academic advising at the individual, course, program, and institutional levels. Data from outcomes assessment programs will not be used for cross-campus rankings or comparisons of individual faculty.
- Unique assessment models, tied to a multiplicity of goals represented by the different institutions and incorporating the principles adopted by the committee, are appropriate to the CSU.

Who Should Assess?

- Programs to assess student outcomes should be campus-based, faculty-centered, and student-responsive.
- Faculty of the individual campuses have the primary responsibility for deciding how to assess student learning. This extends to the design or selection and administration of assessment methods, the interpretation of the results, and how the data will be used to improve programs.
- Consistent with the principle of institutional responsibility, the resources appropriated for assessment should support the development and operation of programs at the campus level. System and state efforts should be directed to helping campuses devise assessment programs. For this reason, the CSU opposes creation of a centrally administered state assessment program.
- Data collected through institutional assessment programs should be governed by recognized codes of ethics treating research with human subjects.

At What Cost Should Assessment Be Done?

- Student outcomes assessment, when appropriately carried out, is just one of several institutional practices that must exist in order to achieve educational excellence. Student outcomes assessment should be linked with the academic program review process presently mandated by the CSU Board of Trustees.
- While the evaluation of student learning is a regular faculty responsibility, implementation of comprehensive assessment programs will add significantly to faculty workload and will be costly. These costs must not be borne directly by students. Supplemental funding is essential to the

development and operation of effective assessment programs. In the absence of adequate support, program implementation must be limited.

- Before substantial resources are requested for, or invested in, comprehensive assessment programs, it must be established that they provide effective means to improve the quality of educational programs. Because of their high cost and the need to evaluate their effectiveness, assessment programs should be implemented experimentally and incrementally within the CSU. It should be noted that assessment has the potential of identifying problems in educational programs that require additional resources for solution.

After lengthy study and discussion about the benefits and potential disadvantages of student outcomes assessment, the advisory committee concluded that: (1) student outcomes assessment programs, if carefully designed and implemented, have significant potential for improving teaching and learning; (2) the CSU cannot afford to ignore educational practices and strategies, including student outcomes assessment, that show great promise for the improvement of teaching and learning; and (3) it is the responsibility of CSU faculty and administrators to adopt practices and strategies demonstrated to be effective, including assessment, as appropriate to the classroom, the discipline, and the campus.

From the study of assessment programs in other states, members of the advisory committee had learned that it is not compliance with a requirement but responsiveness to what is learned that can make student outcomes assessment a powerful and useful tool to improve academic programs and institutions. If it is to become part of the institutional culture—the "culture of evidence" encouraged by the Western Regional Accrediting Association—assessment must measure what is educationally significant to the faculty in intellectually credible ways.

The Second CSU Assessment Conference

The committee had a choice. It could have reported its findings, conclusions, and recommendations for systemwide policy to the chancellor and considered its work finished. It chose instead to carry the collegial approach to the level of the campus. With the chancellor's support the advisory committee helped plan a second systemwide conference on assessment for the fall of 1988.

The 1986 conference had served well the purpose of informing key CSU constituencies about outcomes assessment and stimulating their further inquiry. The goal of the second conference, *From Projects to Policy: Seize the Agenda,* was to share what campuses had learned from their investigation and experimentation, and to discover whether there was sufficient common ground to develop a system policy on assessment.

The fall 1988 conference was the definitive event in enabling the committee to move from consideration of issues to the formulation of policy recommendations. A number of faculty and students spoke positively about their own experiences with outcomes assessment. This was most enlightening for the faculty, since skepticism still runs high regarding the merits of investing time and energy in assessment that might be devoted to other, more direct efforts to improve teaching and learning. Every controversial issue was put on the table for discussion: the impact of assessment on workload, the increased costs for assessment, issues of validity and need, and real and implied threats to academic freedom.

Open discussion of these issues was very useful in enabling the advisory committee to move forward. From the debate it became clear that a substantial number of faculty wanted to know what the impact of their efforts on students was, and how they could determine whether students were meeting the broader educational goals associated with the degree program. To answer these questions, evidence would have to be gathered from several sources. Getting such information had to be the shared responsibility of faculty, both as individuals and as members of departments, the administration at all levels, and of students. It was recognized that existing mechanisms were in place that could be used, and that the process should be incorporated into program review, which was already well established through Trustee policy.

Recommendations for Systemwide Policy on Assessment

The advisory committee formulated three sets of recommendations to four separate constituencies. The first set of recommendations relates to the improvement of teaching and learning.

> At the **level of the individual:** Evaluations of students in specific courses should include elements that assess their progress in terms of the overall programmatic goals of the major, the minor, and the campus general education program.
>
> At the **department level:** The department faculty should have ways of collectively evaluating individual student attainment in the major that go beyond course grades.
>
> At the **campus level:** Each campus should have mechanisms to assess how well students are meeting goals of the general education program, and for examining the interaction between academic programs, student support services, the campus environment.
>
> At the **system level:** The central system administration should seek to restrict the proliferation of standardized testing in favor of supporting integrative assessment that meets programmatic and curricular goals.

One of the main reasons for opposing proliferation of standardized testing was recognition that reliance on readily available instruments would circumvent the critically important work of clarifying goals and determining appropriate means to measure their achievement.

A second set of recommendations focuses on improving the quality of assessment itself.

At the **level of the individual:** Each CSU faculty member should review his or her current evaluation practices in the light of recent findings on testing and on assessment.

At the **department level:** As a means to develop resident expertise in assessment, each department should encourage some members of its faculty to engage in research and scholarship related to assessment in the discipline.

At the **campus level:** The administration should assist departments in acquiring resources for staff development in the area of assessment.

At the **system level:** The central system administration should assist campuses by supporting dissemination, experimentation, and professional development in assessment.

The third set of recommendations aims at improving communication with and about students.

At the **individual level:** Students are entitled to information that will enable them to understand what the results of evaluations of their performance mean in terms of the broader goals of the program and of the institution.

At the **department level:** Departments should ensure that faculty evaluation of students' work includes elements that link individual students' achievements to overarching programmatic goals. They should also ensure that this information is used appropriately for advisement of students and for analyzing the effectiveness of departmental curricula.

At the **campus level:** Campuses should assist departments by systematically collecting and providing information on current and former students' characteristics, development and opinions.

At the **system level:** The central system administration should help campuses to report information gathered by the system and the campus in ways that are useful to departments in analyzing the effectiveness of their programs.

In March 1989, drafts of the *Guiding Principles and Recommendations* were sent to the campuses. These drafts served as the basis for discussions at regional meetings. These sessions were well attended by interested faculty and administrators and were extremely useful to the committee in developing its report. The 1988 conference had been a turning point for many campus people as well as for the advisory committee. Several drafts and faculty reviews later, a resolution urging the Board of Trustees to endorse the recommendations was unanimously approved by the Academic Senate in March 1990.

In September 1990, the CSU Board of Trustees accepted the advisory committee's report, adopted the *Guiding Principles,* directed the chancellor to implement recommendations pertaining to the central system administration, encouraged the campuses to pursue the goals of the recommendations, and requested that periodic reports on assessment be given by the presidents.

Politics and Collegiality

Donald Langdenburg, chairperson of the board of the directors of the American Association for the Advancement of Science, says that discussion about higher education is beset with slogans for complex and controversial public issues that represent significant bodies of scientific knowledge. Examples of such slogans include: accountability, educational reform, the environment, political correctness, global change, technology transfer, and assessment, to name just a few.

According to Langdenburg, slogans reduce controversial and complex public problems to labels. Use of slogans suppresses reference to the underlying body of research that should be an essential feature of the debate, but often is not, and thus permits debate to be shaped by the politicians and opinion makers, rather than the educators or scientists. When the time comes to move from debate to action, action is shaped by the values and perceptions of the leaders of the debate. There is no assurance that these values or perceptions are based on relevant data.

What seems to be missing in many debates about higher education, says Langdenburg, is engagement between those who understand and those who act; i.e., between the faculty, who are trained to value complexity and to respect evidence, on the one hand, and the public policy makers, who must often make politically defensible decisions quickly, on the other. Faculty must do their "civic duty" and become more involved with public debate about controversial issues. Not only must they do the appropriate research, they must also convey the results of that research to those who have the power to act.

If it is to yield valid information, assessment of the outcomes of a college education must be undertaken with critical awareness of the characteristics of the particular students being assessed and of the effects of human biological, cognitive and social development. In the absence of such understanding, changes in student behavior and knowledge may be attributed to the educational programs or institutions that actually have little to do with these students' experiences on campus. Outcomes assessment carefully and correctly done can be a powerful tool for improving teaching and learning and for providing the kind of evidence of institutional effectiveness policy makers should have. Assessment programs designed around outcomes that can be measured easily and with statistical reliability, but which do not reflect the specific influences of the program or institution, may satisfy demands for accountability, but have little potential to help institutions understand and improve themselves.

In many states, outcomes assessment of questionable validity has been imposed on higher education by state government. In California, this has not occurred. The early participation in the public policy debate of CSU faculty and administrators who were knowledgeable about assessment and who could point to steps that had already been taken to involve campuses in the assessment debate was a major factor contributing to legislative restraint. This restraint, in turn, permitted the collegial development of a systemwide assessment policy which has the strong support of the campus and systemwide academic senates, the administration and governing board.

Chapter 4

Portfolios in the Major: A "Success" Story?

Lu Mattson
Professor, Hutchins School of Liberal Studies
Sonoma State University

Introduction

Time alone will determine the effects of the outcomes assessment efforts of the late 1980s. Whatever the results for higher education in the state of California, they are presently having their impact in the Hutchins School of Liberal Studies at the Sonoma State University where several kinds of assessment have taken place within the past half dozen years. Assessment of attitude shifts occurring as students completed Hutchins' lower and upper divisions enabled the faculty to understand the development of the program's undergraduates. Surveys of majors five to ten years after graduation and focus groups which interpreted the survey results helped identify aspects of the Hutchins program which were serving students well and others which required improvement. These first assessment projects were completed within the year in which they were begun, were reported to the faculty, and resulted in several curricular and staffing decisions. Those efforts set out to answer specific questions about students and the program, and having secured the answers they sought, they were considered finished. Another attempt at assessment, now in its third year, is still evolving, however, and will probably continue to do so for some time to come. This assessment effort involves the systematic use of portfolios in the Hutchins School upper-division liberal studies major. Unlike Hutchins' first approaches to assessment, this one has grown beyond the protocol developed for it through a CSU systemwide Academic Program Improvement Grant. What we are presently learning is that strategies that were easy enough to describe before we really had begun sometimes worked, sometimes did not; goals that seemed easy to attain have not proven to be so in all instances. While we are in the third year of incorporating the portfolio in our upper-division program and are beginning to see the results, it is safe to say that we have only begun to explore the

problems and possibilities attendant upon the use of such a tool for students' purposes and our own.

A few words about the special requirements we have decided to make of any assessment tool we employ and about the Hutchins program itself are in order to explain assessment at Hutchins. With regard to assessment itself and our view of it, we need to make clear that our interdisciplinary program, like many spawned in the '60s, started out with a claim of being "student-centered." Twenty-five (and more) years after the establishment of the Hutchins School, we still do assert that our primary aim is to produce life-long, autonomous learners who feel capable of moving around among various disciplines, who are confident of their ability to analyze and synthesize ideas. Given that fundamental goal, our courses are interdisciplinary, and our pedagogy does as much as we have discovered to do to encourage student participation which is proactive rather than re-active. Because we are interdisciplinary, we do not owe particular allegiance to the content of any single discipline, and try instead to provoke curiosity about and competent exploration of ideas from a variety of fields. That said, it can be imagined that outcomes assessment which uses standardized approaches and is based on disciplinary content alone offers little to Hutchins in its efforts to improve its program. As we began our assessment attempts, we discovered that we needed a set of goals to help us select among the many avenues of outcomes assessment. We derived these even as we were arguing about whether we should apply for a grant to help us introduce portfolios into the major. There is nothing exceptional about the goals we articulated, but as the discussion of our use of the portfolio unfolds here, mention of them should help clarify the bases for some of our decisions about how to improve the portfolio that we have developed. As a faculty we agreed that the following goals, in roughly the order listed, should inform our portfolio work. The portfolios were to be developed so that they could:

1. lead the student to a better understanding of the nature of interdisciplinary enquiry in Hutchins;
2. help enhance student self-awareness;
3. improve faculty advising;
4. draw the educational experiences of the major together into a coherent whole;
5. provide insight into the academic program's effectiveness; and
6. make the achievements of the program understandable by others.

Predictably, the degree to which the goals have been met is, so far, varied. At the outset, our faculty agreed to stick with the decision to implement a portfolio until all members had had a chance to work with it in and out of class. Only after the faculty had attained wide experience with it would it be thoroughly evaluated to see if its use did indeed achieve those goals, and then it would be modified as necessary. That agreement proved to be a good one: Faculty who were skeptical about the attempt joined in willingly when it was agreed that the portfolio's continuation in our program would eventually

depend upon consensus regarding its usefulness. Some of the concerns that will finally have to be addressed will appear at the end of this chapter; suffice it for now to say that at present most of the faculty have had a chance to use the portfolio in classes. The number of students who have completed the portfolio is just now becoming substantial. We are, in other words, perhaps within a year of achieving the level of expertise regarding the portfolio that will allow us to make a decision about the "final" function it will have within our major. It is apparent already, however, that addition of the portfolio to the upper division has indeed helped cure some problems, even as it has introduced new ones. Perhaps others about to launch portfolios of their own will find our experience helpful. But before we can detail what the portfolio presently means to our program, a few words about the program itself and the problems the portfolio was intended to cure are necessary.

Hutchins' Major Program

Inaugurated in the days before the term liberal studies had been narrowed to a designation for "teacher preparation," Hutchins has always had at its heart the type of course which takes as its focus an issue, a question, a problem, or an idea that requires the approaches of many disciplines. The student's experience begins with a year-long course which radically reexamines American "truths," then widens out to explore the immigrant experience and several third world "countries of origin." The purpose of the course is to get the student into a stance which is more analytical regarding that which is familiar and more comprehending regarding that which is foreign. Beyond this junior-level course, most of our offerings fit into one of four broad interdisciplinary areas: *society and self,* which, as the title implies, emphasizes issues regarding the person in social roles; *the individual and the material world,* which includes courses raising questions about science, technology and values; *arts and human experience,* which looks at ways in which the nondiscursive allows us to understand; and *consciousness and reality*, which deals with perception, cognition, and reality formation. Roughly, but only roughly, the first three areas correspond to the social sciences, the natural sciences, and the humanities. If one were to analyze what kinds of units are included in any of those courses, one would discover that the units for any single course come from several disciplines. A typical course taken in the second core area, for example, might have a unit of science because the course incorporates cosmology, a unit of social science because of the exploration of social history during the sixteenth century, and one of humanities because of an analysis of the philosophical systems which affected and were affected by the Copernican Revolution. Students must complete work in each of the four Core Areas, emphasizing one of them rather more substantially than the others.

A different type of course stands at the beginning and at the end of the student's career with us. At the beginning, *Libs 302: Introduction to Liberal Studies* is intended to acquaint the student with Hutchins' approach to learning and its intellectual foci. *Libs 402: Senior Synthesis* allows the student to draw the undergraduate education to a close through a capstone experience

at the end. These latter two courses are of particular concern as we describe how the portfolio works for us.

The content of Hutchins courses is somewhat different, then, from what a typical junior entering the major will probably have encountered in a traditional general education program. Differences mark the pedagogy as well. Working mainly with primary sources and virtually never with textbooks, we expect in all instances that the student will learn to read punctually and carefully and to think independently about the materials and experiences that make up course content. Punctual, thoughtful reading is crucial since nearly all courses in the major are seminars—colloquia, really—where students discuss what they have read, dig out meanings, and find applications of ideas under the guidance of a seminar leader. Requisite skills include the ability to read sophisticated material critically on one's own and make interpretations of it, to draw the attention of one's colleagues to major ideas and say why they are important in a discussion, to listen attentively to others' points of view and respond to them substantively. Writing about one's ideas is required in all courses, both in "preparation papers" before discussions and in formal essays. These writings have little to do with reports, so typical of what most students have been taught to produce, which simply summarize the ideas of others. Writing in Hutchins has mainly to do with the student's struggle to articulate his or her own insights. Finally, beyond the discursive, students are encouraged to find other ways to learn and to present themselves: through art projects, community work and internships. Where pedagogy is concerned, then, the program emphasizes modes which are usually strange to the student who comes from a traditional educational background.

We began to develop the portfolio for use in Hutchins to respond to some of the problems inherent in what has been described. Students who enter the major usually come from other institutions, primarily community colleges. For many, confronting our set of expectations is a disorienting experience. Not only are the categories of information that served to get the student through the general education labyrinth superseded in our upper division, so are many of the students' ideas about intellectual authority and their relationship to it. Before we recognized the need for the introductory course, students generally rode out this period of confusion and eventually learned to assume responsibility for their own knowledge, but the first semester was usually finished before that transformation occurred. The student did not really see what the range of Core Areas implied, even though catalog copy explained it. Furthermore, the student often heard only peripherally that other avenues of exploration besides the formal essay were open and considered appropriate. What constituted a fruitful question around which to formulate a paper or project remained a mystery until the student was perhaps as much as half-finished with the major. Certainly, the notion that there is a web of relationships tying together the ideas of all the courses in a Core Area (and others besides) is news to people who have been taught to divide concepts in ever smaller pieces in the interest of precision. Getting this and similar information about the dimension and possibilities of the major to the student more clearly

and quickly was important to us. Given our peculiar set of courses and intentions regarding student skills, an introduction was scarcely enough. We viewed the portfolio, incorporated in Libs 302, as a potential means of clearing up the opacity of our major program.

There was, as well, the issue of the capstone experience. Hutchins has always talked in terms of one—to allow the finishing senior students to synthesize the disparate parts of their educations, or to allow them to look toward the future and the work they will do after graduation. Indeed, many students have even accomplished senior projects which did one of those things very well. But over the years the senior theses and senior projects have been hit-or-miss affairs, well-focused if the student knew how to take charge of the project, less so if the student was not assertive enough in developing the project. We thought that the portfolio, used as a means of conferring coherence on what might otherwise simply remain a collection of classes and projects, might provide a platform on which the student could build a senior project.

Advising had been a problem for us as well. Because Hutchins was approved as a waiver program, relieving the student of the need to take the National Teachers' Examination, we had become a major that served substantial numbers of individuals seeking the elementary teaching credential. These extremely goal-directed people had to complete a long list of credential requirements; accruing x units in this area, y units in another. The consequence of this situation was to turn advising into a bean-counting exercise, where actuarial practice substituted for talk about what a student was learning and how that person was developing. The portfolio was seen as a means of keeping an advisor abreast of the student's development, since it would be used at the end of each semester as the student was being advised.

Finally, an overview of collective student achievement had always eluded the faculty. We thought we knew what we were accomplishing with the groups, but we had never really "measured" it. At Hutchins, the faculty know the students unusually well, owing partly to relatively small seminar-discussions where the students do express their own ideas. In addition, the faculty has worked unusually closely with one another precisely because there is no single, agreed-upon canon of material that must be included in the interdisciplinary curriculum. Co-taught, co-designed courses and an atmosphere of dialogue about what students must learn has kept us more knowledgeable about one another's thoughts regarding the curriculum than is necessary in a more traditional setting. However, in a program such as ours, precisely because of its latitude, one remains unsure whether the range of student intellectual growth is what it should be. For all our evidence that most students were, as individuals, being induced to become the autonomous, imaginative, interdisciplinary, critical readers and thinkers we intended to foster, we couldn't say what was true for the group. The portfolio was seen as a means of attaining a wide perspective on the effectiveness of the program, one founded on data we could all see. Once having made these faculty observations regarding the program's overall effects, we

believed we would know better where we were having our best success and could make corrections where necessary.

These, then, were the promises the portfolio seemed to hold out to us as we began its implementation.

The Portfolio Protocol: Theory and Practice

In 1990, after a year spent designing the book and its contents, determining what to ask of the students and piloting and refining our first efforts, the portfolio became a feature of the Hutchins upper-division program. A year later, a Hutchins report explained the portfolio protocol, describing how the portfolio worked in the context of the introductory course, the student's middle semesters in the major, and the capstone course, and discussed how it was used to provide program evaluation. That procedure is, to a large extent, the one followed today. The report documented:

> In Libs 302, the student buys the portfolio, a three-ring notebook, handsomely embossed, we think, with several labelled dividers and a good many instructions and assignments inside. Four of the sections correspond to the four Core Areas. When the student opens a Core Area section, one page describes what kinds of issues and ideas make up the Core Area, another sets forth a series of questions or assignments, each of which is an example of the kind of enquiry the student might engage in during an appropriate course. These questions are mainly meant to be illustrative, but in Libs 302 they form the focus for in-class discussions which explore the kinds of materials and considerations a student would need to consider in drawing together a work to include in the portfolio. So, for example, the student reading about the first Core Area would find the following:
>
> "The first of the Hutchins Core Areas takes as its focus the relationship between the individual and all kinds of human groups. It is, after all, in the context of human interaction that the individual finds the dimensions of the self. Between society and the individual flow those ideas, attitudes, beliefs which result in the political and economic arrangements that make life-in-common possible. . . ." Included among the questions one finds: ". . .5. investigate a specific public policy issue or political controversy. Identify the major actors. What interests or values are under dispute? How can this issue be resolved? How does this case shed light on concepts such as social justice, human rights, authority, community, freedom, and so forth?" Similar descriptions and questions introduce the Libs 302 student to the other Core Areas, and provide the springboard for classroom discussion of the area as well as some preliminary writing to be shared in class. By the time the classes have worked through the four sections, the student has had a chance to ask the questions that make manifest what our interdisciplinary major is all about.

Because we want the students to reflect on why they are in college and in our particular major, we ask them to spend some time in Libs 302 thinking not just about the coursework they will need to complete but about their goals beyond graduation and the skills they will need to develop if they are to become the active learners we profess to develop. The report described the part of the protocol related to these items as well:

> Too often, students come to our program prepared to meet our demands, complete our courses, fulfill our graduation requirements without consulting themselves adequately about their own needs and expectations. A portion of the portfolio therefore has as its aim the opening up of that personal dialogue whereby the student begins to learn to identify his or her own strengths, weaknesses, aspirations, fears, potentialities, and see them within the context of their work in the major. To that end, at the very beginning of Libs 302 the student is asked to answer a series of questions that begin: "Why are you in college? What do you hope to be doing 5 years from how?" The student goes on to fill out and discuss with the instructor the Kolb Learning Style Inventory. The student includes in the portfolio the "Evaluation of Transfer Credit" form from the Admissions and Records Office and receives a Requirements Checklist indicating just what will be included in the major. The student then lists all previous coursework on an advising chart, confers with the advisor about required work remaining to be done either in general education or for the major. Next he or she works through a brief section entitled "Developing Skills." Here we have an opportunity to let the student know the kinds of abilities we expect to see develop over the semesters: ability to enter discussion effectively, analyze written material, communicate through writing, make formal oral presentations, communicate through the arts. We are able to ask the student about "respect for other points of view, willingness to discuss/debate issues, relevance of comments, appropriate quantity of talk" and so forth, thereby making it clear that these items and others like them, both large and small, go together, we believe, as part of the educated person. At the end of this section of the portfolio, the student is asked to "write a brief synopsis...of your own most important learning goals and the ideas about learning you take with you as you go...." This Synopsis of Educational Goals is not about requirements to be met but about that other layer of development that can't be captured by GPA's.

In sum, we ask the student to have an important individual board meeting just as the major begins. This provides us with a starting point, with "formative evaluation," if you will, that lets the student know not only the technical dimensions of the major but the scope of internal development we hope for in each of them and the latitude we think should be available to them to realize their aims while they meet our requirements.

Our primary goal of first serving student needs determined two aspects of the portfolio: one of which sought to introduce the student to the interdisciplinary major and its pedagogy, the other of which attempted to get the student to stop and reflect about personal educational goals just at that point at which general education had been completed and work in the major was about to begin. Our further intention of improving our advising through the agency of the portfolio was reflected in plans for its use in the middle semesters of a student's upper-division career:

> When the student has completed Libs 302, the ongoing development of the portfolio is the student's responsibility. But it furnishes a focus previously lacking in our student advising as well. It accompanies the student to an appointment with an advisor each term. Discussion of what work has been completed, how the development of the various Core Area sections is progressing, whether the student has changed aims or has met them proceed naturally from the introductory work the student has done at the outset.

Each of us sees about twenty advisees, often only when it is time to decide on classes for the coming semester or when it is time to apply for graduation, unless the student gets into some kind of academic difficulty. The faculty had long complained that advising for the credential, which really did consist of looking to be sure that the student had met a list of requirements and had earned an appropriate number of units in various fields, had turned us away from the primary student/faculty relationship, the planning of an education. The portfolio, as it was being developed, would provide us with a growing body of material, all of which was generated by the students themselves:

> Each student is required to collect at least three items for each Core Area before graduation. For one of the areas, the student must present half a dozen examples of his or her work. Certainly, most of those items will be essays written for seminars. The students are urged, however, to include other kinds of materials as well—tapes, art projects, records of activities, annotated bibliographies, and so forth; but each section must include at least one substantial formal paper. Furthermore, at least one of the formal essays must address a multi-cultural issue. Students are instructed to include materials as they were returned to them by their instructors. In other words, we do not want to see re-written, re-worked pieces, for the important reason that we expect to be able to trace student growth and development by reviewing portfolio contents. In this respect, the Hutchins portfolios differ from those prepared by students in the arts, for example, who use the portfolio to draw together the most finished version of a work possible.

We also sought, by placing heavy emphasis on the portfolio in the final senior semester, to bring the student to a sense of closure with his undergraduate education:

A capstone course, new to the program, has been designed for the graduating seniors. While *Liberal Studies 402: Senior Project* gives the student a chance to work with colleagues in making presentations regarding their "major" Core Area, it is above all a place to bring the portfolio to completion. This last effort aims at conferring coherence on what may otherwise simply remain a collection of classes and projects. By way of completing the portfolio, the student must write a synopsis for each Core Area, a job that may have been completed earlier in the program. The primary writing assignment for Libs 402, however, asks the student to develop a "Senior Synthesis," a substantial paper bridging the areas represented by work the student has included in the Portfolio. More than simply a review of what the student has done, it also discusses where the student stands intellectually at present, describes how the individual got to that point, and speculates about where he or she thinks the future will take him as a "lifelong learner." The student is urged to use this final paper as a chance for reflection and in it to produce an instance of his best work. The Libs 402 instructor, then, engages in summative assessment to close for the student the cycle begun with the formative assessment of Libs 302.

Once the student's work with the portfolio is completed, the document is used to cast light on program effectiveness:

> Program assessment can begin when the students enrolled in Libs 402 turn in their portfolios at the end of the term. At that point the instructors for Libs 402 meet together to "read" randomly selected portfolios and to judge them according to a set of criteria devised by the faculty. At this reading, the question is not whether a given student has or has not successfully completed the major, for the Libs 402 instructor has already made that determination. Now the question is whether the program seems to be achieving its aims: have the students developed the skills proclaimed as necessary at the outset; do the students have the ability to analyze and synthesize ideas; is the range of projects evidence that the program is rigorous enough, imaginative enough, challenging enough? What seem to be the areas of greatest program strength? What look to be its consistent points of weakness? Holistic reading techniques are employed to enable the faculty to look over a substantial amount of material without getting lost in detail. Perry's scheme of intellectual and ethical development is applied to determine whether the students as a group seem to have progressed adequately in the direction of responsible, resourceful autonomy.

Anyone acquainted with the distance between a grant proposal's description of an as-yet-untried procedure and the actual practice which eventually emerges will probably recognize a fair amount of a *priori*-ism in

those descriptions, and what has been called a "triumphalist tone" in the delivery. In fact, we were on the right track in many of our original ideas, and we do use the portfolio substantially as described above. But we are finding that many of our ideas, which sounded good as our faculty committees were planning the portfolio, are not what we need in reality. Our own critique of those first attempts has led to a number of changes, and more are yet to come. The critique itself is helpful: Trying to discover how to strengthen the portfolio has led us to be clearer among ourselves about certain aspects of the program.

In the belief that a discussion of what has not worked may save others from unnecessary difficulty, and is therefore more important than explaining what has worked, we offer a number of our own criticisms of what we designed. A critical reader has probably already guessed some of the difficulties that our use of the portfolio, for all its possibilities, has introduced. Some of them are relatively insignificant, but the most important probably has to do with simplicity, or the lack thereof. We were unaware how perplexing the portfolio could be for faculty and students alike. Nor did we calculate how formidable portfolios can seem in courses that use them extensively. Both problems—lack of simplicity and formidability—show up in the introductory course. The portfolio does indeed include important information about the interdisciplinary major, about graduation requirements, precredential requirements. It does get the student to respond to serious questions about goals and to analyze his or her own level of achievement in previous courses. After working on the portfolio, the students seem to understand where they are headed more clearly than was previously the case. However, that first course has been seriously affected by the needs of the portfolio itself. Asking the students to read and write about each of the four Core Areas can consume too much of a semester. Admittedly, the students do have a clearer grasp of what can be included in each area, but that understanding is gained at a price. In fact, the students need that course in order to learn to seminar and to write. That first aim is achieved under present circumstances, but the problem is that course materials are now selected so that each Core Area is explored by a few brief works or one longer one. The writing, too, is directed to each of the Core Areas. The papers are adequate, but not better. The problem for us is that we overdesigned the use of the portfolio in the introductory course, thereby sacrificing course focus. Unlike our other courses that do take a concentrated look at a single issue or idea, Libs 302 now takes four looks, and the result is superficiality. The students do not learn in this course how to dig deeply into a topic, and that affects the quality both of their writing and of their verbal discourse.

The portfolio is not presently used during advising in the middle semesters of the student's stay with us. In fact, the student's ongoing questions are not about portfolio development at all, nor do the advisors pursue the issue. In this instance, we have probably underdesigned rather than overdesigned our protocol. If we do intend to switch to advising about the student's intellectual growth, then we need to provide time in the semester's schedule for it to occur, and we need to work together to explore what we mean by the growth we are looking for in each student as represented through the collec-

tion of products. The idea of ongoing advising based on portfolio materials is not a bad one; it is simply an undeveloped one.

Our assumptions about the final course also need reexamination. We realize that we sent two messages to the students regarding the Senior Synthesis paper, and each had a weakness inherent in it. One of our instructions was for the student to return to the collected materials, synopsize the work present for each Core Area, and write a major essay "bridging" two or more of them. That is exactly the kind of thing that sounds wonderful during planning stages. The reality of the situation is that no one, neither faculty nor student, knows what a "bridging essay" is. If we did know, we would probably recognize that writing one is far too formidable a task for an undergraduate. Besides, who wants to put out a major effort to revisit last semester's work?

The second message we sent the students was that they could write a Senior Synthesis which looked at where they had been intellectually, where the program had taken them, and where they expected to go in the future. It is not hard to imagine which of the two essays the students choose to write. It is also not difficult to imagine how self-serving and sycophantic those essays can become. When pushed, many of the students will cop to it that they select the latter option because it is easy, and because it is political to praise us. That, needless to say, is not what we intended. We could, however, have guessed what we would get. The first batch of portfolios presented the latter type of essay exclusively, and it was nice to read them. So did the second batch, but by then we began to understand what we had invited. Obviously, another piece of work remains for us to get straight among ourselves what we should be asking for. The exercise so far has been illustrative.

Use of the portfolio for program evaluation, too, presents some problems. In fact, sitting down with colleagues over several students' collected works is quite an interesting exercise. At present, the procedure looks to be more promising than we at first assumed it was. Our intention to use it to see how a student had (or had not) developed under our tutelage led us to instruct the students to include their papers in the condition in which they received them from their instructors. We wanted to see what kind of guidance the student was receiving. We had not calculated that the students are reluctant to put anything less than their best foot forward. They want to edit their work or to exclude any of it that got less than raves from the instructor. And yet, sanitized papers will not tell us as much as we want to know. This use of the portfolio, which goes down a different path than that used by art majors who collect only their best work, needs to be made more effective. We recognize that having the students place materials in four categories with the Core Areas defeats our ability to see easily how a student has developed through time, and so a proposal is before us simply to have them arrange their work chronologically. In our attempt at elegance, we missed the obvious.

So far, this report says little about the achievements of the portfolio. That is probably as well, for we are truly in a middle state right now in which we are beginning to understand what is, for us at any rate, a new medium.

Even if we did not recognize that the portfolio would seem formidable to some students, we were right in believing that they, as well as we, would find satisfaction in signs of their growth, and that should be mentioned. The formidability is itself interesting in that we didn't anticipate it: We assumed that the student would see the portfolio as an opportunity to create something uniquely one's own, and, for most of them, that is proving to be the case. But there is always the overlooked angle. The student who doesn't see the portfolio as a creative challenge is likely to see it as a frightening task. At the outset the book is, after all, empty except for material we provide describing the program, and they are to fill it. It turns out that we need to remember that not all juniors will be thrilled at the prospect, even while we know that they will probably be proud of the outcome.

We know we are still in search of the elusive capstone project or experience, but now we are clearer than before that it needs to have an inner necessity of its own to be truly engaging of the students' abilities. It also needs to push the student to some limit not yet achieved in the regular coursework demands. The culminating senior synthesis paper will achieve that, when we have experimented a bit more with it. The faculty working with this semester's group of 402 students have pressed the issue. While the students met these increased demands with complaints at the outset, they now see the importance of their doing an extended piece of writing on a matter of substance in their last semester. Perhaps we are finally learning to be precise about what constitutes a last best effort from the student.

Our criticisms about the portfolio's commanding too much time in both the introductory and the closing courses result from our not having anticipated that we would overdo our emphasis on it. In both instances, the remedy is fairly simple: We need to allow the portfolio itself to move further into the background by reconceptualizing the intellectual issues we want to approach in the introductory course and bringing them to the fore. We need, as was mentioned before, to let the students find their own ways to significant issues or problems around which to develop the Senior Synthesis, which still will be the centerpiece of each portfolio. These errors are remediable. On the positive side, the portfolio has forced us to rethink the opening and closing courses in the major: How do we get the students up and running in our program? How do we help them achieve a sense of completion?

We said that we thought we would be able to draw some conclusions about program quality on the basis of what we received in the portfolios, and we can. For instance, we claim that the students learn to write with us, and many of them do. But in the portfolios we see less writing than we thought we would, and we do not see the range of challenge that we think should be there. We will be able, we now are confident, to be quite specific in recommending program changes with regard to writing assignments. A kind of sameness in the types of assignments we presently make is being revealed to us, and we will want to think about that. We have already noticed a tendency on the parts of the students to skew what could (or should?) be projects in the natural sci-

ences toward presentations focusing on the social sciences. In addition, we can see that it is not the natural sciences that are the students' greatest weakness; it is the arts. For the first time, we are getting a look at what, on balance, the student's upper-division experience is. Our discussions will force us to program considerations we had not noticed were waiting for us.

In conclusion, one would have to say that the portfolio at Hutchins is a mixed success. It is alive, and we are wrestling with it. We had to start someplace, and so we say that it is all right that we planned too much in part, too little in part. Assuming that the portfolio remains a feature of our program, the next few months should bring us to much greater clarity about what is possible and what is worth doing.

Chapter 5

Integrated Program Review: Lessons from Assessment Research at California State University, Fresno

Priscilla Chaffe-Stengel
Professor, Information Systems and Decision Sciences Department
Campus Coordinator of Assessment
California State University, Fresno

Overview

Student outcomes assessment has had an interesting, perhaps telling, history on the campus of CSU Fresno. Over the course of our six years of experience with the implementation of assessment and program review, we have worked in the absence of a statewide assessment mandate. The benefits, as we found, of having assessment goals defined locally include:

- Statewide efforts neither define nor restrict local efforts.
- Assessment programs can be focused on educational impact at the level of the student and on issues deemed important to the institution.
- The identification of local concerns is itself a valuable activity with direct impacts on the campus, community, and, in our case, the university system.

Although we have operated with uncertain funding sources throughout these years, we have managed to use systemwide grants, local funds, faculty and student committees, master's degree research projects, coursewide projects, independent studies and existing survey, testing and accreditation activities to accrue valuable information for relatively little cost.

Our assessment experience is a tale of the importance of timing, consultation, negotiation and perseverance. As a preview to this chapter, a list of valuable lessons we have learned is presented as follows.

- Timing and pace need to reflect the number of actors and their involvement in the assessment activities. Specific assessment grant researches managed by a small team of faculty flourish when conducted at a rapid pace, while assessment activities that involve the consensus of a broad circle of faculty and students require a slower, more methodical pace. In a word, patience encourages consensus and trust. As a corollary, coordination of piggy-backed and multidimensional efforts is important and cannot be rushed. This includes coordination of developing various instruments for data collection, using these instruments across various populations by various data gatherers, and reporting results to various audiences. Effective coordination of multiple tasks across multiple constituents takes time.

- Faculty consultation is necessary but not sufficient to the success of assessment programs. While faculty consultation plays a vital role in the success of any major campus initiative, representative faculty governance alone is not sufficient to deliver adequate and accurate information to all faculty. Our development of an interdisciplinary, campuswide body to recommend policies for the implementation of integrated program assessments gave visibility and the clear indication of our intention to form assessment plans and policies in an open arena. Forming policy in an open environment enhances faculty trust.

- Negotiation for administrative and faculty support is a very important procedure in the implementation of assessment programs. Administrative support is essential, especially in the absence of a state mandate and state funding. Faculty consensus must be built through the faculty governing bodies. Agreement on the distribution of various data to various audiences needs to be negotiated at the beginning of the review process. To share sensitive data too widely not only limits faculty trust, it limits the effectiveness of diplomacy in negotiating actions to be taken in response to assessment results. Faculty need assurances that results will not be summarized in a way that allows comparison of individual faculty.

- The design of a review process that allows for the iterative reform of the process itself gives the flexibility to address significant outcomes as they emerge from the prior year's review cycle. One year to the next, we inverted the reporting process to encourage faculty dialogue and ownership of results; we launched a student survey to probe a troublesome finding from the prior year.

- Student survey results, based on our findings, are reasonably correlated with student test results, at least in reading and writing. The collection of survey results is easier, less expensive and less onerous than the collection of test results.

- The presentation of data needs to precede the discussion to inspire faculty interest. This finding is consistent with Peter Ewell's adage.

- Good assessment need not be expensive.

Introduction

Our story began in fall 1986, when a request to study the writing competencies of our upper division students was initiated by one of the campus Academic Senate bodies. Because it carried a significant implementation cost, the request was reviewed thoroughly by both the Academic Senate and campus administration. What was initiated in fall 1986 as a request was delivered after review by the Academic Senate in spring 1987 as a mandate to conduct the study. Our request for permission to act had become a local mandate to act. Within weeks thereafter arrived the announcement that the campus had been awarded a grant from the CSU system to augment the campus commitment to the study. But the cooperation of key faculty had not yet been gained, largely because we did not know until late spring whether the campus would support the cost of implementation. The first few weeks of fall semester 1987 were critical to the success of the study that was eventually undertaken and, we thought then, to the success of assessment research itself on our campus.

Implementation of the 1987/88 assessment research project went very well. In spring 1988, the campus bid for and won a second CSU assessment grant for the 1988/89 academic year. The researchers and administrators involved, bolstered by the success of the first year's activities, were heartened by the award of a second assessment grant from the CSU system. They planned a half-day program on academic assessment as part of a three-day fall 1988 retreat for 90 members of the campus faculty and administration. A panel presentation was formed with representatives from the local campus, the CSU system and the federal government. Faculty reaction to the presentation, however, was severe, one member labeling the panel members "traitors to academe," instructing those without the requisite faith and loyalty to get out of higher education.

One year later, the Academic Senate accepted with unanimous support the campus comprehensive plan for student outcomes assessment. Much happened in the intervening year.

Assessment of Writing Competency

Assessment research is necessarily complex, drawing from and recording the rich diversity that exists on today's university campus. Assessment that generates understanding of how diverse groups interact within the institution must by its very nature be multivariate, multiple measures from multiple sources. But coordinating the review of proposed assessment activities is often complicated by the very diversity it intends to record. Timing can be of the essence in accomplishing a thorough but smooth review. A case in point follows.

A 1979 systemwide mandate required each of the system's campuses to certify that all graduating baccalaureate students are competent in writing at the upper division level. Each of the campuses in the system has implemented

that requirement in its own unique way. Over the past dozen years, our campus has allowed students the choice of either passing a standard writing examination or passing a one-semester upper division writing-intensive "W" course. The written examination, called the Upper Division Writing Examination (UDWE), includes a nationally normed multiple choice English test, two locally generated tests of spelling and punctuation, a 45-minute essay of personal experiences and a 45-minute essay of argumentation. The campus offers 21 different writing-intensive "W" courses across 19 different departments. Each of the "W" courses requires a minimum of 5,000 words of student composition accomplished in at least five different assignments.

In 1986, the Writing Competency Subcommittee of the campus Academic Senate sought to assess the comparability of the student outcomes evident from the two alternatives available to students. The subcommittee proposed that all students enrolled fall semester 1987 in a "W" course be given the UDWE. The proposed study would produce a common measure of the effectiveness of the "W" courses in preparing students to exit with the minimum writing competence as measured by the UDWE. The cost of testing and scoring the approximate 1,100 "W" course student examinations was estimated to be $14,000, and, if approved, would be supported entirely by campus funds. At the same time, the subcommittee chair submitted to the CSU Academic Program Improvement (API) Grants Program a complementary proposal to add a test of comprehensive reading skills to the writing measures the campus planned to gather.

The on-campus proposal to establish the undergraduate writing competency baseline was forwarded from the Writing Competency Subcommittee, which is composed of nine faculty (one from each school on campus and a representative from the writing program housed in the English Department). The subcommittee meetings were regularly attended by seven other campus personnel, ranging from representatives from the Office of Testing to the chair of the English Department. Additionally, two other faculty members participated in the design of the proposal: the director of the Learning Resource Center and the chair of the Educational Equity Subcommittee. So, in its conception, the proposal draft directly involved 18 campus personnel. The proposal was discussed before the Academic Policy and Planning Committee (nine members), the Executive Committee (eight members) and the Academic Senate (70 members). There were then a total of 105 campus faculty, student and administration members who were aware of, discussed, and voted on the campus proposal to gather the test data. In June 1987, we were notified that our complementary API proposal would be funded. Later that month, a letter was mailed from the academic vice president to all "W" course instructors of record to inform them of the action carried by the Academic Senate, of the API Assessment Grant award, and of the activities planned for the coming academic year. The year-long path that the proposal had taken involved every level of faculty and administrative representation.

In spite of our very real outreach, some of the "W" course faculty remained uninformed of our project, its goals and its purposes. Representative faculty governance did not succeed in delivering adequate information to all constituents. During the first few weeks of the fall semester, numerous telephone calls from "W" course faculty were logged by the Office of the Associate Vice President for Academic Affairs. They objected to their students' participation in the required testing. Their complaints ranged from statements that they were not given adequate notice of the testing, through concerns for the time conflict students had with jobs, through skepticism about the adequacy of a one-sitting test, to distrust of the ends toward which the results would be used. Several students called to complain as well. Their major concerns were the lack of information about the nature of the examination, the additional time it required, and the use of the results. Misinformation was evident.

Faculty and student reactions were heightened when they were told that the study was going to be done; our request to conduct the study and spend campus funds had been passed by the Academic Senate as a mandate to conduct the proposed study. In October 1987, a breakfast meeting was held. All "W" course instructors were invited. Members of the Writing Competency Subcommittee, Office of Testing, English Department, central administration and I answered questions faculty had about the testing. Critical to the cooperation negotiated at the meeting was our guarantee that no results would be summarized that allowed comparison of individual faculty.

In retrospect, it is not clear we could have changed the sequence in which events occurred. It was clear that the cooperation of all "W" course instructors was important to capture the variety of students attracted to the different "W" courses. But we could not notify the "W" course instructors until we had campus approval to spend the funds required to offer the test. By the time we had campus approval, the test was mandated. Perhaps the timing of the events could have been quickened; if the breakfast meeting had taken place prior to the end of the 1986-87 academic year, the faculty directly involved would have had more time to reflect on the purposes of the project, and might have been more positively disposed to participate in the fall. There is evidence in the literature indicating that initially faculty are wary of assessment in general and of student outcome measures in particular. In retrospect, perhaps we should not have been surprised by the initial resistance of some of the faculty involved. Ironically, as a result of the vocal opposition of some of the "W" course faculty, awareness of the project and its goals was enhanced from the on-campus publicity generated in the dialogue. In the end, we had a 61 percent show rate among the "W" course students.

Faculty Involvement in Assessment

Many faculty were involved in our first API assessment project. They talked with their students about participating in the examination and they received the student results prior to the close of the fall semester. A small

cadre of faculty designed the study, oversaw the data collection, interpreted the results and disseminated reports. The nucleus of faculty with primary involvement, however, was small. With successes from the first API grant, with the award of a second API assessment grant, we welcomed the opportunity to plan a half-day presentation for the fall 1988 faculty retreat in Asilomar.

The three-day faculty retreat was planned for 90 members of the campus faculty and administration. The panel presentation on assessment was scheduled for a three-hour block the morning of the last day at the retreat. We arranged for an officer from the Fund for the Improvement of Postsecondary Education (FIPSE), Dr. Constance Cook, to join us from Washington, D.C., to open the panel discussion with a talk about the programmatic benefits of student-based assessment. We arranged for the chair of the CSU Advisory Committee on Student Outcomes Assessment, Dr. Bernard Goldstein, to talk about work at the systemwide level toward a policy to guide the development and implementation of campus-based assessment. And we arranged for two local faculty, Drs. Priscilla Chaffe-Stengel and Penelope Dyer, and the university test officer, Dr. William Stock, to talk about what we had learned thus far from our own campus assessment activities, where we needed to go over the next few years, and how we might get to our intended goals.

Early in the retreat, Dr. Helen Roberts, director of the API Grant Program in the systemwide Chancellor's Office, addressed the role her program plays in facilitating systemwide development of issues in teaching and learning. In opening the assessment panel discussions, Dr. Leonard Salazar, associate vice president for academic affairs, introduced Dr. Cook, whose talk set assessment as the tool of programmatic development that can focus on instructional improvement, motivate faculty development and provide the essential link between teaching and learning. Dr. Goldstein shared the progress made to date by the systemwide Advisory Committee and emphasized the committee's support of assessment that is campus-specific, multidimensional, and longitudinal. [The story of the development of systemwide principles is treated in Chapter 5 of this monograph.] The panel recommended seeking to identify what our institution is in terms of what we accomplish. The audience was left with these questions: What do students know? How did they learn what they know? How do they feel about what they know? And what have they done in terms of what they learned here?

A question and answer period followed the panel discussion. Strong reaction was established by a faculty member who labeled the panel members "traitors to academe." He argued that to assess students meant to call into question the wisdom of the classroom professor. He instructed those without the requisite faith and loyalty to get out of higher education.

As we reflected on the timbre of the meeting, we realized the meeting had introduced a caution that had not before been an explicit element of our planning. We continued the second API grant activity with the same verve we had held before the meeting. What changed was the manner in which we planned for the increased involvement of campus faculty.

Undergraduate Program Review

In fall 1989, the Planning Commission for Undergraduate Program Evaluation and Review was convened by the vice president for academic affairs in cooperation with the Academic Senate. Its fundamental charge was to make recommendations for the integration of program evaluation/review and strategic planning with measures of institutional climate and student achievement. Specifically, the commission was charged with the responsibility to:

- examine different measures of student achievement;
- explore alternative methods of gathering evidence about institutional climate and culture;
- identify campus priorities for program review that encompass advising, teaching, and academic support functions;
- recommend the manner in which measures of institutional climate and student achievement will be utilized in conjunction with program review;
- develop appropriate methodologies to accomplish the integration of these measures into program review;
- recommend policies for the implementation of integrated program evaluation and review.

The Planning Commission was composed of seventeen voting members representing all eight schools, five Academic Senate committees, Student Affairs, the CSU System Academic Senate and the Office of the Vice President for Academic Affairs. The seventeen members also represented staff, faculty, department chairs and school deans. Besides their composition with respect to school, committee and staff/faculty/administrative representation, members of the commission were invited for their active participation in previous campus and systemwide initiatives, and for their willingness to be open yet careful and critical.

Commissioners reviewed a wide spectrum of materials and literature circulated prior to its first meeting, including the working draft of the systemwide policy on student outcomes assessment, materials from our regional accrediting agent, and literatures from national publications and agencies. The commission set the parameters for its work during its first meeting, in which various student populations were identified for inclusion in comprehensive program evaluation. Highlighted were several stages of student progress: (1) matriculation, (2) withdrawal/stop out/drop out, (3) completion of General Education requirements, and (4) completion of baccalaureate major requirements. The commission postponed to a later date its consideration of the review of students at matriculation and withdrawal and focused its second and third meetings on the identification of those student outcomes measures that give a comprehensive picture of the skills and attitudes undergraduate students should have upon completion of their courses at CSU Fresno. The university

mission statement and the goals of the General Education program were considered in determining those broad attitudes and fundamental skills we associate with an education person.

The commission focused its last four meetings on the discussion of departmental review of the major. Discussion at commission meetings recognized the difficulties encountered in assessing the General Education program: that many students transfer with much of their GE requirements already satisfied on other campuses; that there is a multiplicity of courses from which students may choose; and that objectives framed to account for the inherent flexibility in the GE program must also be translatable into appropriate and reasonable ways to assess the levels of success with which the institution is meeting these goals. The commission developed, field-tested and revised a seven-page survey for graduating seniors that attempted to measure student achievement on key attitudes and skills found in the General Education program and the mission statement of the university. The survey was designed as a mechanism to assess at the final undergraduate semester how well students felt they met the goals of the General Education program. By design, it was also seen as a vehicle for helping departments evaluate student achievement beyond individual course goals and grades. These functions are featured recommendations in the CSU Policy on Student Outcomes Assessment. Specifically, the CSU policy states that "the faculty of each department or program should have ways of evaluating student attainment in the major that go beyond the evidence provided by course grades." (CSU Advisory Committee on Student Outcomes Assessment, 1989, p.14) Further, at the campus level, "The faculty of each CSU campus should have mechanisms to assess how well students are meeting the goals for the General Education program of the university." (p.14) And, at the department or program level, "CSU campuses should encourage some faculty in each academic department to engage in assessment research related to teaching and learning." (p.15)

The CSU Fresno Survey of Seniors was designed as one of three major components of the departmental undergraduate program review. Departments undergoing undergraduate program review, it was planned, will include a comprehensive self-study, incorporating information from departmental and school strategic plans, external accreditation reports, and description and analysis of programs by departmental faculty. A team of three faculty will be identified to conduct an on-site visitation; the composition of the team is to include a content expert from another university and two CSU Fresno faculty one from outside the school and the other from within the school but outside the department undergoing review. In addition to the accrual of faculty input from the self-study and the visitation on the consistency and integrity of the program, the commission recommended that each department undergoing review be required to analyze, interpret and react to the results of the student survey for their own graduating seniors. The university Office of Testing was arranged as the hub for the receipt, scanning and dissemination of descriptive

results of the CSU Fresno Survey of Seniors. The commission further recommended that departments undergoing review should receive a data disk of results for their own seniors to facilitate the process of engaging faculty in the assessment of student learning and the incorporation of student outcome measures in the academic review process.

On April 30, 1990, the recommendations of the Planning Commission on Undergraduate Program Evaluation and Review were presented to the Academic Senate, which gave its unanimous support to the commission's design of the review process for undergraduate programs.

What happened between fall 1988 and spring 1990 to change the campus climate toward assessment occurred on three major fronts. First, the award of two consecutive API challenge grants in assessment (1987/88 and 1988/89) extended the development of our assessment activities, and gave visibility and credence to the quality of the assessment research being done on campus. Second, the campus initiative to form the seventeen-member Planning Commission on Undergraduate Program Evaluation and Review indicated that campus administrators were committed to the development of broad-based student outcomes assessment on our campus. It also indicated that the development of assessment plans would proceed in an open arena with input sought from all constituents on campus. And, third, the recommendations made by the commission were reasonable. They provided for the development of a campus-based instrument, the Survey of Seniors, and outlined the assistance academic departments and schools would be given in collecting, analyzing and reporting results. The recommendations embedded the assessment of student outcomes into undergraduate program review, and placed with the faculty of the academic unit the primary responsibility for the determination of those measures to be featured in program review. The recommendations provided a means for faculty to evaluate student skills and attitudes beyond individual course performance. They provided a vehicle for departmental faculty to assess how well the General Education program is meeting departmental priorities in addition to how well departmental majors are achieving the goals of the General Education program itself. The commission's recommendations parallel many of the recommendations made in the systemwide policy on student outcomes assessment.

General Education Program Review

The last review of our General Education program was completed in 1987. At that time, the review was done by six ad hoc committees of 10-15 members each. Coordination was difficult. Attempts to gather faculty perspectives and student feedback on courses were made, but the sheer size of the task to be accomplished kept such attempts to a minimum.

In planning the next review of the General Education program, we felt an ongoing process that reviewed a portion of the program each year was more manageable and hence more promising in terms of the depth of the issues

such a review could probe. In fall 1990, the campus coordinator of assessment, Priscilla Chaffe-Stengel, met with the General Education Subcommittee to design the process to be used in reviewing the six GE Core categories:

1. Written English
2. Speech
3. Quantitative Reasoning
4. Critical Thinking
5. History
6. Political Science

Faculty surveys were formed to explore issues relating to faculty expectations for student work, common course goals and requirements, enforcement of prerequisites and multi-section coordination. Because there is a concern on campus that students are not taking their GE Core courses early in their academic program, enrollment data were accrued to reflect the distribution of students by class level. And evidence linking per section enrollments, grade distributions and the employment status of the section instructors (professor/lecturer/assistant and full- or part-time) were collected as well. Student data specific to GE courses were not gathered as part of this first review of the GE program, but information from the Survey of Seniors was included with the review.

In terms of the review itself, summary of the surveys was a bigger job than had originally been planned. While there are only six GE Core categories, they contained 243 different sections taught by 157 different instructors. Information from last year's GE Core review led this year to the creation of a tiered review process for the next segment of the GE program, a process that will have faculty reporting to course coordinators, who will in turn report across GE divisions to a small review team of campus faculty. The small review teams will report to two members of the GE Subcommittee responsible for generating a final report to the associate vice president for academic affairs and the Academic Senate.

The review of GE Core revealed several important findings which led to the redesign of this review process. The first major finding was that although the reports of the review were shared with various administrative and faculty senate groups, the integrated understanding of the issues that arose from the review was too concentrated. The review did not have a significant presence among the very people who embodied the program — its faculty. Specifically, the faculty who taught in the program, who offered the courses, who interacted with the students by and large remained unaware of the results of the review. Ironically, this occurred in concert with one of the major findings of the review, that faculty rarely met other than informally in the hall or over coffee to discuss the progress of their commonly taught courses. Faculty coordination across multiple sections happened only informally within a department. Faculty interaction across courses that satisfied common General Education divisions occurred rarely if ever across campus departments.

We explicitly redesigned the second year of the GE program review to, in a sense, invert the process we had used the year before. The tiered review process was endorsed because it would (1) require faculty who teach multiple

sections to come together to discuss the course with their course coordinators, and (2) then have course coordinators come together with other coordinators from across common GE divisions to discuss their findings. It was because the GE review was conducted in year-long segments that we had the opportunity to redesign the review process itself, to better address the issues as they surfaced from the prior year's review.

The second major finding of the review of GE Core was the frequency with which some of the courses were offered in large-class format. Four of the six GE Core requirements had class sections that were enrolled at 94 to 184 students. Approximately 20-27 percent of all students enrolled fall 1990 in a course satisfying these four GE Core requirements were in a large section. Two interesting developments came on the heels of this finding.

First, when the report was shared with our senate subcommittee, faculty were basically uncomfortable with the statement that class sizes of some GE courses were too big. They questioned everything, from how representative was the time frame within which evidence was gathered, to how calculations were performed to support the statement that the class sizes of some sections were too high. In support of an adage shared by higher education specialist Peter Ewell (NCHEMS of Boulder, Colorado), when data preceded the discussion, faculty interest was sparked. When we put data on the table, the evidence was so compelling that faculty no longer questioned the validity of the process or the conclusion. Where just the week before, we had been struggling to convince faculty that the courses were too large, the presentation of section enrollments with ranges and average enrollments across each course led one member to quip after a brief review of the data, "What would happen if we put an automatic cap on section enrollments effective next fall?" In fact, the presentation of data was compelling in a way that prior argumentation had not been. [We did not, by the way, move to cap per section enrollments; the problem is currently under discussion with other senate officials, however.] To be fair in our recommendations to assessment practitioners, we want to caution against the overexposure of sensitive data. Distribution of evidence needs to proceed carefully and needs to be negotiated at the beginning of the review process. To share sensitive data too widely not only limits faculty trust, it limits the effectiveness of diplomacy in negotiating actions to be taken in response to assessment results. It is nonetheless valuable to remember that. when initial faculty reaction to assessment activities is uninspired, introducing the discussion with data sparks faculty interest.

The second interesting development came when one administrator asked in light of these large GE Core sections, "So what?" We had assumed the answer was obvious that a student enrolled in a large class would have a different experience, one marked by more standardized testing, fewer opportunities to participate in class discussions, reduced interaction with the faculty teaching the course, etc. Upon reflection, however, we realized that we did not have any evidence to indicate the effect of class size on the quality of the experience students have. To that end, we are currently designing a 50-item

student survey to administer later this spring to approximately 8,370 students enrolled in GE Breadth courses, in which we will ask questions intended to develop our understanding of the relationship between class size and student experience. Again, this revision of the review process itself was possible because the review was cast in short, one-year segments, which allowed us to reform the process at regular intervals.

The review of GE Core courses also revealed that there were no written guidelines for GE Core categories in speech, history and political science. Guidelines for speech and political science have been proposed and approved; guidelines for the GE Core history category are yet pending before the subcommittee. Issues of multisection coordination were discussed with the faculty and department chairs involved. These discussions led to commitments for increased training of part-time instructors and for the creation of instructor's manuals for one course.

Results from the CSU Fresno Survey of Seniors established our first campuswide baseline for self-reported skill and attitude levels for graduating seniors. Campuswide results were disseminated along with an electronic database of responses for participating students from each school and/or department, to encourage each academic unit to probe the characteristics and experiences of their own students. Campuswide, concerns were raised by the fact that just over 23 percent of our graduating seniors reported that, over the last six months, they had not read a book other than those assigned in their classes. In response, the vice president for academic affairs is initiating the development of a campus reading list and supporting program to encourage students to read more frequently and more broadly.

Several issues remain for the GE Subcommittee to consider before a final report on the GE Core Review is issued to the campus Academic Senate. A summary of the distribution of students by their class level indicated that too many students are waiting until their junior or senior year to complete GE Core requirements. Data from the review are being used as a basis for a major initiative to develop a comprehensive redesign of our computerized student information management system. Specifically, the data will provide a justification for the development of a system that will allow on-line electronic access to student transcripts from feeder institutions as well as our own campus. The goal of computerized checks for course prerequisites may well be several years away for this campus, but progress toward that end is being made.

Budget Issues in Assessment

Much discussion has occurred recently in California about the fiscal crisis in state-supported higher education. The story of assessment in the CSU system cannot adequately be told without reference to the budget cuts to the system and its campuses. CSU Fresno has moved the assessment agenda forward despite deep cuts to the campus allocation, however, and has done so in large part because of (1) the administrative initiative of the leadership in Academic Affairs from Drs. J. Leonard Salazar, Judith Kuipers and Alexander

Gonzalez, (2) some good luck, and (3) fiscally sound and sometimes innovative assessment design proposed by both the commission and the campus coordinator of assessment, Dr. Priscilla Chaffe-Stengel.

In the absence of a statewide mandate to assess student outcomes, campus progress in developing and implementing comprehensive assessment is a direct result of administrative initiative. In the absence of a statewide mandate to assess and under the impact of a deteriorating budget, administrative commitment is essential. We have also had some good luck to complement the administrative commitment evident on our campus. We began assessment on our campus in 1987 when budgetary times were not so lean. The infusion of funds from the two API challenge grants amplified the impact that locally committed funds could have in advancing the assessment agenda. By the time the budget worsened to the extent that the campus support for assessment was questioned, assessment activities had become a routine feature of Academic Affairs. That was fortunate. We had enough time before the budget crisis descended to establish the merits of assessment research.

But when the budgetary knife did fall, the review of campus assessment activities showed that fiscally sound design had been used even in less lean years. In the collection and analysis of assessment results, we have used faculty committees, student committees, master's degree researches, coursewide projects, independent studies and information that is routinely gathered by other campus agencies. Innovative use of assessment initiatives as class projects produced a database of 1,400 student records gathered at the exit to the campus library, where the students were asked the frequency with which they visit the library, the typical activities they perform there and their specific purpose in the visit sampled. A second class project produced an in-depth analysis of the "stop out/drop out" population on campus, those students who were qualified to reenroll but did not. Both of these projects were completed for the reproduction cost of the survey instrument and a few long distance telephone calls. As an aside, from every class involved in such a project, two to six students have continued the following semester working in independent study on some special aspect of the class project. Over the past several years, several graduate students completed master's theses that extended the application of our student surveys to different student populations, one to the student users of our Learning Resource Center and another to student athletes. Much valuable information has been developed through the participation of these student researchers, and they report as quite valuable the hands-on experience they take away from these involvements.

As a detailed example of fiscally conservative design, in a larger study that was undertaken to establish a campus baseline for the assessment and analysis of undergraduate reading and writing competencies, we developed a rationale for using student survey data in lieu of more expensive reading and writing test data. Our analysis showed that the interrelationship between student reading and writing skills is reasonably stable whether reading is defined

in terms of test scores or self-reported activity levels. The study's central premise that reading and writing are directly related was supported when both reading and writing competencies were measured in terms of test scores, and was supported whether the test scores were objective or essay in nature. We found that both better readers and better writers do maintain higher GPAs. Some minor differences surfaced when reading skills were defined in terms of self-reported activity levels rather than test scores; the number of books read in the last six months affected most of the objective tests and the argumentation essay scores, while the number of minutes per day spent pleasure reading affected both essay samples but not the objective tests. The more frequent browsers in the campus library wrote better essay samples and were better spellers, for example. We concluded that survey data can be used with moderate reliability in lieu of test scores. Campus personnel interested in assessment can be encouraged to use survey data which are cheaper and less onerous to gather than student test data. As campuses grapple with deepening budgetary crisis, this is good news to share.

Overall, the assessment activities to date have cost the campus photocopying expenses for reproducing surveys, 9 weighted teaching units in reassigned time per year for the coordinator of assessment, and approximately $800 a year for the production of the customized CSU Fresno Survey of Seniors, which can be processed by an optical mark reader. Time to scan, summarize and report results of the Survey of Seniors is not included here; processing is ably done by our university test officer, Dr. William Stock. Some intersession time was extended once to complete the summary of the unexpected number of GE Core faculty surveys. Good assessment need not be expensive.

Conclusion

In summary, student outcomes assessment has had both an interesting and a telling history on our campus. We have learned some interesting lessons as implementation unfolded. The faculty request that became an Academic Senate mandate bore difficulties in coordinating participating faculty. The development of specific assessment grant research thrived under a full press, but the pace to successfully involve a broader circle of faculty in assessment activities was slower, more careful, more deliberate. Timing and pace are critical elements to monitor as implementation of assessment research unfolds.

We have used a number of vehicles to carry the germ of assessment. The creation of the campuswide seventeen-member commission to plan the policies and procedures for undergraduate program review became our campus vehicle for increasing faculty awareness about assessment. The review of our General Education Core program will be an example to carry through the Academic Senate structures to demonstrate the benefits of program review that integrates faculty and student input. As we bring undergraduate program review on line, we will be able to implement an academic review process that

can provide the tools with which we can fine tune our educational programs. Program review can operationalize our educational goals as we appraise the educational significance of what we do. The emphasis on measures of student achievement promises the opportunity to trace the impact of courses and programs on individual students and offers faculty the potential for applied research into educationally significant issues. Assessment offers the potential to investigate individual student differences in response to different education programs occurring in varied institutional settings. In short, it advances the opportunity to investigate the allocation of educational resources and their impact on the relationship between teaching and learning. These are exciting times for educators.

Reference

CSU Advisory Committee on Student Outcomes Assessment. (1989). *Student outcomes assessment in the California State University: A report to the Chancellor.* Long Beach, CA: The California State University, Office of the Chancellor.

Chapter 6

Assessment and Equity in the California State University

Ming Lee
Assistant Director
Institute for Teaching and Learning
California State University

Rapid changes in demographic and social realities pose enormous challenges to higher education in the United States. How to adapt instructional methods and materials, various academic and student service programs, and assessment instruments and practices to make education more accessible and equitable to everyone is a big challenge facing every higher learning institution. As the largest four-year public educational system in the country, the California State University (CSU) has already had a diverse student population close to what is projected for American society in the twenty-first century. CSU experiences in adapting systemwide assessment programs to meet various needs of their students may be worthwhile sharing with other institutions.

Profile of CSU Student Population

In fall 1990, the 20-campus CSU system enrolled almost 370,000 students (CSU Division of Analytic Studies, 1991). Thirty-six percent were nonwhite, 55 percent were women, 42 percent were adults aged 25 or older, and 40 percent were part-time students. The percentage of nonwhite enrollment ranged from 72 percent at the Los Angeles campus to 11 percent at the Chico campus. During the past decade, the CSU Asian or Pacific Islander enrollment increased the fastest at about 158 percent, while the African American enrollment increased the slowest at 17 percent and Latino enrollment increased 94 percent. This growth of enrollment among ethnic minority students in the CSU outpaces that in the nation, which shows an increase of Asian or Pacific Islander, African American, and Latino students at 111 percent, 7 percent, and 63 percent, respectively (Schantz & Brown, 1990). The overall minority component of more than one-third of the CSU students is also significantly larger than the national average of 18 percent.

It is within this diverse context that the CSU systemwide policy recommendations were formulated, systemwide assessment practices have been taking place, and pilot programs have been implemented on campuses. The chancellor-appointed Advisory Committee on Student Outcomes Assessment recognized the impact of student diversity in assessment activities, noting that "the value of outcomes assessment is limited if it is not accompanied by an understanding of student characteristics and opinions and how these relate to general institutional patterns over time." (CSU Advisory Committee on Student Outcomes Assessment, 1989, p. 17)

CSU Systemwide Assessment Practices

Several CSU systemwide assessment programs have been in place for a number of years, including the following:

- **The English Placement Test (EPT)** — Established in 1976 by the CSU Board of Trustees and jointly developed by the CSU and the Educational Testing Service (ETS), this test measures the writing skills of entering lower division students and places them in appropriate composition courses.
- **The Graduation Writing Assessment Requirement (GWAR)** — The writing skills policy approved by the CSU Board of Trustees in 1976 also included a provision that all students demonstrate writing competency as a requirement for graduation and as a prerequisite to classified standing in graduate programs. No systemwide uniform examination has been adopted for fulfilling this requirement. Instead, each campus is asked to tailor programs to suit local situations and particular needs of the disciplinary majors.
- **The Entry-Level Mathematics Examination (ELM)** — The CSU introduced the ELM examination in 1983 as part of its basic skills assessment process for first-time freshmen and undergraduate transfers. The ELM must be taken and passed by all CSU students prior to enrolling in courses that satisfy the general education requirement in quantitative reasoning.

In addition to these entrance and exit requirements for English and mathematics competencies, the CSU also collects student data through the following systemwide surveys of current and former students:

- **Student Needs and Priorities Survey (SNAP)** — Every 4 to 5 years, the CSU surveys currently enrolled students regarding: (1) life goals and educational priorities; (2) personal lives and family backgrounds; (3) levels of satisfaction with various aspects of their academic and social experiences on campus; and (4) obstacles or problems, whether institutional or personal in nature, which might hinder progress toward their educational goals. SNAPS surveys were conducted in 1981, 1984, and 1989 with more than 35,000 students from up to 18 CSU campuses participating. Special attention has been paid to the particular needs and satisfaction level of ethnic minority students.

- **Survey of Spring Graduates** — Since 1975 the CSU has conducted a biennial employment survey of all spring graduates from all CSU campuses. Data collected in this survey provide campuses with answers to such questions as: (1) What do CSU students do after graduation? (2) What are rates of employment for women, minorities, and older graduates who seek employment? (3) Do CSU graduates get the jobs for which their major programs prepare them? (4) How do CSU graduates find jobs and what are their starting salaries?

Data collected from surveys of current and former students can be integrated with various administrative databases for research purposes. We are currently in the process of attempting to combine several of these databases so that, we hope, next year we will be able to compare student needs and priorities with a variety of outcome measures to evaluate the impact and quality of specific programs over a long period of time. These data also allow campuses to compare characteristics, opinions, and achievements of students from their own campus over time and with those of CSU students in general.

These systemwide assessment programs and practices are not designed especially for nontraditional students. Campuses supplement the systemwide assessment programs with a wide variety of instruments designed particularly for nontraditional students. For example, although research data indicate that the English Placement Test (EPT) is as appropriate for measuring competency of language minority students as of native students, the test is perceived to provide insufficient diagnostic information about ESL students' writing skills. Some campuses then use supplementary instruments for placing ESL students, such as Michigan Test of English Language Proficiency, Michigan Test of Aural Comprehension, Clinical Evaluation of Language Functions — Advanced Screening Test and Diagnostic Battery, etc. An alternative recommended by the CSU Advisory Committee on English as a Second Language (1988) is to develop a separate version of the EPT containing items with appropriate cultural content and more depth of diagnostic information. The committee also recommended to norm the test scores on nonnative speakers of English.

Another example of supplementing systemwide requirements with appropriate alternatives to make assessment more equitable for nontraditional students can be found in the Graduation Writing Assessment Requirement (GWAR). For most campuses, no allowances are made in the ways the GWAR tests are given or scored. However, some campuses do permit ESL students extra time to complete the examination and/or adjust scoring practices to compensate for minor mistakes characteristic of second language speakers. It is believed by these campuses that faculty readers and interpreters of test results need to be sensitive to the writing of language minority students and not penalize them for mistakes in idiom or other minor errors.

The CSU just recently adopted a systemwide policy to provide reasonable and necessary accommodations to students with disabilities in the

EPT/ELM programs, and in any campuswide examination administered to satisfy the CSU GWAR requirement. The accommodations include braille and large print test versions, cassette recordings of test items, Signed Exact English (SEE) and/or American Sign Language (ASL), extra testing time, alternative test locations, etc. The purpose of this policy is to ensure that test results accurately reflect the student's English and/or mathematics skills rather than the constraints resulting from the student's disability. This policy serves as another example of adaptations that assessment programs can include in order to reach the goal of equity.

CSU experiences with student outcomes assessment programs suggest that adaptation for equitable assessment is possible. However, for any new assessment programs, equity should be included as one of the ultimate goals set at the very beginning of the planning stage.

References

CSU Advisory Committee on English as a Second Language. (1988). *Report of the English as a Second Language Workgroup.* Long Beach, CA: The California State University, Office of the Chancellor.

CSU Advisory Committee on Student Outcomes Assessment. (1989). *Student outcomes assessment in the California State University: A report to the Chancellor.* Long Beach, CA: The California State University, Office of the Chancellor.

CSU Division of Analytic Studies. (1991). *Enrollment by ethnic group: Fall term 1990.* Long Beach, CA: The California State University, Office of the Chancellor.

Schantz, N.B. and Brown, P.Q. (1990). *Trends in racial/ethnic enrollment in higher education: Fall 1978 through Fall 1988.* Washington, D.C.: U.S. Department of Education, National Center for Education Statistics.

Chapter 7

CSU and Assessment — Second Down and Eight Yards to Go: A View from the Scrimmage Line

Ralph A. Wolff
Associate Executive Director
Western Association of Schools and Colleges

The Senior College Commission of the Western Association of Schools and Colleges (WASC) adopted a formal standard requiring assessment of institutional effectiveness at the time of a significant revision of our standards of accreditation in 1988 Standard 2.C (under Standard 2, Institutional Purposes, Planning and Effectiveness) states:

> *The institution has developed the means for evaluating how well, and in what ways, it is accomplishing its purposes as the basis for broad-based continuous planning and evaluation.*

Since adopting this standard, we have conducted visits to 15 of the 20 CSU campuses; visits to two more are scheduled for 1991-92 Eight of the visits already conducted were comprehensive visits, occurring after an extended period of campus self-study. The remaining seven visits were shorter, interim visits with smaller teams, but several team reports commented on assessment issues.

Because the language of Standard 2.C is not self-implementing, WASC also has been developing a more specific and focused assessment initiative. Part of this initiative has been to include at least one evaluator with extensive experience in assessment on all comprehensive evaluation teams. We also have developed a set of clear directions, if not expectations, for all institutions and a resource manual to help institutions engage in assessment. Two drafts of this resource manual have been circulated to institutions for comment, and it is now ready for final Commission approval and printing.

Some general patterns emerge as a consequence of the overwhelming cycle of visits. There are four areas I will focus on in this paper: (1) general

findings of our accreditation visits to CSU campuses; (2) how CSU institutions are reporting assessment activities in self studies; (3) general observations about the current state of assessment within CSU; (4) concluding recommendations that include future directions for WASC.

Drawing conclusions about assessment on all 20 CSU campuses immediately runs the risk of making generalizations that are too broad. Some of the statements made may not reflect the conditions on all 20 campuses, but they do reflect general patterns. Within institutions as large and complex as the typical CSU campus, it is difficult to generalize even within a single institution. Nonetheless, my goal is to create a springboard for further discussion and action.

General Findings of Accrediting Teams to CSU Campuses

Think of where we are now in assessment in terms of a football game. We are suited up in the game, and there has been a first down effort on nearly every campus. More often than not, it has been a credible effort of assessment. And as on any team, there has been an effort even before the game started to find the talented players, engage in some training, even attempt to define some of the rules, and then come to the stadium to engage in the real activity. But to give you my conclusion at the outset, we are only at second down. We have a long way to go, and even though the goal posts are visible, the view at the line of scrimmage is formidable. The defensive line against which we are playing is worse than facing the San Francisco 49ers, and a game-winning strategy is not yet evident.

There is good news and bad news to report as I see things from the line of scrimmage. First the good news: There are a lot of things going on within CSU that are solid, positive and, although unheralded, valuable assessment. At every campus, our teams found a number of very credible and useful assessment activities under way. These take the following forms:

- **Basic research on selected issues** — At every campus, even those without a formal office of institutional research, there is an extensive array of data collected. They deal with such important issues as enrollment, degrees earned, attrition and persistence, faculty workload and dozens of other variables. On several campuses there is movement toward a student tracking system that appears to have great promise. All campuses have student evaluation of teaching in some fashion.
- **Diagnosis of basic skills and provision of remedial support** — A fundamental assessment activity is ensuring that students admitted have the basic skill level to undertake collegiate level work. An unsung role of the CSU has been the testing of basic skills in composition and mathematics and then offering the necessary remedial courses for students to progress to degree-oriented programs.
- **The systemwide requirement of an upper division writing requirement** — The CSU has taken an important stand in establishing a writing requirement, and various campuses have developed responses to the

requirement and support systems for those who fail to meet the requirement.

- **Program and general education reviews** — All campuses have some form of program review, and many have undertaken formal reviews of their general education programs. These reviews are a form of assessment and have the potential to reveal significant areas of needed improvement.
- **SNAPS Survey** — The SNAPS survey yields useful information for campus planning and provides the basis for follow-up studies and programming.
- **Departmental assessment efforts** — On each campus our teams have found some departments attempting to assess student learning and outcomes. Often these departments are in the sciences or professional areas, such as nursing or engineering. Their assessment efforts typically involve surveys of graduates and employers.
- **CSU system-sponsored projects** — The assessment pilot projects sponsored by the Office of the Chancellor are extremely valuable examples of leadership and experimentation. These have not been widely described in our evaluation team reports, but my participation in CSU systemwide conferences on assessment has enabled me to see the enthusiasm and change occurring as a result of these efforts.

These seven forms of assessment activity are those most commonly cited by our evaluation teams. But there also have been common areas of concern or weaknesses found in CSU campus approaches to assessment. These weaknesses can be summarized as follows:

- **Data collected are not well disseminated or used.** Uniformly our teams have reported that the extensive array of data collected on CSU campuses is not widely distributed; and even those data that are adequately distributed are not effectively employed in campus planning and decision making. My own observation is that the initiating source for much of the data collected is the CSU system office, rather than an individual campus. Thus, a considerable amount of data goes from the campus to the Office of the Chancellor but does not get distributed on the campus.
- **The focus of assessment within CSU is still not on what students have learned.** Perhaps the single most important emphasis of the national assessment movement is on improving student learning. That requires considerable attention to the teaching/learning continuum. For a variety of reasons (some of them tied to the culture of CSU, which I will comment on later), there has not been as significant attention placed on student learning outcomes as there has been data collection around student movement through the institution.

- **Program and general education reviews do not emphasize such important issues as the clarity of learning objectives for the curriculum, student learning outcomes, or departmental climate.** Program reviews function much as accrediting reviews have historically. A departmental self study is reviewed by a group of visitors from other departments or outside the institution. The focus is more on what courses are taught, resources available, faculty workload and scholarly accomplishments than on the effectiveness of the curriculum in accomplishing its goals. Rarely does a campus require that a common set of questions be addressed in program and general education reviews which address assessment issues.

- **Faculty are not systematically engaged in asking powerful questions and getting answers to these questions that impact and improve student learning.** The fundamental assessment question, which should be built into the fabric of the curriculum and be the province of the faculty, is whether the faculty are asking the questions and collecting the evidence that fosters constant improvement of teaching and the curriculum. Apart from isolated examples within a campus (often externally stimulated), our teams and my own personal observations do not find systematic faculty inquiry about issues that should be relevant to faculty.

  Models such as the Harvard assessment seminar are widely available for replication. They provide a structure for addressing such questions as what key factors lead women and minorities to graduate school or to major in certain disciplines. A study at UC Berkeley has recommended that the classroom is the best resource for engaging students in projects that promote real interaction among students, especially students from different racial and ethnic backgrounds. There are many areas for follow-up study from this and other studies on diversity. In addition, classroom research is a marvelous tool for engaging faculty in creative ways to improve teaching without major budget commitments.

- **Nearly all campuses have undertaken significant campuswide planning efforts yet these efforts are largely disconnected from assessment efforts or program reviews.** Strategic planning has been a major thrust on a number of campuses. Data from assessment activities have not informed these campus plans; and in the few cases where priorities have been set, there is little effort to reevaluate them through assessment projects designed to determine institutional or departmental impact. In some cases, the planning exercises did not take into account the current budget crisis. As a consequence, the plans are not as useful as they could be.

- **The commitment from CEOs and CAOs for assessment has been variable throughout the system. But it is rare to find sustained attention from campus leadership.** In a national study of directors of institutional research at more than 900 institutions, 80 percent of the

respondents believed that the visible commitment of the CEO and/or CAO was the key factor for "facilitating" implementation of assessment. The leading factor cited by 60 percent of the respondents as impeding implementation of assessment was the indifference or resistance on the part of the faculty.

- **Much of what the campus is doing well in assessment is not widely known.** As I have talked with members of campus assessment committees, I commonly have found that there are examples of assessment "discovered" when departments are surveyed about assessment. These models need to be reviewed and highlighted. Management theory emphasizes learning from the "winners" as much as learning from mistakes.

How Institutions Have Addressed Assessment in Self-Studies

There are three areas in campus self studies where assessment should play an important role. Under Standard 2.C, institutions should set forth an overview of assessment and analyze approaches the institution has taken, indicating how assessment efforts are (or are not) linked to institutional planning. Under Standard 4 on educational programs, there should be an analysis of what information the institution has on the effectiveness of general education, not just its design and rationale. Evaluations of the program review process (Standard 4.F) should indicate how assessment data are incorporated in program review and lead to improvement of teaching, curriculum emphasis, course requirements and the like. Under Standard 7, there should be evidence of the *effectiveness* of student service programs rather than a description of structure, staffing and resources. A special focus should be on campus climate for women, minorities, adult students, and other special categories, as well as for majority students.

Little of this is happening now. The typical self-study provides a description of the *process* used to address assessment under Standard 2.C and typically indicates that the campus is just getting under way. The usual first step is a campus assessment committee. The committee is usually working on definitions and a campus policy statement on assessment. Occasionally, the self-study report will identify specific assessment activities. Little information is provided, however, as to whether such activities are effective, how they have contributed to needed changes, how the results or process used were disseminated to others as a model, or what new assessment experiments will be initiated in the future.

Under Standard 4, there is typically very little evidence or data about effectiveness. Most of the discussion is about design of curricula, processes of review, and resource allocations. Chapters on Standard 7 (Student Services) focus more heavily on describing student service programs, staff and resources than on evidence of campus climate or the impact of student service programs in fulfilling the educational mission of the institution.

These common observations about our team reports and campus self-studies on assessment lead me to some comments on the culture of the California State University system. Organizational literature places considerable attention on how the operational culture of an organization shapes institutional response to problems and individual behavior. My observation of CSU campuses leads me to some observations about how the culture of the CSU has affected campus responses to assessment. I state these as neither good nor bad, but as a commentary on how the culture of the system appears to shape behavior.

- **CSU campuses approach new issues structurally rather than entrepreneurially.** As a consequence of this operational culture, the first effort of dealing with assessment was to develop a systemwide policy statement; the same has been true on most campuses where initial attention has been placed on campus policies on assessment. This approach has been taken rather than starting with many small experiments with different assessment models. While this has created a policy environment for assessment to begin, it has stymied the initiation of assessment efforts "until our policy is developed."

- **The formula approach to funding CSU creates conceptual barriers to addressing student learning and curriculum issues, thereby affecting assessment motivation and methodology.** The formulaic paradigm of CSU is so deeply embedded that many discussions about curriculum are shaped by resource allocation questions rather than evidence of learning effectiveness. In addition, there has been a reward system for thinking that more students are better because they provide more resources. There have not been financial or other rewards for addressing how effective each campus has been with the students it can afford to serve. As campuses move to set priorities in times of great constraint, evidence of effectiveness will be crucial but will not be readily available.

- **There is an environment of zero sum thinking when it comes to assessment (and other issues).** It is hard to talk about assessment without addressing the "add on" syndrome. "If it is added on to my existing workload, I won't do it. Take something away first." Assessment needs to be seen not so much as something new but as something that already is a fundamental part of our work as faculty and administrators.

- **Issues of limited trust and communication affect the willingness of faculty to become involved in something new.** Problems of communication between the campus and the Office of the Chancellor, or within the campus itself, can make assessment more a political football than a serious educational issue.

The Current State of Assessment Within CSU

What is the current status of assessment within CSU? It is not possible to talk about assessment without first talking about the effects of the state's

fiscal crisis on campuses. The current state of campuses is grim; there is considerable demoralization in some places, little optimism, and no sense of control over one's own destiny. There is more a sense of preserving what we have rather than enthusiasm for embarking on major new journeys. Campus leadership is unable to portray a clear picture for the future, and presidents themselves do not know from day to day what new round of budget cuts may be ordered next.

Whither WASC: Concluding Recommendations

Within this context, assessment from the line of scrimmage looks a lot like razzle dazzle, and there is a desire to stay with the tried and true plays. In some cases, assessment *initiatives* are being put on hold. Campus assessment committees may continue to operate, but without much fervor in some cases. Thus, I see a time out being called on some campuses in order to wait until some of the budget-cutting dust has settled.

Unfortunately, the external world, including WASC, will not wait. While we will indeed be sensitive to the current situation, the need for campuses to progress with assessment remains. We are at second down because there is a good base from which to work at most campuses — there are considerable data already available; there is a systemwide policy statement on assessment that is supportive and effective in responding to the principal concerns of faculty; many campuses have or are near completing campuswide statements of assessment; and there are some models of assessment on every campus worthy of highlighting. But we have a long way to go... eight yards for a first down, and a touchdown is still a ways off.

There are some concluding recommendations for us all, WASC as well as CSU.

- **Our goal should be the creation of a culture of evidence.** That is the intent of our assessment initiative. An institution of quality, especially one confronting hard choices and priorities, needs usable and effective data to act. Such data need to be in the hands of the responsible parties — administrators and faculty alike. Much can be done to distribute data more effectively and include data more wisely in our decision-making processes.

- **Assessment should be embedded in the institutional culture, not added onto it.** The thrust of our initiative at WASC is to incorporate assessment in general education and program reviews, and in campus planning efforts. These activities already occur; now we must shift the attention on such processes to include the clarity of learning goals and the success with which they are accomplished, and then to link these back into real improvements.

- **The focus of assessment should be on student learning and diversity issues.** Building a culture of evidence means that it is important to

have better evidence to support all aspects of an institution's functioning. These two areas need special attention.

- **Attention should be paid to scale issues.** Most CSU campuses are large, complex organizations. As institutional research attempts to generate broad, institutionwide data on important topics, attention also needs to be paid to small entrepreneurship at the school and departmental levels. Support needs to be given internally for individual projects that start small and allow the campus to build incrementally. Emphasis can be placed on classroom assessment and research as well as other strategies.

- **The motivating force for assessment will need to shift from outside the institution to inside.** Assessment should be done because it demonstrably improves the institution. Yet the primary motivating force for most assessment is external mandates. There is no legislative mandate in California yet. But the Commission on Teacher Credentialing is requiring clear competence assessments; professional accrediting agencies, which wield great influence over departments, are moving in this direction, and so will WASC. The purpose is not "to prove" something to us, but to do it to reconnect us all to why we have chosen education as our profession.

- **Assessment is not something that can get started along with the beginning of a self-study process. It only is effective if ongoing and continuous.** Too often I see institutions treat assessment as an accreditation mandate that only needs to be taken seriously episodically, two years before the WASC visit. The kinds of assessment projects that are effective within institutions, even building a subculture of entrepreneurship, take years.

- **The CSU needs to be more directly and visibly aligned with national movements for the improvement of higher education.** There are many great stories to tell about the CSU, about each of the 20 campuses. There are large-scale stories about the enormous remediation job being performed without significant resources, about the service provided to adult students and first generation college students, about major curricular improvements that are under way; there are individual stories as well, about specific faculty and students.

  But CSU is not visibly in a leadership role with the national movements on general education reform, assessment, or the redefinition of scholarship or the development of teaching portfolios. Even the new call for emphasis on undergraduate teaching by research universities should bring a visible response, with evidence drawn from assessment already under way, that CSU is a leader that research universities can follow.

WASC wants to and is prepared to work with institutions and support their efforts. Our efforts are to support institutions as much to prod. We look forward to working together with the CSU and other higher learning institutions.

Appendix A

Evaluation of Student Outcomes Assessment Pilot Projects in the CSU

Project Steering Committee

Priscilla Chaffe-Stengel
Professor of Information Systems
and Decision Sciences
California State University, Fresno

Peter Ewell
Senior Associate
National Center for Higher Education
Management Systems

Bernard Goldstein
CSU Faculty Trustee and
Professor of Biology
San Francisco State University

Olita Harris
Professor of Social Work and
Associate Dean of Undergraduate Studies
San Diego State University

Becky Loewy
Professor of Psychology
San Francisco State University

Angel Sanchez
Associate Director
Analytic Studies
California State University
Office of the Chancellor

Daniel Stufflebeam
Director
Evaluation Center
Western Michigan University

John Tafel
Director of Authorization and
Director of Program Excellence
Initiatives
Ohio Board of Regents

Frank Young
Director
California Academic Partnerships
Program
California State University
Office of the Chancellor

Project Director

Helen Roberts
State University Dean
Instructional Improvement
California State University
Office of the Chancellor

Assistant Director

Ming Lee
Assistant Director
Institute for Teaching and Learning
California State University
Office of the Chancellor

Appendix B

Student Outcomes Assessment in the California State University

Campus Pilot Project Directors

Year	Director/Department	Campus	Project
1986-87	Andrew Moss/English	Pomona	Enhancing Quality by Assessment: A General Education project
1986-87	Kenneth Nyberg/ Sociology	Bakersfield	An Empirical Evaluation of Five Baccalaureate Social Science Programs
1987-88	Catherine Dezseran/ Theatre Peter Grego/Theatre	Northridge	Student Outcomes Assessment in Academic Program Improvement in Theatre
1987-88	Priscilla Chaffe-Stengel/ Information Systems and Decision Science	Fresno	Assessment of Undergraduate Writing Competence
1987-88	Harry Polkinhorn/ English	San Diego	Student Outcomes Assessment: Liberal Studies Major
1987-88	Leigh Mintz/ Geological Science Lu Ann Duffus/ Economics	Hayward	Assessment of Majors: A Three-Campus, Three-Discipline Model
	Richard Giardina/ Political Science Newman Fisher/ Mathematics	San Francisco	
	Leon Dorosz/Biology Howard Shellhammer/ Biology	San Jose	
1987-88	Kenneth Nyberg/ Sociology	Bakersfield	An Empirical Evaluation of the Impact of Undergraduate Curriculum: Year Two
1988-89	Eugene Clark/ Political Science	Bakersfield	Knowledge and Attitudes in General Education: CSU-Community College Joint Assessment

Campus Pilot Project Directors (cont'd)

Year	Director/Department	Campus	Project
1988-89	Chris Cozby/Psychology Jeffry Young/Gerontology	Fullerton	Student Outcomes Related to Curricular Variety in Gerontology
1988-89	Priscilla Chaffe-Stengel/ Information Systems and Decision Science	Fresno	Assessment of Undergraduate Reading Competence
1988-89	Catharine Lucas/English	San Francisco	Assessing Outcomes for English Teacher Candidates
1988-89	Lu Mattson/ Liberal Studies	Sonoma	Curriculum and Pedagogy Effectiveness
1989-90	Bessie Marquis/Nursing	Chico	Outcomes Assessment of Four Classes of Nursing Graduates
1989-90	Mary Cullinan/English	Hayward	Assessment of Student Outcomes. A Basic Writer's Writing Program
1989-90	Lu Mattson/ Liberal Studies	Sonoma	Integrating Student Outcomes Assessment into the Curriculum

This publication was supported by a grant from the
United States Department of Education
Fund for the Improvement of Postsecondary Education

California State University
Institute for Teaching and Learning
400 Golden Shore, Suite 132