

University of San Diego

Digital USD

Dissertations

Theses and Dissertations

2008-05-01

Evaluating the Effectiveness of Provosts in Building a Student Learning Assessment-Supportive Organizational Culture: A Multiple-Site Evaluation within the California State University System

Chris Procello PhD
University of San Diego

Follow this and additional works at: <https://digital.sandiego.edu/dissertations>



Part of the [Leadership Studies Commons](#)

Digital USD Citation

Procello, Chris PhD, "Evaluating the Effectiveness of Provosts in Building a Student Learning Assessment-Supportive Organizational Culture: A Multiple-Site Evaluation within the California State University System" (2008). *Dissertations*. 789.

<https://digital.sandiego.edu/dissertations/789>

This Dissertation: Open Access is brought to you for free and open access by the Theses and Dissertations at Digital USD. It has been accepted for inclusion in Dissertations by an authorized administrator of Digital USD. For more information, please contact digital@sandiego.edu.

**EVALUATING THE EFFECTIVENESS OF PROVOSTS IN BUILDING A
STUDENT LEARNING ASSESSMENT-SUPPORTIVE ORGANIZATIONAL
CULTURE: A MULTIPLE-SITE EVALUATION WITHIN THE CALIFORNIA
STATE UNIVERSITY SYSTEM**

By

CHRIS PROCELLO

A dissertation submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy
University of San Diego

May 2008

Dissertation Committee

Fred Galloway, Ed.D., Chair
Lea Hubbard, Ph.D., Member
Carole Huston, Ph.D., Member

ABSTRACT

Although research on cultivating support for student learning assessment at the institutional level points to the necessary involvement of all campus stakeholders, researchers have commented on the particularly important role of institutional administrators. Most research on the role of administrators in building support for assessment to date has not, however, focused on provosts, even though they are critical because of their power to determine the internal allocation of institutional resources. To address this issue, this study used a 27-question, Likert-scale survey to estimate the extent to which provosts in the California State University system have been successful in building an assessment-supportive organizational culture on their respective campuses.

All presidents, provosts, associate provosts, deans, and associate deans in the 23-campus system were surveyed using eight Total Quality Management constructs. Based on the opinions of the 195 administrators that responded, provosts within the system were found to be more effective than not with an overall score of slightly more than seven on a ten-point scale (with ten as “very effective”). Provosts were rated as most effective in terms of “shared vision” and “involvement” and least effective in terms of “quality at the same cost” and “collaboration”, although the average scores on all eight of the constructs were fairly tightly bunched. In addition, multivariate analysis revealed that two measures of institutional size, total enrollment and the number of academic affairs administrators, as well as provostial tenure and percent of graduate students were helpful in explaining variation in overall effectiveness; in particular, higher enrollments were associated with greater effectiveness.

Beyond its significance at the institutional and university system levels, this study was important in that it explored the extent to which the collegiate student assessment movement has been institutionalized. However, the study's grounding in Total Quality Management was questioned by many respondents; further research might consider a different theoretical approach. Examining perceptions among different strata of academic affairs administrators could assist in this endeavor. Finally, future researchers might examine other large public university systems to begin painting a national picture of the effectiveness of provosts in building a student learning assessment-supportive culture.

ACKNOWLEDGEMENTS

First, Karin, my wife, I could not have completed this degree without you and your support (and editing!). I love you.

Thank you to everyone at the University of San Diego who offered support. My committee chair, Fred Galloway, was outstanding. Carole Huston and Lea Hubbard, my other committee members, were equally excellent and sources of insight and good (paradigmatic) debate. Cheryl Getz, my advisor for during the first couple of years of my program, was always helpful, as was Athena Perrakis. Thanks to all the Leadership faculty, whose humane approach to their students' lives is deeply appreciated, and to faculty who offered their feedback on this study, my survey instrument in particular. Michael Strait, Beth Dobkin, Tammy Tu, Beth Yemma, Kate Sheridan, Andre Branch, Oksana Shcherbak, and Richard Winn, thank you for your support and encouragement.

Finally, this study would not have been possible without my experiences at San Diego State University; thank you to Geoff Chase, Pat Patterson, Brock Allen, and Melody Kilcrease. I am indebted to Lorie Roth in the CSU Chancellor's Office, whose feedback on my work was pivotal, and to the many administrators within the CSU system who took the time to respond to my survey.

TABLE OF CONTENTS

	Page
ABSTRACT.....	iii
ACKNOWLEDGEMENTS	v
TABLE OF CONTENTS.....	vi
LIST OF TABLES	viii
LIST OF FIGURES	ix
 CHAPTER	
1. INTRODUCTION	
Background to the Evaluation.....	1
Statement of the Problem.....	8
Purpose of the Evaluation	10
Research Questions.....	13
 2. REVIEW OF THE LITERATURE	
Collegiate Student Learning Assessment.....	17
External Accountability.....	19
Internal Improvement.....	22
Student Learning Assessment and Institutional Administration in Higher Education.....	27
Organizational Culture.....	31
A Note on the Fundamental Role of Interpretivism.....	34
Organizational Culture as Inspired by Functionalism.....	36
Organizational Culture as Inspired by Critical Science....	43
Organizational Culture as Inspired by Postmodernism.....	51
Organizational Culture: Variable or Metaphor?.....	59
Evaluation Utilization.....	61
Evaluation Utilization and the Organizational Context.....	66
Evaluation and Quality.....	74
Total Quality and Evaluation.....	76
Total Quality, Evaluation, and Organizational Culture.....	81
Conclusion.....	96
 3. RESEARCH DESIGN AND METHODOLOGY	
Purpose.....	99
Scope, Focus, and Units of Analysis.....	100
Sampling Strategies.....	103
Analytical Approach, Data Collection, and Data Management	

and Analysis.....	104
Project Management, Logistics, and Survey Procedures.....	106
Delimitations and Limitations	
Research Methodology.....	107
Theoretical Frameworks.....	108
4. FINDINGS	
Exploratory Data Analysis.....	112
Research Analysis.....	117
Effectiveness by Institutional Characteristic	
Carnegie Classification Type.....	120
Total Enrollment.....	122
Number of Academic Programs.....	124
Length of Provost Tenure.....	125
Number of Academic Affairs Administrators.....	127
Percent of Graduate Students.....	129
Age of Institution.....	131
Effectiveness by Institutional Characteristic and Total Quality Construct.....	132
Research Analysis Summary.....	140
5. DISCUSSION	
Conclusions.....	142
Implications beyond the CSU.....	144
Recommendations for Further Research.....	146
REFERENCES.....	148
Appendix	
A. Dissertation Concept Map.....	179
B. Survey Cover Letter.....	181
C. Survey Instrument.....	183
D. List of Survey Statements as Connected to Total Quality Culture Construct.....	189

LIST OF TABLES

Table 1. Habermas' three cognitive interests.	p. 44
Table 2. Model of Total Quality Management Values and Beliefs	p. 83
Table 3. Total Quality Management Values and their Opposite	p. 85
Table 4. Principles of Good Practice for Assessing Student Learning Compared to the Basic Elements of Total Quality	p. 93
Table 5. Descriptive Statistics for Each Total Quality Construct	p. 115
Table 6. T-test Results for Each Total Quality Construct	p. 118
Table 7. Significance Scores by Institutional Characteristic and Total Quality Construct	p. 134
Table 8. Significance Scores for Total Enrollment Groups by Total Quality Construct	p. 135
Table 9. Significance Scores for Number of Academic Affairs Administrators Groups by Total Quality Construct	p. 136
Table 10. Significance Scores for Length of Provost Tenure Groups by Total Quality Construct	p. 137
Table 11. Significance Scores for Percent of Graduate Students Groups by Total Quality Construct	p. 137
Table 12. Regression Coefficients by Total Enrollment, Length of Provost Tenure, Number of Academic Affairs Administrators, and Percent of Graduate Students by Total Quality Construct and All Responses	p. 139

LIST OF FIGURES

Figure 1. Preskill and Torres' Theory of Developmental Evaluation	p. 72
Figure 2. Gap Chart Example	p. 90
Figure 3. Bar Graph of Campus Responses to Survey	p. 113
Figure 4. Histogram of All Responses	p. 114
Figure 5. Histogram of Responses to Total Quality Construct— Collaboration	p. 116
Figure 6. Histogram of Responses to Total Quality Construct—Quality at the Same Cost	p. 116
Figure 7. Gap Chart of Provost Effectiveness—All Responses	p. 119
Figure 8. Gap Chart of Provost Effectiveness—Carnegie Classification Type Groups	p. 121
Figure 9. Gap Chart of Provost Effectiveness—Total Enrollment Groups	p. 123
Figure 10. Gap Chart of Provost Effectiveness—Number of Academic Programs Groups	p. 125
Figure 11. Gap Chart of Provost Effectiveness—Length of Provost Tenure Groups	p. 126
Figure 12. Gap Chart of Provost Effectiveness—Number of Academic Affairs Administrator Groups	p. 128
Figure 13. Gap Chart of Provost Effectiveness—Percent of Graduate Students Groups	p. 130
Figure 14. Gap Chart of Provost Effectiveness— Age of Institutions Groups	p. 132

Evaluations can be threatening because, inherently, they create the possibility that negative findings about a particular program or organization will emerge. Like positive findings, these can be seen as a reflection of the work of particular groups or individuals. The objective of any evaluation is not to identify persons or groups who are performing poorly, but rather, for most program evaluations, it is to make known what otherwise would not be known for the purpose of continuous improvement.

Darlene Russ-Eft and Hallie Preskill, *Evaluation in Organizations*, 2001

Why do colleges seem so reluctant to make major changes in the way they conduct their educational programs? The reason is not that they are indifferent to the welfare of their students, and critics are wrong to suggest the contrary. Most college presidents and deans genuinely care about undergraduates and want to see them educated well. The great majority of professors enjoy their teaching, like their students, and devote much time to their classroom responsibilities. Yet enjoying teaching and caring about students do not necessarily bring a willingness to reexamine familiar practices and search for new methods that could serve the purpose better.

Derek Bok, *Our Underachieving Colleges*, 2006

CHAPTER 1: INTRODUCTION

Background to the Evaluation

Throughout the past twenty-five years, the student assessment movement has been one of the major reform initiatives in higher education (El-Khawas, 2002), impacting practice and policy at the national, state, and university system levels. And at each of these three levels, fundamental tensions within the movement have surfaced. Since its beginning, the collegiate student assessment movement has served two oftentimes conflicting purposes: external accountability and internal improvement. The resulting discord between these two purposes—seen as being “in fundamental opposition, one ‘wrong’ and the other ‘right’” (Ewell in Gray, 2002, p.63)—have shaped and continue to shape the movement by prompting research on how assessment can and should be used for both purposes (Palomba & Banta, 1999).

At the national level, assessment in the name of accountability emerged from voices outside the academy whose overall impact was new accountability demands and a stricter regulatory environment for higher education (Association of American Colleges and Universities, 2002). At this level, federal government initiatives have set the context and provided the initial impetus and continuing momentum for the student assessment movement (Peterson & Einarson, 1997). Informed by several landmark reports from the mid-eighties that decried the quality of higher education and demanded reform, the federal government has pressed for assessment since the beginning of the movement.¹ For example, in 1987, the federal government revised the guidelines for the Council on Postsecondary Education, which oversees all regional accreditation agencies, by

¹ It is important to mention the U.S. Department of Education’s 1983 *A Nation at Risk* here for although it focused on K-12 education it was central to setting the stage for the debate (and, eventually, the policy) about educational quality.

proposing that accreditation focus on student achievement (Nelson, 1991). The following year, the U.S. Department of Education established new criteria of institutional outcomes for recognition of all accrediting bodies, calling for a focus on “educational effectiveness” (Palomba & Banta, 1999; Wright, 2002), effectively changing the criteria for the distribution of federal funds to postsecondary institutions (Advisory Committee on Student Outcomes Assessment, 1989). In 1990, the National Education Goals Panel established the nation’s first objectives for collegiate learning and continued the call for the development of valid and reliable assessments (Ewell, 1993a, 2002) and around this year, federal funds, in particular the Fund for the Improvement of Postsecondary Education, began to support assessment scholarship (Banta, 1993; Banta, Lund, Black, & Oblander, 1996; El-Khawas, 2002; Ewell, 2002; Palomba & Banta, 1999). By 1993, the Government Performance and Results Act passed, which focused government activity on outcome-related results rather than on inputs or process, causing a shift of emphasis felt throughout the public and not-for-profit sectors (Donaldson & Scriven, 2003; Russ-Eft & Preskill, 2001). Finally, initiatives at the national level have spurred reform initiatives within state governments and regional accreditation associations, strengthening the force of the movement.

Assessment in the name of internal improvement, in contrast to assessment for external accountability, at the national level emerged from four traditions inside the academy: first, the tradition that examines collegiate student learning as an application of educational and developmental psychology; second, research on student retention; third, the burgeoning of program evaluation that was related to the larger movement toward “scientific management;” and fourth, the mastery- and competency-based learning

movement that began in elementary and secondary education (Ewell, 2002). All four traditions shaped practice and language in the early 1980s, leading to a 1984 U.S. Department of Education report entitled *Involvement in Learning: Realizing the Potential of American Higher Education* (Study Group on the Conditions of Excellence in American Higher Education, 1984) and the First National Conference on Assessment in Higher Education in the fall of 1985 (Ewell, 2002). Voices from within higher education dominated this report and conference and others like them soon to come. Of these voices, six are particularly important—those of Alexander Astin, Derek Bok, Richard Light, Ernest Boyer, Lee Shulman, and K. Patricia Cross—as these “idea champions” have captured national attention since the mid-1980s and have directed attention toward teaching and learning, making them subjects of scholarship (Lazerson, Wagener, & Shumanis, 2000). The reason for these reformers’ placement of student learning above other indicators of educational effectiveness is because of the mission of higher education: to improve student learning and personal development (Boyer Commission on Educating Undergraduates in the Research University, 1998; Ewell, 1989; Palomba & Banta, 1999; Pike, 2002). At the national level, as a result of the work of these six researchers and many others, research on assessment of student learning and personal development for the sake of improvement has blossomed over the past 25 years, leading to research at the classroom, course, program, general education, institutional, and university system levels.

At the state level, almost every state has enacted requirements to monitor the performance of all public agencies, program by program (Mark & Henry, 2004). One result of this trend toward performance-monitoring systems in higher education has been

a progressive increase in the number of states that have enacted student assessment initiatives since the mid-1980s: before 1982, no state required reporting on outcomes but by the mid-1990s virtually all states had assessment mandates for public colleges and universities (Banta, 1993; Chun, 2002; Ewell, 1993a, 2002; Lazerson, Wagner, & Shumanis, 2000; National Association of State Universities and Land-Grant Colleges, 2006; Peterson & Einarson, 1997; Peterson & Vaughan, 2002). States have pressed for greater evidence on the outcomes of collegiate study and some have imposed new requirements on colleges and universities, including annual reporting on institutional performance or a revamping of academic programs (El-Khawas, 2002). Accordingly, there has been an upward trend in the number of institutions reporting assessment activities. In 1987, 55% of institutions claimed that they had an assessment program; by 1993, this proportion had risen to 98% (El-Khawas, 1993; Ewell, 2002).

In California, there have been three major turning points in the history of collegiate student assessment. The first occurred in 1984 when the California State Legislature commissioned a review of the 1960 document, *A Master Plan for Higher Education in California* (Master Plan Survey Team, 1960), producing a 1987 report entitled *The Master Plan Renewed: Unity, Equity, Quality, and Efficiency in California Postsecondary Education* (Commission for the Review of the Master Plan for Higher Education, 1987). In this 1987 report, in a section on enhancing educational quality, the issue of quality in undergraduate instruction is addressed: "...campuses must establish clear curricular objectives, match curricular development to those objectives, and institutionalize systematic evaluation of program and individual curricular offerings" (Commission for the Review of the Master Plan for Higher Education, 1987, p.31). The

report also speaks to regional accrediting commissions, recommending that colleges and universities be held accountable for clear expectations for student learning and program assessment. The second turning point came in 1990, when the state moved past recommending to mandating. Largely due to the work of Assemblyman Tom Hayden, who had authored several bills proposing the establishment of mandatory outcomes assessment since 1986, the California State Education Code was amended (Higher Education Assessment Act of 1990) to mandate the use of assessment, thereby reinforcing the ideas of public accountability and student learning as the primary goal of higher education. The third turning point came in 1999 when the State Legislature called for the creation of a new master plan, resulting in *The California Master Plan for Education* in 2002. Reflecting advances in assessment research, the document argues for use of “authentic,” or direct, measures of student learning and a focus on general education that invites consensus on a common body of knowledge and skills; this new plan bolstered the case for assessment with emerging research to reaffirm student learning as the primary focus. These three turning points in California’s educational history reflect the development of collegiate student assessment policy at the state level, which has generally progressed from initiation to codification to reaffirmation of codification.

Because accreditation associations have a significant influence on institutional assessment efforts (Ewell, 1993a; Peterson & Einarson, 1997; Peterson & Vaughan, 2002; Welsh & Metcalf, 2003), a discussion of California’s history with assessment ought to include a discussion of the region’s accrediting organization. Formed in 1962 “to promote the welfare, interests, and development of education in the Western Region,” the

Western Association of Schools and Colleges (WASC) includes three accrediting commissions, one for K-12 schools, one for community and junior colleges, and a third for colleges and universities. The Accrediting Commission for Senior Colleges and Universities' chief goals are to promote institutional engagement with issues of educational effectiveness and student learning, to develop a culture of evidence that informs decision-making, and to foster active interchange among public and independent institutions (Western Association of Schools and Colleges, 2007). WASC has experienced two waves of reform around assessment. The first came in 1988 when, in revising its *Handbook of Accreditation*, a series of accreditation standards on institutional and program quality and effectiveness were introduced (Nelson, 1991; Wright, 2002). It was around this time, it is important to note, when the organization began using the phrases "a culture of evidence" and "a culture of inquiry," which imply that the use of assessment information should be a part of an institution's culture (Nelson, 1991; Wright, 2002). The second period of significant reform came in 2001. After years of study, WASC established a new framework for accreditation in 2001 that was organized around two "core commitments"—institutional capacity and educational effectiveness—and moved to a multiple-part review cycle (Wright, 2002). The rationale behind this split, in which the capacity review focuses on institutional policies, structures, and resources and the educational effectiveness review focuses on the institution's educational vision, its organization for learning, and evidence of student learning, is to spotlight the issue of educational effectiveness and student learning so that institutions do not revert to the old compliance-driven model (Wright, 2002). In this way, and through the institutional

review process, WASC encourages institutions to become learning organizations (Western Association of Schools and Colleges, 2001).

Similar to WASC, the California State University system has experienced two major waves of reform with regard to student learning assessment, one in the late 1980s and another close to the year 2000. After the first CSU Assessment Conference in 1986, held to identify the “essential characteristics” of then emerging assessment theories, practices, and their contexts, the CSU Chancellor created the Advisory Committee on Student Outcomes Assessment in 1987. Consisting of members from the Academic Senate, the California State Student Association, campus administration, alumni association, and the Chancellor’s Office, the committee consulted with a range of CSU stakeholders before publishing *Student Outcomes Assessment in the California State University* in 1989. This report proposed 12 “guiding principles” for the CSU in assessing student outcomes, the first stating that “the only legitimate purpose of assessing student outcomes is to improve teaching, learning, and academic advising at the institutional, course, program, and/or institutional level” (Advisory Committee on Student Outcomes Assessment, 1989, p.11). The second period of reform came nearly a decade later when in 1998 the CSU Board of Trustees endorsed the *Cornerstones Report*, a planning framework that called for a system-wide accountability process. The following year, the Trustees introduced the accountability process, beginning a continuous self-analysis of educational effectiveness and the dissemination of regular, cyclical reports to the public. These reports, submitted by each campus every other year and by the system annually, use quality of baccalaureate degree programs as the first of nine performance measures (California State University, 2007). Implementation of this process was phased: from

2000 to 2002 a narrative describing the processes for establishing and assessing student learning outcomes in general education and in the majors along with assuring that students are achieving core competencies for the degree was required. Starting in 2004, a narrative summarizing campus academic program reviews was required. Thus, within the CSU system, the conversation about student assessment originated at the same time as the beginning of the collegiate assessment movement, but full implementation of assessment policies has only recently been introduced.

Statement of the Problem

Research on cultivating support for student learning assessment at the institutional level points to the necessary involvement of campus stakeholders, including students, staff, faculty, and administrators, because all significant organizational changes require broad-based support if they are to be successful and sustainable (Banta, 2002; Boggs, 1999; Palomba, 1997; Palomba & Banta, 1999; Walvoord, 2004; Welsh & Metcalf, 2003; Young & Knight, 1993). However, many researchers have commented on the particularly important role of institutional administrators. They have pointed to their “legitimate power,” their role and position in the organization that grants them authority (French & Raven, 1960; Weber in Gerth & Mills, 1964; Merton, 1957; Salancik & Pfeffer, 1977; Wilson, 1989).

To be successful, assessment initiatives must be made a priority by institutional administrators as their power allows them to direct the behavior of others and mobilize symbolic-, material-, human-, and time-based resources toward specific ends (Banta, 1993; Bok, 2006; Gray, 1997; Light, 2001; Palomba & Banta, 1999; Watt, Drennen, Rodrigues, Menelly, & Wiegel, 1993). That is, institutional administrators can make an

explicit commitment of institutional resources to student assessment activities and, second, can link student assessment activities to the institution's internal resource allocation process (Peterson & Einarson, 1997; Peterson, Einarson, Augustine, & Vaughan, 1999). For assessment to become institutionalized, then, there must be effective leadership that has made an administrative commitment to, which includes adequate resources for, assessment on a campus (Banta, 2002; Banta, Lund, Black, & Oblander, 1996; Bok, 2006; Light, 2001; Peterson & Einarson, 1997; Peterson & Vaughan, 2002; Peterson, Einarson, Augustine, & Vaughan, 1999; Task Force on Teaching and Career Development, 2007). This commitment extends beyond the allocation of resources, however. One of the ten principles of good practice for assessing student learning in higher education states that "assessment is most effective when undertaken in an environment that is receptive, supportive, and enabling" (American Association for Higher Education, 1992; Banta, 1997; Banta, Lund, Black, & Oblander, 1996), which, for institutional administrators, implies an institution-wide approach to shaping campus climate and culture.

While strong support from the president and other senior administrators is viewed as crucial for an institution's success in student assessment (Peterson, Einarson, Augustine, & Vaughan, 1999), most case studies on this issue to date have focused on the president and not other senior administrators.² More common in the literature is a clumsy lumping together of all "institutional leaders" that assumes no role differentiation between these administrators, as if their responsibilities were the same. Few case studies

² For example, each of the case studies in *New Directions for Higher Education's* 1997 "The Campus-Level Impact of Assessment," many of the case studies in Banta's 1993 *Making A Difference* and Banta, Lund, Black, and Oblander's 1996 *Assessment in Practice*, and many of the campus profiles in the popular journal *Assessment Update* mention presidential leadership.

or other forms of research on institution-wide assessment initiatives mention the role of the provost and, if so, only in passing.³ This is surprising considering that, while the precise role of a provost varies from institution to institution, the title generally refers to the chief academic officer, responsible for oversight of all academic affairs and activities. The negligible amount of research on the role of the provost here becomes even more surprising given what is known about the relationship between academic affairs and assessment: At the institutional level, primary responsibility for student assessment is usually positioned in academic affairs, executive responsibility for student assessment planning is most often vested in academic affairs administrators, assessment plans and policies are usually subject to approval from the chief academic officer, and operating responsibility for day-to-day student assessment activities is likely to be given to an academic affairs administrator who reports directly to the chief academic officer (Peterson, Einarson, Augustine, & Vaughan, 1999). *Thus, considering the importance of assessment to the provost's office, combined with a provost's power to control a campus' internal resource allocation process, a problem lies in the lack of research on the specific role of the provost in building institutional support for student learning assessment.*

Purpose of the Evaluation

Of the theoretical approaches to understanding the principle “assessment is most effective when undertaken in an environment that is receptive, supportive, and enabling,” the organizational culture approach has received the most attention in the collegiate student learning assessment literature. This is because, ultimately, cultural changes at the

³ The greatest exception to this might be the National Center for Postsecondary Improvement study in which their survey was directed to the chief academic officer on each campus. However, the survey instructed this person to send the survey to “the appropriate person or persons” for completion, thereby implying delegation (Peterson & Augustine, 2000; Peterson, Einarson, Augustine, & Vaughan, 1999).

organizational level are perceived to be the only way to institutionalize assessment initiatives. The goal of institutional administrators, therefore, ought to be shaping campus culture toward supporting student learning assessment. Richard Light captures this sentiment well:

At the beginning of this book, I quoted a dean from another university who said the strategy at his college was to admit a talented group of students and then just “get out of their way.” It seems clear to me from the dozens of anecdotes and examples in this book that campus leaders should...do exactly the opposite of that dean’s recommendation. They should make a thoughtful, evidence-based, purposeful effort to get *in* each student’s way. In fact, shaping a certain kind of campus culture may be the biggest contribution campus leaders can make. (2001, p.209)

It is because of the prominence of this sentiment in the assessment literature that I propose a focus on organizational culture. If “ultimately, it’s about the culture” in relation to educational effectiveness (Kuh, Kinzie, Schuh, & Whitt, 2005a) and changing campus culture ought to be the ultimate goal, then we must work toward organizational improvement through changing organizational culture.

The primary purpose of this evaluation, then, is organizational improvement. This study operates with the general assumptions and objectives of organizational development theory and research defining organizational development as “a system-wide and values-based collaborative process of applying behavioral science knowledge to the adaptive development, improvement and reinforcement of such organizational features as the strategies, structures, processes, people, and cultures that lead to organizational

effectiveness” (Anderson & Anderson, 2001, p.xxi). Organizational development, as this definition suggests, is an approach to organization change based on applied behavioral science knowledge in which interventions, which typically focus on values, are implemented to foster organization-wide improvement and effectiveness (Burke, 2002; Cummings & Worley, 2004; Rothwell & Sullivan, 2005). Organizational development ultimately seeks to bring about organizational change, in particular what Anderson and Anderson (2001) call “developmental change,” which represents the *improvement* of an existing skill, method, performance standard, or condition that does not meet current or future needs.

The primary purpose of organizational improvement leads to the task of evaluating effectiveness. The motive to determine and evaluate effectiveness, however, must be linked to a governing theory or approach for systematic inquiry. This study operates with Michael Patton’s definition of evaluation:

...the systematic collection of information about the activities, characteristics, and outcomes of programs to make judgments about the program, improve program effectiveness, and/or inform decisions about future programming...evaluation [is] done for and with specific, intended primary users for specific, intended uses. (2002, p.10)

Besides his emphasis on improvement, Patton’s focus on “intended primary users,” those who are likely to be the real users of the evaluation for intended purposes, fits this study well. In this case, CSU provosts are the intended users and organizational improvement is the intended use. If provosts, as stewards of an institution’s academic affairs and activities, hold “legitimate power” and can control a campus’ internal resource allocation

process, then they certainly ought to be the intended primary users. And if provosts within the CSU system carry a special responsibility regarding student learning as a result of CSU's mission centering on teaching and learning, then they certainly ought to be interested in improving their campus in the area of student learning assessment.⁴

Research Questions

The task of evaluating effectiveness with the purpose of improvement, the need for empirically- and theoretically-based research on the role of provosts in supporting assessment, and the mission of the CSU leads to a formative evaluation of the effectiveness of provosts in building a culture supportive of student learning assessment within the CSU system. A formative evaluation, according to Patton (2002), aims to make improvements to a specific set of activities at a specific time and place and with a specific group, and should not set out to criticize users but to help them. Thus, the overarching question here is based in improvement with a focus on provosts and student learning assessment within the CSU:

- *How can CSU provosts be more effective in building a student learning assessment-supportive culture at their campuses?*

Of the formative evaluation types, this question lends itself to an implementation evaluation, which in monitoring the fidelity of program delivery or policy implementation, relies on distinguishing between the actual and the ideal (Love, 1991; Patton, 1997, 2002). This study's theoretical orientation leans toward positivist and realist

⁴ Improvement efforts must be linked to an organization's mission (Russ-Eft and Preskill, 2001). And the mission of the California State University system, as related in the California Master Plan for Higher Education, centers on teaching and learning: "The state colleges shall have as their primary function the provision of instruction in the liberal arts and sciences and in the professions and applied fields..." (Master Plan Survey Team, 1960, p.2). This study's focus on student learning is rooted not only in the national conversation about assessment but also in the specific mission of the CSU.

approaches, operating with the presumption that there is a “real world” with verifiable patterns that can be observed and predicted (Patton, 2002). The “patterns” here are evident in the search for “value gaps” in perception between provosts’ ideal values regarding building an organizational culture supportive of assessment and the provosts’ values-in-use. That is, the goal of this study is evaluative in that it aims to discover “gaps” where provosts can be more effective in the interest of organizational improvement assuming that, in the end, “bridging the gap between what is happening and what is possible is what change management is all about” (Pascale & Sternin, 2005, p.73).

It is important to note that the history of this “value gap” approach began with Argyris and Schön’s 1974 *Theory in Practice*. In a discussion of theories of action, Argyris and Schön make the distinction between espoused theory, the words we use to convey what we do or what we would like others to think we do, and theory-in-use, theories that are implicit in what we do as practitioners and managers. Related to this is the much older “is-ought distinction” and the distinction between facts and values; Argyris and Schön’s theories of action are connected to the older philosophical debate about making descriptive statements about what “is” (facts) and prescriptive statements about what “ought” to be (values). This distinction, in Argyris and Schön’s terms, has been used by numerous researchers.⁵ The “espoused” position represents the ideal, in this case the “ideal culture profile” for a university supportive of student learning assessment; the question becomes one of measuring deviations from the ideal. This evaluation’s subsidiary questions, therefore, are:

⁵ For example, see Barr and Tagg (1995), Bright and Cooper (1993), Buch and Wetzel (2001), Burns (1994), Chang (1996), Detert, Schroeder, and Cudeck (2003), Kilmann, Saxton, Serpa, & Associates (1985), and Kuh, Kinzie, Schuh, & Whitt (2005a, 2005b).

- *What, if any, are the gaps between provosts' values-in-use (the "is") and ideal values (the "ought") regarding building a student learning assessment-supportive culture?*
- *What are the differences and similarities in "value gaps" among CSU campuses regarding provosts' building a student learning assessment-supportive culture?*

CHAPTER 2: REVIEW OF THE LITERATURE

This review integrates areas of three bodies of literature—collegiate student learning assessment, organizational culture, and evaluation—to propose a formative evaluation of provosts' effectiveness in building a student learning assessment-supportive organizational culture at their campuses. My aim here is to discuss the history, scope, and components of these literatures, the areas within them that are most relevant to my study, and significant gaps in these literatures my research can fill. Justification for this evaluation rests on it being relevant to organizational improvement. At the institutional level, it would add to research on the role of institutional administrators in organizational improvement efforts regarding student learning.

There are four main points to my argument. First, although research has found assessment at the institutional level to be tightly connected to academic affairs and provosts to have power in a campus' internal resource allocation process, there exist no empirical studies that attempt to develop a theoretical understanding of the specific role of the provost in building institutional support for student learning assessment. Second, the organizational culture literature is expansive but clustered into three paradigmatic groups: research inspired by functionalism, by critical theory, and by postmodernism. This study, operating with the assumption that organizational culture is a variable, bases its approach to organizational culture in the functionalist tradition. Third and fourth, evaluation research that integrates the organizational context is underdeveloped and weaving in literature on contemporary management strategies, particularly the principles of Total Quality, would help with its development. My study, then, addresses gaps in

three bodies of literature while simultaneously bringing areas of these literatures together in the interest of organizational improvement through educational reform.

Collegiate Student Learning Assessment

Since its beginning more than 20 years ago, the student assessment movement within U.S. higher education has served two oftentimes conflicting purposes: external accountability and internal improvement. Conflicting political and intellectual traditions characterize this movement with, generally, the purpose of assessment seen as for accountability *or* improvement (Banta, 1993; Dowd, 2005; Ewell, 2002; Gray, 2002; Johnson, McCormick, Prus, & Rogers, 1993). That is, these two purposes have traditionally been seen as being “in fundamental opposition, one ‘wrong’ and the other ‘right’” (Ewell in Gray, 2002, p.63). The resulting tensions between these two purposes have and continue to shape the student assessment movement by prompting research on how assessment can and should be used for both purposes (Palomba & Banta, 1999).

Assessment in the name of internal improvement emerged from four traditions: first, the tradition that examines collegiate student learning as an application of educational and developmental psychology; second, research on student retention; third, the burgeoning of program evaluation that was related to the larger movement toward “scientific management;” fourth, the mastery- and competency-based learning movement that began in elementary and secondary education (Ewell, 2002). All four traditions shaped practice and language in the early 1980s, leading to a 1984 U.S. Department of Education report entitled *Involvement in Learning: Realizing the Potential of American Higher Education* (Study Group on the Conditions of Excellence in American Higher Education, 1984) and the First National Conference on Assessment in Higher Education

in the fall of 1985 (Ewell, 2002). Voices from within higher education dominated this report and conference and others like them soon to come.

Assessment in the name of accountability emerged from a set of voices outside the academy, primarily in state legislatures and the U.S. Senate (Ewell, 2002). Calls for greater accountability were a byproduct of widespread dissatisfaction with education in the mid-eighties, best symbolized by the U.S. Department of Education's 1983 *A Nation at Risk* (Bok, 2006). This dissatisfaction also led to the National Governors' Association's *Time for Results* (1986) and the Education Commission of the States' *Transforming the State Role in Improving Undergraduate Education* (1986). These reports (1) drew attention to the performance of public institutions, (2) argued for a more proactive role of state authorities in higher education, thereby forcing colleges and universities to develop comprehensive programs to measure student learning, and (3) ultimately, led to states adopting assessment mandates (Advisory Committee on Student Outcomes Assessment, 1989; Ewell, 1993a, 2002; Wright, 2002). The overall impact of these external voices was new accountability demands and a stricter regulatory environment for higher education (Association of American Colleges and Universities, 2002).

Thus, the student assessment movement began with a "dichotomy of purpose"—external accountability versus internal concerns for institutional improvement (Ewell, 2002; Peterson & Vaughan, 2002; Pike, 2000). To best understand each purpose, the two will be discussed separately here, beginning with demands for accountability.

External Accountability

In 1996, Stanford University's National Center for Postsecondary Improvement, sponsored by the U.S. Department of Education, commissioned a five-year project that provided the first comprehensive picture of institutional definitions of, approaches to, support for, and uses and impacts of student assessment in postsecondary education (Peterson & Einarson, 1997; Peterson, Einarson, Augustine, & Vaughan, 1999).⁶ One aspect of the project examined influences on student assessment from the external environment and found these influences to have an important role in initiating and shaping assessment efforts (Peterson & Einarson, 1997). In particular, they found the major external influences to be (1) the federal government, (2) state governments, and (3) regional accreditation associations (Peterson & Vaughan, 2002), a finding that has been supported by others (Ewell, 1993a; National Association of State Universities and Land-Grant Colleges, 2006; Welsh & Metcalf, 2003). These three constituents external to higher education are discussed below.

Generally, federal government initiatives have set the context and provided the initial impetus and continuing momentum for the student assessment movement (Peterson & Einarson, 1997). The discussion about student assessment at the federal level has been informed by several landmark reports from the mid-eighties that decried the quality of higher education, including *Involvement in Learning* (Study Group on the Conditions of Excellence in American Higher Education, 1984), *To Reclaim a Legacy* (Bennett, 1984),

⁶ The first in-depth study of the status of higher education's assessment movement was completed in 1990 (Johnson, Prus, Anderson, & El-Khawas, 1994). In this study, a survey was sent to a stratified sample of 455 colleges and universities. The National Center for Postsecondary Improvement study, in contrast, sent a survey to all 2,524 public and private postsecondary institutions recognized by the U.S. Office of Education that offer undergraduate programs at the associate or baccalaureate degree level (Peterson, Einarson, Augustine, & Vaughan, 1999).

Access to Quality Undergraduate Education (Southern Regional Education Board, 1985), and *Integrity in the College Curriculum* (Association of American Colleges, 1985) (Advisory Committee on Student Outcomes Assessment, 1989; Peterson & Einarson, 1997). Taken together, these reports criticized the state of baccalaureate education and demanded reform (Wright, 2002). And reform followed: in 1987, the federal government revised the guidelines for the Council on Postsecondary Education, which oversees all regional accreditation agencies, by proposing that accreditation focus on student achievement (Nelson, 1991). The following year, the U.S. Department of Education established new criteria of institutional outcomes for recognition of all accrediting bodies, calling for a focus on “educational effectiveness” (Palomba & Banta, 1999; Wright, 2002), effectively changing the criteria for the distribution of federal funds to postsecondary institutions (Advisory Committee on Student Outcomes Assessment, 1989). In 1990, the National Education Goals Panel established the nation’s first objectives for collegiate learning and continued the call for the development of valid and reliable assessments (Ewell, 1993a, 2002) and around this year, federal funds, in particular the Fund for the Improvement of Postsecondary Education, began to support assessment scholarship (Banta, 1993; Banta, Lund, Black, & Oblander, 1996; Bok, 2006; El-Khawas, 2002; Ewell, 2002; Palomba & Banta, 1999). By 1993, the Government Performance and Results Act, which focused government activity on outcome-related results rather than on inputs or process, passed causing a shift of emphasis felt throughout the public and not-for-profit sectors (Donaldson & Scriven, 2003; Russ-Eft & Preskill, 2001).⁷

⁷ The debate about the means by which colleges and universities should be held accountable for educating their students, and the role of assessment therein, has only intensified since the early 1990s (Dowd, 2005).

State governments' assessment initiatives have paralleled those at the federal level, and have had more of a direct impact (Bok, 2006; Peterson & Einarson, 1997). Almost every state has enacted requirements to monitor the performance of all public agencies, program by program (Mark & Henry, 2004). One result of this trend toward performance-monitoring systems in higher education has been a progressive increase in the number of states that enacted student assessment initiatives since the mid-1980s: before 1982, no state required reporting on outcomes but by the mid-1990's, virtually all states had assessment mandates for public colleges and universities (Banta, 1993; Chun, 2002; Ewell, 1993a, 2002; Peterson & Einarson, 1997; Peterson & Vaughan, 2002; Lazerson, Wagner, & Shumanis, 2000; National Association of State Universities and Land-Grant Colleges, 2006). States have pressed for greater evidence on the outcomes of collegiate study and some have imposed new requirements on colleges and universities, including annual reporting on institutional performance or a revamping of academic programs (El-Khawas, 2002). Accordingly, there has been an upward trend in the number of institutions reporting assessment activities. In 1987, 55% of institutions claimed that they had an assessment program; by 1993, this proportion had risen to 98% (El-Khawas, 1993; Ewell, 2002).

Paralleling the states, all regional accreditation associations now require assessment (Angelo, 2002; El-Khawas, 2002). And like the states, accreditation associations have experienced waves of reform. Between 1984 and 1992, for example, all

This is evidenced by the continuous flow of reports from higher education associations on this topic (see Association of American Colleges and Universities, 2004, 2005; Business-Higher Education Forum, 2003, 2004; The Council for Higher Education Accreditation, 2003; National Center for Public Policy and Higher Education, 2005a, 2005b, 2005c; State Higher Education Executive Officers, 2005). Within this debate, wrangling over the subject of state- or nationally-imposed standardized tests and assessments for all college graduates has caught the attention of many in the academy.

six regional accrediting associations established or revised policies, standards, or statements related to student learning assessment (Ewell, 1993a; Peterson & Einarson, 1997; Wright, 2002). And since 1992, a second wave of accreditation activity around assessment has gained momentum (Ewell, 2002; Wright, 2002), reflecting these associations' increasing focus on what students actually learn versus the traditional focus on what the institution offers (Ewell, 1999; Norris, 2006). Accreditation associations appear to have a significant influence on institutional assessment efforts (Ewell, 1993a; Peterson & Einarson, 1997; Peterson & Vaughan, 2002; Welsh & Metcalf, 2003). In the National Center for Postsecondary Improvement study, preparing for accreditation was the highest-rated institutional purpose for engaging in student assessment (Peterson & Vaughan, 2002); in many cases, accrediting associations have replaced the state as the primary external stimulus for assessment (Ewell, 1993a). "Arguably,...the single most powerful contributor to assessment's staying power has been its championing by regional and professional accreditors" (Wright, 2002, p.253), who have drawn on their power as the source of accreditation status and eligibility for federal and state funds (Peterson & Einarson, 1997). The collegiate student assessment movement has made accreditation associations more relevant, giving them new purpose and increasing their clout (Wright, 2002), although they have been careful in using their newfound power by allowing institutions wide flexibility in proceeding through the accreditation process (Ewell, 2002; Palomba & Banta, 1999; Wergin, 2005).

Internal Improvement

Assessment in the name of internal improvement is the second major force behind the student assessment movement in higher education. This aspect of the movement, as I

have described, is rooted in four traditions and in the mid-1980s was couched in terms of “educational effectiveness” where student assessment should not be an end in itself but used for educational and institutional improvement (Peterson & Vaughan, 2002). It was not this coalescence of traditions alone that fueled the movement, however; “idea champions” are important to any reform movement (El-Khawas, 2002). Six higher education reformers’ visions of valuing student learning for its own sake are particularly important (Lazerson, Wagener, & Shumanis, 2000). First, Alexander Astin of U.C. Los Angeles’ claim that most measures of institutional quality were flawed because they said nothing about student learning and that a college’s quality and prestige should be measured by the “value added” to its students’ learning (Astin, 1993) heavily influenced *Involvement in Learning* and the movement thereafter. Second, Derek Bok and Richard Light of Harvard University took up research on learning environments, agreeing that colleges and universities had to demonstrate that they genuinely added to students’ knowledge (see Bok, 2006), leading to Light’s examination of Harvard’s learning environment in the Harvard Assessment Seminars (Light, 1990, 1992, 2001) and in his *Making the Most of College* (2001). Third, Ernest Boyer of the Carnegie Foundation for the Advancement of Teaching redefined “scholarship” to incorporate a wide variety of faculty work, arguing for the scholarship of teaching in *Scholarship Reconsidered* (1990). Fourth, Lee Shulman of Stanford University and the Carnegie Foundation for the Advancement of Teaching continued Boyer’s efforts to create a scholarship of teaching by connecting teaching to the disciplines, arguing that good teaching enacted teachers’ understanding of their disciplines. Fifth, K. Patricia Cross of U.C. Berkeley and the American Association of Higher Education redefined the concept of “classroom

research” in 1985, leading to the development of concrete ways to change collegiate classrooms toward improved student learning in *Classroom Assessment Techniques* (with Thomas Angelo, 1988, 1993) (which includes the popular “one-minute paper” [Light, 2001]) and *Classroom Research* (with Mimi Harris Steadman, 1996). These “idea champions” have captured national attention since the mid-1980s and have directed that attention toward teaching and learning, making these two areas subjects of scholarship.

The reason for these reformers’ placement of student learning above other indicators of educational effectiveness is because of the mission of higher education: student learning and personal development (Association of American Colleges and Universities, 2007; Boyer Commission on Educating Undergraduates in the Research University, 1998; Ewell, 1989; Palomba & Banta, 1999; Pike, 2002). “Learning” is typically defined in terms of general education or a discipline—specific knowledge, skills, and values fundamental to a liberal education or a discipline—and “development” usually implies professional and career development as well as personal growth and development (Gray, 2002). Because of the primacy of student learning, the general term “assessment” in higher education is assumed to be a truncation of “student learning assessment”: “After fifteen years of a wide-spread assessment ‘movement,’ there is now broad agreement among accrediting agencies, disciplinary and professional associations, administrators, and faculty opinion leaders that improving student learning is (or should be) the primary goal of assessment” (Angelo, 2002, p.188). To organize my discussion of student learning assessment for the sake of internal improvement, I discuss below two internal purposes for institutional engagement in student assessment. First, assessment in the interest of student achievement and, second, assessment in the interest of improving

institutions and academic programs are discussed. These two (internal) purposes ranked as second and third in institutional purposes for engaging in student assessment, after the first-ranked accreditation purpose, in the National Center for Postsecondary Improvement study (Peterson & Vaughan, 2002).

Student learning assessment in the interest of student achievement is rooted in the student outcomes tradition, which has focused on classroom- and course-level assessment and generally been directed toward faculty and, ultimately, individual students (Ewell, 1989; Palomba & Banta, 1999). The essence of this literature is captured by Barr and Tagg's "learning paradigm" (1995). In "From Teaching to Learning—A New Paradigm for Undergraduate Education," Barr and Tagg argue for a shift from an "instruction paradigm," where colleges and universities exist to provide instruction and mistake a means (instruction) for an end (learning), to a "learning paradigm," where colleges and universities exist to produce learning. That is, they argue for a shift away from faculty, traditional research, and instruction and toward students, scholarship, and learning (Angelo, 2002; Boggs, 1999; O'Banion, 1997, 1999; Tagg, 2003). Restructuring institutions to produce better learning would include learning assessment and information systems at all levels—the "systematic gathering of information for sustained improvement" (Light, 2001, p.224)—making assessment one key data-collection methodology to measure student learning (Norris, 2006). For Barr and Tagg, "the place to start the assessment of learning outcomes is in the conventional classroom; from there, let the practice grow to the program and institutional levels" (1995, p.14). At the classroom level, no book has had more of an impact than Angelo and Cross' *Classroom Assessment Techniques* (1988, 1993), a toolkit for faculty that brought the "learning paradigm" into

the classroom by involving teachers in ongoing, systematic study of teaching and learning within their own classrooms in order to understand and improve them (Angelo, 2002; Obler, Slark, & Umbdenstock, 1993). It was these tools, along with other popular ones like Chickering and Gamson's seven principles for good practices in undergraduate education (1987) and Bloom's taxonomy (1956), that were used by the emerging centers for "teaching and learning" and "academic excellence" in the late 1980s (Borden, 2002). These centers devoted to faculty development needed to be created as most faculty had no formal training in assessment or evaluation (Banta, 2002; Banta, Lund, Black, & Oblander, 1996).

Student learning assessment in the interest of improving institutions and academic programs is rooted in the institutional effectiveness tradition, typically employs program evaluation techniques, and has generally been directed toward academic administrators (Ewell, 1989). In the most general terms, it has to do with the ways an institution allocates its human and other resources and organizes learning opportunities and services to foster student learning (Kuh, Kinzie, Schuh, & Whitt, 2005a). This strand of assessment has led to assessment at the departmental (the majors), general education, institutional, and university system levels (Amiran, Schilling, & Schilling, 1993; Banta, Lund, Black, & Oblander, 1996; Black & Kline, 2002; Borden, 2002; Johnson, McCormick, Prus, & Rogers, 1993; Palomba, 2002; Palomba & Banta, 1999; Peterson, Einarson, Augustine, & Vaughan, 1999). Work along this strand has led to, for example, the redefinition of roles of offices within institutions, such as offices of institutional research and undergraduate studies, and to the use of popular assessment tools, such as the National Survey of Student Engagement, the College Student Experiences

Questionnaire, and the Collegiate Learning Assessment. Assessment at the institutional and program levels has been the subject of reports since the beginning of the assessment movement. *Involvement in Learning*, in fact, stressed the importance of feedback not only in the classroom but at the institutional level where colleges and universities could “learn” from feedback on their own performance (Ewell, 1989, 2002). Other landmark documents like *To Reclaim a Legacy, Integrity in the College Curriculum*, and *Greater Expectations* have joined the argument for curricular reform, specifically coherency and connectedness in curricular experiences, and for ongoing monitoring of student learning and development (Ewell, 2002, 2004).

Student Learning Assessment and Institutional Administration in Higher Education

Research on cultivating support for student learning assessment at the institutional level points to the necessary involvement of campus stakeholders, including students, staff, faculty, and administrators, because all significant organizational changes require broad-based support if they are to be successful and sustainable (Banta, 2002; Boggs, 1999; Palomba, 1997; Palomba & Banta, 1999; Welsh & Metcalf, 2003; Young & Knight, 1993). However, many researchers have commented on the particularly important role of institutional administrators. They have pointed to their “legitimate power,” their role and position in the organization that grants them authority (French & Raven, 1960; Weber in Gerth & Mills, 1964; Merton, 1957; Salancik & Pfeffer, 1977; Wilson, 1989).

To be successful, assessment initiatives must be made a priority by institutional administrators as their power allows this group to direct the behavior of others and mobilize symbolic-, material-, human-, and time-based resources toward specific ends

(Banta, 1993; Bok, 2006; Gray, 1997; Light, 2001; Palomba & Banta, 1999; Watt, Drennen, Rodrigues, Menelly, & Wiegel, 1993). For example, institutional administrators can communicate to a campus, through sanctioned channels, that attention to student learning is the priority, thus employing symbolic means to express their commitment to the issue (Ewell, 1989; Kuh, Kinzie, Schuh, & Whitt, 2005a; Light, 2001; Magruder, McManis & Young, 1997; Palomba & Banta, 1999; Peterson & Augustine, 2000; Young & Knight, 1993). They can commit “symbolic resources” in that they can use their power to influence what issues get attention (Smircich & Morgan, 1982). Institutional administrators can also increase the budget of the Office of Institutional Research to provide financial support for new assessment activities (Williford, 1997), staff an Office of Assessment (Palomba & Banta, 1999) or a Center for Teaching and Learning in order to provide professional development on assessment (Peterson & Augustine, 2000), create a university-wide Assessment Committee (Palomba & Banta, 1999), or link assessment to faculty evaluation and rewards (Peterson & Augustine, 2000). Time can also be a resource; senior administrators can allow adequate time for faculty or others to understand, accept, and carry out assessment (Banta, 2002; Bok, 2006; Garvin, 1993; Palomba & Banta, 1999). The point here is institutional administrators can make an explicit commitment of institutional resources to student assessment activities and, second, can link student assessment activities to the institution’s internal resource allocation process (Peterson & Einarson, 1997; Peterson, Einarson, Augustine, & Vaughan, 1999).

For assessment to become institutionalized, then, there must be effective leadership that has made an administrative commitment to, which includes adequate

resources for, assessment on a campus (Banta, 2002; Banta, Lund, Black, & Oblander, 1996; Bok, 2006; Light, 2001; Peterson & Einarson, 1997; Peterson, Einarson, Augustine, & Vaughan, 1999). This commitment extends beyond the allocation of resources, however. One of the ten principles of good practice for assessing student learning in higher education is “assessment is most effective when undertaken in an environment that is receptive, supportive, and enabling” (American Association for Higher Education, 1992; Banta, 1997; Banta, Lund, Black, & Oblander, 1996), which, for institutional administrators, implies an institution-wide approach to shaping campus climate and culture. While strong support from the president and other senior administrators is viewed as crucial for an institution’s success in student assessment (Peterson, Einarson, Augustine, & Vaughan, 1999), most case studies on this issue to date have focused on the president and not other senior administrators. More common in the literature is a clumsy lumping together of all “institutional leaders” which assumes no role differentiation between these administrators, as if their responsibilities were the same.

Few case studies or other forms of research on institution-wide assessment initiatives mention the role of the provost and, if so, only in passing. This is surprising considering that, while the precise role of a provost varies from institution to institution, the title generally refers to the chief academic officer, responsible for oversight of all academic affairs and activities. One set of authors described the provost’s role this way, “In some ways, the Provost is the most important internal leadership position in a university, since in most universities, the President’s job is outwardly-focused, dealing mostly with the public and with regents or trustees. In contrast, the Provost is the ‘inside’

person in charge of academic affairs” (Eisenberg, Murphy, & Andrews, 1998, p.6). The negligible amount of research on the role of the provost here becomes even more surprising given what is known about the relationship between academic affairs and assessment: At the institutional level, primary responsibility for student assessment is usually positioned in academic affairs, executive responsibility for student assessment planning is most often vested in academic affairs administrators, assessment plans and policies are usually subject to approval from the chief academic officer, and operating responsibility for day-to-day student assessment activities is likely to be given to an academic affairs administrator who reports directly to the chief academic officer (Peterson, Einarson, Augustine, & Vaughan, 1999). Thus, considering the importance of assessment to the provost’s office, combined with a provost’s power to control a campus’ internal resource allocation process, a problem lies in the lack of research on the specific role of the provost in building institutional support for student learning assessment.

This problem of a lack of research on the role of the provost is further complicated by almost all of the collegiate student assessment literature at the institutional level being primarily descriptive or prescriptive in nature. Most of this literature consists of descriptions of student assessment practices at single institutions or prescriptive guidelines for how institutions should support student assessment efforts (Peterson, Einarson, Augustine, & Vaughan, 1999). Descriptions of the roles of institutional administrators are almost always based on descriptions of successful student assessment approaches at various campuses—not empirically-based, much less evaluation-, conceptually-, or theoretically-based (Peterson, Einarson, Augustine, & Vaughan, 1999). There are almost no systematic empirically-based examinations of the

relationship between different forms of leadership support for collegiate student assessment and levels of institutional support for assessment (Peterson & Einarson, 1997; Peterson, Einarson, Augustine, & Vaughan, 1999). Instead, there are uncritical single-institution “case studies” written by administrative assessment champions. *The literature on collegiate student assessment at the institutional level, specially that on how institutional administrators can build support for student learning assessment, has stalled at descriptive and prescriptive studies. There is a need for research on the role of the provost in building institutional support for student learning assessment that is rooted in empirical evidence and attempts to develop a theoretical understanding of this issue.*

Organizational Culture

The implications of the principle “assessment is most effective when undertaken in an environment that is receptive, supportive, and enabling” are myriad (American Association for Higher Education, 1992; Banta, 1997; Banta, Lund, Black, & Oblander, 1996). If these implications are approached at the organizational level and in theoretical ways, this number declines, but not by much. Organizational theorists have developed approaches to examine an organization’s internal and external environment, strategy and goals, use of technology, decision making, power and political realities, conflicts and contradictions, controlling processes and ideology, ability to learn and change, and climate and culture, among other approaches (Hatch, 1997). Of these theoretical approaches, however, the organizational culture approach has received the most attention in the collegiate student learning assessment literature.⁸ This is because, ultimately,

⁸ See Andersen, Gardner, & Kuh, 2006; Association of America Colleges and Universities, 2002; Belcher, Malmberg, & Parkman, 2002; Bennino & Harris, 2005; Fultz & Wong, 2001; Gordon, 1999; Grackin, 2005; Gray, 2002; Kerr, 2002; Konidari & Abernot, 2006; Kuh, Kinzie, Schuh, & Whitt, 2005a, 2005b; Light, 2001; Taylor, 2002; Wright, 2002; and general literature from all regional accreditation associations.

cultural changes at the organizational level are perceived to be the only way to institutionalize assessment initiatives. Institutional administrators' goal, therefore, ought to be shaping campus culture toward supporting student learning assessment. Richard Light captures this sentiment well:

At the beginning of this book, I quoted a dean from another university who said the strategy at his college was to admit a talented group of students and then just 'get out of their way.' It seems clear to me from the dozens of anecdotes and examples in this book that campus leaders should...do exactly the opposite of that dean's recommendation. They should make a thoughtful, evidence-based, purposeful effort to get *in* each student's way. In fact, shaping a certain kind of campus culture may be the biggest contribution campus leaders can make. (Light, 2001, p.209)

It is because of the prominence of this sentiment in the assessment literature that I move to a discussion of organizational culture. If changing campus culture is the ultimate goal, then we must better understand organizational culture.

Over the course of its development, organizational culture research has drawn from different paradigmatic traditions (Hatch, 1995, 1997; Martin, Frost, & O'Neill, 2004; Ouchi & Wilkins, 1985; Reichers & Schneider, 1990), and this section will use major traditions in the social sciences to organize this literature. The bulk of research on organizational culture can be formed into three groups: (1) research operating with functionalist assumptions, (2) research based on the tenets of critical science, and (3)

research rooted in postmodernism.⁹ While there are authors who borrow concepts from two or more of the traditions or others who integrate all of these traditions, the majority of organizational culture research fits into one of these three groups. A second point to be made about this grouping is that many organizational culture researchers have utilized ideas from these traditions without making the tradition itself the foundation of their work. For example, some research uses ideas from critical science to describe subcultures, but do not move beyond this point towards the emancipation of oppressed groups that constitute a subculture. For this reason, subheadings read organizational culture as “inspired by” (i.e., drawing upon the “spirit,” not necessarily the details of) the traditions.

Each section gives background on the paradigmatic tradition and then defines organizational culture from the tradition’s perspective and elaborates on the implications of this definition. This is followed by a discussion of culture and, first, organizational effectiveness and second, “interests,” or patterns that focus on the well-being of specific individuals or, more likely, groups of individuals (Connell & Nord, 1996). Organizational culture inspired by the functionalist tradition defines culture as that which is shared, treats culture as a variable that can be manipulated for effectiveness, and tends to serve managerial interests. Organizational culture inspired by critical science defines culture as that which is shared within the subcultural boundary, is skeptical of attempts to manipulate culture, and tends to serve the interests of groups that lack the power and status of management. Organizational culture inspired by postmodernism places

⁹ This section relies heavily on Joanne Martin’s “three perspectives” of organizational culture (Martin, 1992, 2002; Martin & Meyerson, 1988; Meyerson & Martin, 1987). While Martin does not discuss paradigmatic traditions to the extent that I do, the assumptions of her three perspectives parallel the three groups I discuss here. The “assumptions” here refer to the epistemological, ontological, ethical, and methodological premises of a paradigm (Denzin & Lincoln, 1998).

ambiguity, complexity, and multiplicity at the center of its definition, tends to make no link at all between culture and effectiveness, and typically serves no interests in taking a descriptive stance. This study operates with the assumption that organizational culture is a variable, reflecting a functionalist philosophy in which organizational culture can be changed by administrators.

A Note on the Fundamental Role of Interpretivism

Organizational culture research, at a fundamental level, is interpretive research. Interpretivism's fundamental ideas are inherently a part of the foundation on which other traditions do their work; interpretivism's approach has "inspired" all of the three traditions. A brief explanation of interpretivism is warranted as it is important not to lose sight of this fact.

The interpretive sciences developed in reaction to the claims of positivism, an paradigmatic tradition that has taken the physical sciences as a role model and applied scientific methodology to the social world to create a "social science."¹⁰ In rejecting the positivist assumption that the social world can be understood in the same way as the physical world, the interpretive sciences have employed the concept of ideationalism, where ideas, values, understandings, and meanings are causal influences and where reality is perceived to be "socially constructed" (Alvesson, 2002a; Berger & Luckmann, 1966; Donaldson, 2003). For interpretivists, humans construct their reality through the ideas and values that constitute the meaning-making process:

¹⁰ To be more specific, it was against the claims of logical positivists that interpretive theories developed (Hatch & Yanow, 2003), which asserted that science can only deal in "observables" and that any proposition within the framework of traditional metaphysical doctrines should be rejected as meaningless (Donaldson, 2003).

Unlike rocks and atoms, humans make meaning, and so a human (or social) science needs to be able to address what is meaningful to people in the social situation under study. This requires understanding how groups, and individuals within them, develop, express, and communicate meaning, something that objective, unmediated observation (if that were even possible) cannot yield. (Hatch & Yanow, 2003, pp.65-66)

This quote points to a second difference between interpretivism and positivism: the aim for interpretivists is to *understand* social phenomena whereas the aim for positivists is to *explain* social phenomena (Charmaz, 2006). Interpretivists seek to understand the meaning-making process, the meaning humans give to both their behavior and to those of others. What unites interpretivists, then, is the exploration of the content of the meaning-making process: the ideas, values, beliefs, and feelings of lived experience (Hatch & Yanow, 2003). This goal of understanding the process of meaning construction leads to the proclamation that knowledge derives from interpreting meaning (Geertz, 1973). Interpretive philosophers, such as Wilhelm Dilthey and Hans-Georg Gadamer, argued that meaning is not expressed or known directly, but rather embedded in the artifacts generated by the meaning-making process (Hatch & Yanow, 2003). Artifacts (objects, language, acts), all that we have direct access to, become the focus of interpretation because they represent and embody meaning. This explains why a portion of organizational culture research has examined objects, such as products, logos, and signs; linguistic artifacts, such as stories, jargon, and humor; and acts, such as rituals, ceremonies, and taboos (Hatch, 1997; Hatch & Yanow, 2003; Trice & Beyer, 1984).

The point here is that ideationalism, the quest for understanding the meaning-making process, a focus on interpreting artifacts, along with other concepts derived from interpretivism, are a part of the platform from which functionalists, critical theorists, and postmodernists in organizational culture research spring. While researchers from these three traditions arrive at different conclusions, they all agree—to some extent—on the validity of interpretivism's approach and employ tools that originally developed out of the interpretive tradition.

Organizational Culture as Inspired by Functionalism

Organizational culture researchers inspired by functionalism typically do not separate the symbolic from the functionalist aspects of culture. They have been informed by theories in the social sciences, in anthropology and sociology especially, that hold that all aspects of the social structure serve a purpose and are indispensable in that they are linked to the survival of society (Donaldson, 2003). Examples of functionalist theorists include Malinowski (1935) and Radcliffe-Brown (1952) in anthropology, and Durkheim (1984), Parsons (1951), and Merton (1957) in sociology. These functionalists oftentimes used mechanical and biological metaphors to explain their theories, seeing groups and societies as machines or living organisms comprised of independent yet interrelated and interconnected parts that form an overall unified structure (Morgan, 2006). An organization, then, is a social system that has a functional unity where all parts of the system must work together with some degree of internal consistency. Functionalism, in a way similar to positivism, has borrowed concepts from the natural sciences and applied them to human groups and societies. It has, in particular, drawn from Charles Darwin's theory of evolution by natural selection (1964) in seeing individuals, groups, and

organizations as struggling for existence in a challenging environment, more often than not competing with others and being forced to either adapt or perish (Collins & Porras, 1994; Donaldson, 2003; Morgan, 2006). From the idea of social Darwinism, functionalism has stressed attributes that are connected to beneficial outcomes (survival).

Early on, functionalist research was underpinned by a perception of organizational culture as a unitary phenomenon, a perception in which cultural unity and survival are linked. Organizational culture inspired by the functionalist tradition focuses on mutually consistent interpretations, or the organization-wide consensus on meaning and interpretation. "Culture is like a solid monolith that is seen the same way by most people, no matter from which angle they view it" (Martin, 2002, p.94). Accordingly, research from this tradition defines culture as what is shared:

Culture is a pattern of shared basic assumptions that was learned by a group as it solved its problems of external adaptation and internal integration, that has worked well enough to be considered valid and, therefore, to be taught to new members as the correct way to perceive, think, and feel in relation to those problems. (Schein, 2004, p.17)

This definition (and others like it), it is important to point out, tend to rely on symbolic interaction theory, Clifford Geertz's symbolic anthropology in particular, where culture is a system of shared symbols and meanings and where symbolic action can be interpreted, read, or deciphered in order to reach understanding (Hatch, 1997; Patton, 2002; Smircich, 1983).

"Organizational culture" as a distinct area of research that was consistently, as opposed to sporadically, developed came into being in the late 1970s and early 1980s

(e.g., Pascale & Athos, 1981; Pettigrew, 1979), and these first developments were rooted in functionalism.¹¹ In the early 1980s, three best-selling books on “strong” corporate culture were published that popularized the organizational culture concept, as defined by this tradition (Alvesson, 2002a; Denison, 1990; Martin, Frost, & O’Neill, 2004; Ouchi & Wilkins, 1985; Smircich, 1983). William Ouchi’s *Theory Z* (1981), Terrence Deal and Allan Kennedy’s *Corporate Cultures* (1982), and Tom Peters and Robert Waterman’s *In Search of Excellence* (1982), all of which argued that a “strong” culture is characteristic of successful organizations, brought the functionalist organizational culture idea into the mainstream.

These three books share three characteristics central to their theory. First, each of these books argues that a self-conscious awareness of what values and beliefs make up a “strong” culture typifies a successful organization. Ouchi (1981), for example, describes “theory Z” cultures as ones in which senior administrators and, to some extent, all employees are aware of the guiding values and beliefs of the organization, and how the culture determines all procedures and rules. For Deal and Kennedy (1982), strong cultures are cohesive, revolving around explicitly understood values and beliefs that all employees share. Likewise, Peters and Waterman (1982) tell that “excellent” companies pay explicit attention to values: “Every excellent company we studied is clear on what it stands for, and takes the process of value shaping seriously. In fact, we wonder whether it is possible to be an excellent company without clarity on values and without having the

¹¹ Explicit interest in a cultural approach to organizational research began in the 1970s and matured in the 1980s and 1990s (Hatch, 1993; Hatch & Yanow, 2003; Martin, Frost, & O’Neill, 2004; Ouchi & Wilkins, 1985). However, social theorists have applied an interpretive approach—an attention to meaning, understanding, and interpretation—to organizations since the early twentieth century. Max Weber (1964) first developed and applied interpretive philosophical ideas to organizational settings and was followed by Barnard (1938), Roethlisberger et al. (1939), Whyte (1943), Selznick (1949), and Boulding (1956).

right sorts of values” (Peters & Waterman, 1982, p.280). Second, there is little need to prescribe rules for every action in these organizations as the cultural consensus is so strong that “two individuals who both understand the underlying theory will derive the same specific rule to deal with a particular situation” (Ouchi, 1981, p.41). The shared values of a strong culture provide a sense of common direction for all employees and guidelines for day-to-day behavior, to the extent that shared values act as “an informal control system” that sets expectations for behavior (Deal & Kennedy, 1982, p.33). The need for detailed procedures and rules declines as “in these companies, people way down the line know what they are supposed to do in most situations because a handful of guiding values is crystal clear” (Peters & Waterman, 1982, p.76). Third, shared values are transmitted through stories, myths, legends, and metaphors and serve as touchstones of the common commitment to certain values and beliefs. These cultural manifestations are the “glue” that holds individuals together and creates a functioning organization (Bolman & Deal, 2003; Ouchi, 1981; Peters & Waterman, 1982).

In agreeing that culture is a unitary whole where consensus, consistency, and clarity are supreme, researchers from this tradition do not value ambiguity (Hatch, 1997; Martin, 1992, 2002). As functionalist sociologists Auguste Comte and Emile Durkheim did, they see groups (organizations) as being kept together or cohering through “value consensus,” where solidarity—and social order—are morally good, primacy is placed on shared values, and a lack of consensus is perceived as abnormal (Durkheim, 1984).

Critics of this approach to organizational culture point to inconsistencies, disruptions, conflicts, and ambiguities in organizational life (Martin, 2002). To this, researchers from this tradition acknowledge that deviations do occur, but that they do so

at the relatively superficial levels that do not represent the deeper essence of organizational culture. Edgar Schein's theory of organizational culture (1985, 1992, 2004) best represents this position; for Schein, it is at the deepest level of culture where tacit assumptions are shared on an organization-wide basis.¹² He proposed three levels of culture—artifacts, espoused beliefs and values, and underlying assumptions—arguing that it is the level of basic assumptions that defines the essence of organizational culture (Schein, 2004). The pattern of shared, basic assumptions, which is linked to our human need for stability, consistency, and meaning, for Schein, is “so taken for granted that one finds little variation within a social unit” (2004, p.31). Consensus on these assumptions constitutes organizational culture:

What this “model” does say, however, is that only what is *shared* is, by definition, cultural. It does not make sense, therefore, to think about high or low consensus cultures or cultures of ambiguity or conflict. If there is no consensus or if there is conflict or if things are ambiguous, then, by definition, that group does not have a culture in regards to those things.

(Schein, 1991, pp.247-248)

The emphasis on what is shared in theories of culture based on the idea of cultural depth implies a normative position, where deviations are seen as problems that need fixing. Or, write James Collins and Jerry Porras in *Built to Last* (1994), visionary companies create a “cult-like culture” where there is such a tightness of fit between organizational culture and employees' values and beliefs that new employees either wholeheartedly buy into the culture or leave. (The same sentiment can be found in Collins' best-selling book *Good to*

¹² In this way, by getting at the assumptions that are taken for granted by people in an organization because they are so deeply understood the people don't even think about why they do what they do, Schein could be seen as both an ethnomethodologist and a functionalist.

Great [2001]). The point here is that deviations from the ideal of consistency, consensus, and clarity are perceived as either regrettable shortfalls that are to be remedied or as temporary setbacks to be followed by equilibrium (Martin, 2002; Martin, Frost, & O'Neill, 2004).

Organizational culture researchers who tacitly or explicitly support this functionalist view argue that culture is important to the extent that it can predict productivity or performance, or help organizations survive (Martin, 2002). In functionalist studies, organizational culture is treated as a variable (either independent or dependent, external or internal), with researchers searching for causal links to outcome variables, such as employee commitment, control, or profitability (Martin, 2002; Smircich, 1983; Willmott, 2003). For example, Kotter and Heskett (1992) sought to determine the relationship between corporate culture and long-term economic performance “to clarify the nature of and the reasons for such a relationship, and to discover whether and how that relationship can be exploited to enhance a firm’s performance” (1992, p.10). They tested the validity of the “strong culture” claim using survey research and quantitative analysis, finding a positive, albeit modest, relationship between strength of corporate culture and economic performance. A second empirical study that analyzed the relationship between culture and financial performance verified that culture does make a difference, with the quantitative evidence pointing to behavioral measures as strong predictors of performance and qualitative case studies showing an organization’s past successes and failures to be most significant (Denison, 1990).

These studies, therefore, assume culture can be changed and they explicitly advocate managerial control of cultural change (e.g., Tushman & O’Reilly, 1997; United

States General Accounting Office, 1992). Most of these studies assume Edgar Schein's position (1999, 2004) that if managers treat culture as a superficial phenomenon, assuming it can be manipulated at will, they will fail in truly transforming organizational culture. Instead, the focus must be on the learned, shared, and tacit assumptions that have to become explicit, assessed, and then modified. For example, in *Gaining Control of the Corporate Culture*, Ralph Kilman (1985) outlines five steps for closing "culture gaps": (1) surfacing actual norms, (2) articulating new directions, (3) establishing new norms, (4) identifying culture gaps, and (5) closing culture gaps (Kilman, Saxton, Serpa, & Associates, 1985). Kilman focuses on norms (the "unwritten rules of the game") instead of assumptions, but the logic is the same as Schein's. For both, the first step in changing culture is to see what previously has been unseen, or to express what previously has been unexpressed—to force tacit knowledge to the surface. And this "push" is management's responsibility.

In framing the ultimate question as how management can change or manipulate organizational culture—as a variable—to enhance performance, these researchers take a managerial point of view. In fact, most organizational culture studies are written in the managerial interest, aiming to help managers improve the productivity and performance of their organizations (Barley, Meyer, & Gash, 1988; Martin, 2002). These studies, first and foremost, seek ways to use culture to improve the effectiveness of an organization by increasing involvement, consistency, and a sense of shared purpose, thus increasing productivity (Denison, 1990). In functionalist terms, these studies seek to foster cultural change in the areas of internal coordination and external adaptation for organizational survival (Schein, 2004; Schultz & Hatch, 1996).

Organizational Culture as Inspired by Critical Science

Critical science has its roots in Marxism and has identified with the critical, emancipatory intent of the Marxist tradition since its beginning (Willmott, 2003). While critical theorists have drawn upon the German tradition of philosophical and social thought, as embodied by thinkers such as Kant, Hegel, Marx, Nietzsche, and Weber, and on the work of many others, including Freud, Husserl, and existential philosophers, to reshape orthodox Marxism and create a new radical interdisciplinary social theory, they have never lost Marx's commitment to emancipation (Kincheloe & McLaren, 1998). One of the best ways to understand organizational culture research inspired by critical science is to discuss Jurgen Habermas' theory of cognitive interests (1972), which illustrates this commitment to emancipation. In describing three competing conceptions of scientific knowledge, Habermas argues that it is the emancipatory interest that holds the most potential, as critical science is ultimately concerned with revealing action and meanings that are embedded in oppressive structures of domination that are subject to challenge and change (Willmott, 2003). Again, not all organizational culture research in this tradition has been absolutely determined by the emancipatory interest, but all at least has been informed by and utilized ideas from critical science, particularly Habermas' conception of critical science.

Three cognitive interests, Habermas contends, underpin what we interpret as knowledge: the technical interest, the practical interest, and the emancipatory interest (1972). The table below summarizes Habermas' description and evaluation of these interests:

Table 1

Habermas' Three Cognitive Interests

Cognitive interest	Type of science	Purpose	Focus	Orientation	Projected outcome
Technical	Empirical-analytic	Enhance prediction and control	Identification and manipulation of variables	Calculation	Removal of irrationality within means-ends relationships
Practical	Historical-hermeneutic	Improve mutual understanding	Interpretation of symbolic communication	Appreciation	Removal of misunderstanding
Emancipatory	Critical	Development of more rational social institutions and relations	Exposure of domination and exploitation	Transformation	Removal of relations of domination and exploitation that repress without necessity

Note. From *The Oxford handbook of organization theory* (p. 95), by H. Tsoukas and C. Knudsen (Eds.), 2003, Oxford: Oxford

University Press.

To relate each interest to the paradigmatic traditions discussed here, the technical interest usually connects to functionalism (and positivism), interpretivism is based on the practical interest, and the emancipatory interest is the foundation of critical science. The emancipatory interest's central purpose is human emancipation, especially by unveiling hidden forms of power that exploit the majority (Jermier, 1991). There are a host of assumptions that underlie critical science's methods of illuminating domination and exploitation, two of which are particularly relevant for organizational culture research: that certain groups in any society are privileged over others and that oppression has many faces, for example, class, race, and gender, that may be connected but cannot be collapsed¹³ (Kincheloe & McLaren, 1998). This second assumption, it is important to note, has led to a combination of social science and philosophy that advances (beyond orthodox Marxism's economic determinism) a social philosophy committed to expose and abolish *all* forms of oppression (Willmott, 2003).

In organizational culture research, these two assumptions have led to a concern not only with the implications of power-based social relationships for different groups but also to an understanding of differentiation. Organizational culture research inspired by critical science argues that culture is experienced differently at different levels or within different parts of the organization—organization-wide consensus becomes a myth (Hatch & Yanow, 2003; Martin, 1992; Martin, Frost, & O'Neill, 2004). By challenging the assumption that consensus exists (or even could exist) in symbolic interpretation within

¹³ Other assumptions that underlie the arguments of critical theorists are: All thought is fundamentally mediated by social and historical power relations; facts cannot be isolated from values; the relation between concept and object or and between signifier and signified is never stable or fixed; language is central to awareness; oppressed groups often accept their position as natural, necessary, or inevitable; and mainstream research (unwittingly) is implicated in the reproduction of existing power relations (Kincheloe & McLaren, 1998).

organizations, this perspective views organizations as “multicultural” and uncovers conflicts of interest and differences of opinion among groups (Gregory, 1983; Martin, 1992, 2002; Trice & Beyer, 1993). Here, conflict and difference replace homogeneity and unity as the focus shifts to subcultures, which may coexist in harmony, in discord, and/or in indifference to each other (Martin, 1992, 2002; Martin & Siehl, 1983). Cultural consensus is limited to inside the subcultural boundary; similar to the functionalist tradition, culture is what is shared, but this sharing is circumscribed to a group (Hatch, 1997; Martin, 1992). It follows that ambiguity is channeled outside the subcultures from this perspective. Within the subculture, clarity reigns as ambiguity is pushed to the interstices between subcultures (Martin, 1992, 2002).

In addition to the focus on subcultures, organizational culture research in this tradition stresses the inconsistent interpretation of cultural manifestations. In her qualitative case study of the cultures at a large multinational electronics corporation, Joanne Martin found three categories of inconsistency in interpretation: action, symbolic, and ideological (1992). Action inconsistency occurs when what is espoused differs from actual practice. In the case of the electronics corporation, employees described the company’s espoused equalitarian values as inconsistent with a variety of informal practices, such as the distribution of perks according to a manager’s “pull.” Symbolic inconsistency between the espoused egalitarianism in the company and cultural manifestations took the form of status-stratified seating arrangements in the company cafeteria, higher wall partitions for the engineers (who needed more “room for thought” than others), and jargon that differentiated groups. Ideological inconsistency is the result of espoused messages conflicting with each other. In the electronics company, the

priority given to “innovation” conflicted with equalitarianism as not all groups were perceived to contribute equally to fostering innovation; the engineers, seen as “more equal” than others, were given preferential treatment. The point here is that organizational culture studies inspired by critical science see not only the formation of subcultures based in groups’ shared meaning but also based in rifts in perception that underlie differences in lived experience.

Some of these studies stop at the point of describing subcultures. Others move past this point to a description of culture more in line with the emancipatory interest. These studies place power and conflict at their center, and every individual is perceived as having a position in relation to the dominant cultural group—no cultural space is neutral (Martin, 1992). Jermier, Slocum, Fry, and Gaines (1991), for example, compared and contrasted the official culture of an urban police organization with its subcultures, finding senior management unable to impose organization-wide conformity with the official culture. Using both qualitative and quantitative data, they found that four of the five subcultures—the “crime-fighting street professionals,” “peace-keeping moral entrepreneurs,” “ass-covering legalists,” and “anti-military social workers”—substantially modified or rejected senior management’s dictates.

In another example of contesting subcultures, Bartunek and Moch (1991) describe the cultural heterogeneity that emerged during the implementation of an effort to improve productivity and the quality of employees’ working lives through greater labor-management cooperation in a unionized commercial bakery. Each subculture responded to the program differently based on their perspective. Senior management was most concerned with control and thus employed the “control perspective;” local plant

management operated with the “paternalistic perspective” seeing employees as needing to be managed in the same way as a parent manages their children; line employees took the “dependence perspective” as there was a lack of a sense of responsibility on their part and they depended on others, especially management, to solve their problems; the machinists, members of a local union, viewed themselves as in competition with other groups, especially the line employees, and took the “competition perspective.” Both of these examples serve to make the point that organizations are not only multicultural and heterogeneous but conflictual as groups act with different power-based interests—organizational culture is “a contested reality” (Jermier, 1991).

Some of the studies in this tradition, not surprisingly, take an oppositional stance toward management. In concentrating on issues of control and individuals who resist being oppressed by management, these studies have led to the development of research around countercultures, or groups that are “pockets of resistance” to senior management’s interpretations (Hatch & Yanow, 2003; Martin, 1992). Martin and Siehl (1983) argue that subcultures can be categorized into three groups: enhancing, orthogonal, and countercultural. An enhancing subculture adheres to the core values of the dominant culture and an orthogonal subculture simultaneously accepts the dominant culture’s core values and a separate, unconflicting set of values. A counterculture, in contrast, proposes core values that challenge those of the dominant group. This challenge is not typically strident, however, but subtle, as employees deviate at the risk of being fired. For example, in “The Smile Factory: Work at Disneyland,” John van Maanen (1991) writes of the subtle resistance by employees in Tomorrowland:

Throughout the year, each land is assigned a number of area supervisors who, dressed alike in short-sleeved white shirts and ties with walkie-talkies hitched to their belts, wander about their territories on the lookout for deviations from park procedures (and other signs of disorder)....Supervisors in Tomorrowland are, for example, famous for their penchant of hiding in the bushes above the submarine caves, timing the arrivals and departures of the supposedly fully loaded boats making the 8½ minute cruise under the polar icecaps. That they might also catch a submarine captain furtively enjoying a cigarette (or worse) while inside the conning tower (his upper body out of view of the crowd on the vessel) might just make a supervisor's day—and unmake the employee's. (van Maanen, 1991, p.69)

Countercultures hold discordant values and explicitly oppose certain aspects of the dominant culture and, therefore, are seen as a threat to senior management (Boisnier & Chatman, 2003). Ironically, and in contrast to the functionalist perspective, when individuals deviate they become heroes who try to resist managerial dictates to conform (Martin, 1992).

It is researchers from the critical tradition who have pointed out that most organizational culture research has been written in the managerial interest. Smircich and Calas (1987), for example, reviewed the organizational culture literature, concluding that the field had become “dominant, but dead,” meaning that the initially innovative work that characterized the field had been dominated by the managerial interest and become a functionalist tool stripped of its vitality and promise (Frost, Moore, Louis, Lundber, &

Martin, 1991). In another review of the literature, Barley, Meyer, and Gash (1988) explained how functionalism had become the dominant paradigm within organizational culture research through the ideological convergence of practitioners and academics. They found that this convergence took place as a result of the academic community adopting a perspective like that of practitioners, who have acted on behalf of managerial interests. Both of these reviews assert that most of organizational culture research serves the managerial (and “technical”) interest, and by implication, are critical of managers’ attempts to engineer values and manage cultural change.

To the managerial interest in organizational effectiveness, the question becomes, “Who benefits from this kind of effectiveness?” (Martin, 1992). Organizational culture inspired by critical science tends to be skeptical of the “value engineering” found in *Gaining Control of the Corporate Culture* (Kilman, Saxton, Serpa, & Associates, 1985), for example, as it assumes culture can be controlled and manipulated by management. This skepticism oftentimes translates into taking an antimanagerial tone and a perspective that is more congruent with the views of groups that lack the power and status of senior management (Alvesson, 2002a; Martin, 1992; Martin, Frost, & O’Neill, 2004). Research inspired by the critical tradition has informed the “culture as control debate” by challenging the assumption that senior managers are the most influential members of an organization, and by showing how management’s opportunity to influence does not always guarantee that their intentions will be enacted (Hatch, 1997). While not all research inspired by the critical tradition has been absolutely determined by the emancipatory interest, the study of cultural differentiation has moved organizational

culture research beyond the normative position of the functionalist tradition and its “neutral” descriptions of the status quo.

Organizational Culture as Inspired by Postmodernism

Postmodern philosophy developed in opposition to modernism. In postmodern theory, modernism is generally thought of as the culmination of the “Age of Enlightenment,” which Michel Foucault characterized as an attitude rather than as a period or project (Willmott, 2003). This “attitude” was (is), in general, to rationalize human culture and society; it is marked by an unquestioned value for rationality and an effort to develop an integrated theory of the universe based on scientific principles and methods (Hatch, 1997). From a postmodern point of view, functionalism and critical science are modernist in that they are oriented toward constructing a unifying view, based on notions of truth and the search for “the one best way” (Hatch, 1997; Schultz & Hatch, 1996). In organizational research, these modernist traditions have deepened confusion and misunderstanding by misrepresenting the complexity of organizations (Martin, 1992). Instead, postmodernists assert that “knowledge is produced in so many different bits and pieces that there can be no reasonable expectation that it will ever add up to an integrated and singular view” (Hatch, 1997, p.44). Here, knowledge is fundamentally fragmented, multiplicitious, and contradictory.

Postmodern sources of inspiration to organization theory include Foucault (1972, 1973), Jencks (1977), Bourdieu (1977), Derrida (1978, 1980), Bakhtin (1981), Lyotard (1984), Baudrillard (1988), and Rorty (1989). Four axioms of postmodernism bring together this array of theorists. First, as opposed to describing the world as *being*—a permanent, static, and unchangeable interpretation of the world, postmodernism insists on

the world *becoming*—a world that is always in flux, changing, and emerging.

Postmodernism rejects modernists' emphasis on "substance, stability, identity, order, regularity, and form" for an emphasis on "process, indeterminacy, flux, interpenetration, formlessness, and incessant change" (Chia, 2003, p.128). Second, it critiques the idea of representation, language in particular, by claiming that the act of naming and symbolic representation treats the world as if it was made up of discrete units, which does not mirror reality; for postmodernists, language is inadequate in expressing thought (Alvesson, 2002b; Chia, 2003). Considering that elements of language—symbols, names, concepts, and categories—are the building blocks of theories, theories of how the world operates can neither tell us how the world really is nor be eternally true (Chia, 2003). This leads to a third point: the loss of power of "grand narratives" means a discrediting of theoretical frameworks and large-scale political projects (e.g., communism) (Alvesson, 2002b). "Against the grand narratives of universal truths, total control, and predictability that define the modernist agenda, postmodernism advocates a more tentative and modest attitude towards the status of our current forms of knowledge" (Chia, 2003, p.130). Finally, instead of thinking in terms of causal explanations, postmodernism privileges the ideas of "reminiscence, resonance, recursion, and resemblance"—the lack of rationality is typically alluded to in explaining causality (Chia, 2003). In this way, organizational culture is not stable, orderly, or static but fundamentally unpredictable, ambiguous, and fragmented.

Organizational culture research inspired by postmodernism has focused on, first, the complexity of relationships between one cultural manifestation and another and, second, the multidimensional nature of interpretation(s) (Martin, 1992). Again, this

approach describes organizational cultures as in a constant state of flux, as complex and multiplicitous. There is no clarity here: there are neither clear consistencies nor clear inconsistencies. That is, the perception of multiple interpretations inherent to this approach translates into seeing neither the organization-wide consensus characteristic of functionalist studies nor the subcultural consensus of critical science-based studies (Alvesson, 2002a; Martin, 1992). Instead there are multiple views on most issues, and these views are constantly in flux.

The third focus of organizational culture research in this area is ambiguity. Ambiguity occurs when there is no clear interpretation and differs from uncertainty in that clarity cannot be found after gathering more facts as the facts support more than one interpretation (Feldman, 1991). The most obvious ambiguity derives from no clear cultural center, producing ignorance and confusion in what actions to take, which ideologies to follow, and in how to interpret symbols. A less obvious, but more pervasive, ambiguity emerges from the complications that the clear oppositions of dichotomous thinking omit. The dichotomous thinking that makes up most of organizational culture research inspired by critical science—labor versus management, men versus women, blacks versus whites—is perceived as oversimplifying and misrepresenting the attributes and viewpoints of all members: An organizational culture is “a web of individuals, sporadically and loosely connected by their changing positions on a variety of issues. Their involvement, their subcultural identities, and their individual self-definitions fluctuate, depending on which issues are activated at a given moment”¹⁴ (Martin, 1992,

¹⁴ The extensive use of metaphors in explaining organizational culture could be a book onto itself. (There are, in fact, books on organization studies in general that use metaphors as organizing frameworks [e.g., Morgan, 2006]). In addition to the “jungle” metaphor wherein culture is understood as a messy “web” (Martin, 1992; Martin, Frost, & O’Neill, 2004, p.16), other popular metaphors for organizational culture

p.153). Here, postmodern organizational theorists have borrowed from Jacques Derrida, who coined the term “deconstruction” for the critical examination of oppositional thinking (that is characteristically dichotomous or “binary”) inherent in Western philosophy since the ancient Greeks (Kamuf, 1991). To undermine and subvert oppositional thinking about difference, Derrida pushed for a mode of thinking that allows for the fact that there is only one way to be the same, while there are many ways to be dissimilar (Derrida’s term for this was “différance”). Applied to organizational culture research, postmodernists would point to a complexity in culture(s) based on difference without opposition, which would include multiple, contradictory meanings that are simultaneously true and false, ironies, paradoxes, and irreconcilable tensions (Martin, 1992, 2002).

Ambiguity, an abnormal void that should be filled with meaning and clarity in the eyes of adherents of the other two traditions, is seen as normal, salient, and inescapable by postmodernists. Debra Meyerson (1991) demonstrated this in her study of hospital social workers, finding ambiguity to be a normal attribute of their work. She found ambiguity at two levels: within the occupational community as a whole and within individuals’ cultural experience. She writes that for this group “to dismiss the ambiguities in favor of strictly what is clear and shared is to exclude some of the most central aspects of the members’ cultural experience and to ignore the essence of their cultural community” (Meyerson, 1991, p.132). In this case of hospital social workers, the nature of their work demands engagement with ambiguity, and effective performance depends upon the acceptance and use of ambiguity (Hatch, 1999).

inspired by postmodernism are improvisational jazz performances (Hatch, 1999), collage art (Hatch, 1997), and garbage cans (March & Olsen, 1976).

Ambiguity is also at the center of questions of organizational effectiveness for organizational culture researchers inspired by postmodernism. Three positions have been expressed: ambiguity brings freedom, ambiguity brings chaos, and ambiguity is neither positive nor negative. For those who argue that ambiguity brings freedom, they point to the fact that there is no apparent right or wrong outcome when expectations, preferences, and evaluation criteria are unclear, and because there is less risk of being wrong a sense of psychological safety is created. And with this safety comes individual autonomy for experimentation (Martin, 1992). This logic applies to organizations as well: Without the constraints of prospective planning and rationalization, a playfulness can emerge that allows for experimentation, there are opportunities to improvise, and the organization can capitalize on the unanticipated benefits of the unpredictable and unexpected (Hatch, 1999; March & Olsen, 1976; Martin, 1992).

Other arguments about the link between ambiguity and effectiveness are less optimistic. Karl Weick (1991), for example, illustrated how separate small failures based in ambiguity became linked to result in a disastrous outcome. In describing an incident at Tenerife Airport in Spain where an unusual confluence of ambiguities combined to result in a collision of airplanes where 583 passengers died, he highlights communication-based ambiguities that led to the disaster. In particular, there was the phrase “we are now at takeoff” said by one of the crew members to the air traffic controllers and the Pan Am crew also preparing for takeoff:

Neither the air traffic controllers nor the Pan Am crew were certain what this ambiguous phrase meant, but Pan Am restated to controllers that they would report when they were clear of the takeoff runway, a communiqué

heard inside the KLM cockpit. When the pilot of the KLM flight was asked by the engineer, 'Is he not clear then, that Pam Am?', the pilot replied 'yes' and there was no further conversation. The collision occurred 13 seconds later at 5:06 p.m. (Weick, 1991, p.118-120)

Weick situates this conversation (that produced confusion and assumptions as to what "we are now at takeoff" meant) in a stressful environment where, among other factors, air traffic controllers were shorthanded, not accustomed to handling 747s, and working in English, which was a less familiar language. The point here is that ambiguity in communication, in a stressful environment with such a potential breakdown in coordination, added to the puzzlement that ended up in a disaster.

The third group makes no link at all between ambiguity and organizational effectiveness, arguing that it is wrong to assume that ambiguity is either positively or negatively related to effectiveness. Martha Feldman's 1991 study of policy analysts at the Department of Energy in Washington, DC, serves as an example. She found a high level of ambiguity, especially in the areas of intention and understanding, for this group that provides information and makes recommendations to aid their supervisors in policy making. First, ill-defined purposes led to little agreement about what the Department was supposed to be doing (and what policy analysts were supposed to be doing for the organization). No more than eight out of 34 policy analysts agreed on any general formulation of the Department's mission, and ten claimed that the Department did not have a mission at all. An ambiguity of understanding meant competing ideas of what was appropriate for the organization to do. Ambiguity stemmed from unclear relationships to outcomes:

Analysts are supposed to analyze relatively well-defined problems and produce solutions that can be implemented by politicians....[But] the ambiguities disrupt the possibility of such a smooth flow between analysis and politics. Analyses do not lead to positions that are promoted through politics. Analyses do not even support positions chosen by politicians. Analyses are being produced, but it is not clear for what or for whom.

(Feldman, 1991, p.155)

The impact of the cultural context can be seen here, as the Department of Energy is an organization that brings together diverse and partly conflicting tasks that are linked to the wider political environment, which is constantly in flux (Alvesson, 2002a; Bush, 2003; Trice & Beyer, 1993). In this case, the cultural context helps shape an organization culture saturated in ambiguity that “just is”—it is neither positive nor negative.

Feldman’s essay can be seen as a case-based illustration of the theory of organizational ambiguity (March & Olsen, 1976). With attention to organizational decision making and choice, the theory finds four features of an organization that may be ambiguous: the ambiguity of intention (e.g., multiple and conflicting goals), the ambiguity of understanding (e.g., multiple interpretations of intentions), the ambiguity of history (e.g., difficulty understanding what happened and why), and the ambiguity of organization (e.g., confusion about who in the organization is responsible for what the organization does) (Feldman, 1991; Hatch, 1999). The theory of organizational ambiguity is based on Cohen, March, and Olsen’s “garbage can model” of the organizational decision-making process (1972) when agreement about goals (or issues) and the means to achieve them are absent and, therefore, decision makers confront

ambiguity and uncertainty. Mary Jo Hatch describes the model as one in which “problems, solutions, participants, and choice opportunities are independent streams of events that flow into and through organizations, much like a random selection of waste gets mixed together in a garbage can” (1997, p.278). Whenever solutions, participants, and choice opportunities connect, a problem may be identified or solved, but because of the randomness of the process, solutions may be proposed where no problem exists or get attached to new problems or go entirely unheeded, participants may enter or leave the decision-making context, and choice opportunities may be made without solving a problem or never happen at all (Cohen, March, & Olsen, 1972; Martin, 1992). Cohen, March, and Olsen find universities to be a good example of this type of “organized anarchy,” where confusion and paradox are the rule and where sub-units are portrayed as relatively autonomous groups that are “loosely coupled” (Weick, 1976) with one another and with the institution itself.

The implication of the “garbage can model” for organizational culture is that, again, ambiguity, unpredictability, and fragmentation “just are;” they serve no interest. Organizational culture research inspired by postmodernism is, on the whole, ostensibly value neutral (Martin, 2002). Many portray organizational culture without making objections to the status quo or offering alternative views on how the culture(s) could and should change. Instead, for instance, they might assert that what one manager sees as desirable another may abhor, or find an equal amount of variance within another “group.” In finding and moving beyond the limits of rationality in organizational culture studies, this perspective tends to take a descriptive stance.

Organizational Culture: Variable or Metaphor?

The functionalist tradition has inspired a perception of organizational culture as a unitary phenomenon, where culture is that which is shared throughout the organization. The critical tradition joins the functionalist tradition in emphasizing what is shared, but confines this sharing to the subcultural boundary. Organizations become “multicultural” in that they are composed of groups with inconsistent interpretations of cultural manifestations. In contrast to both of these traditions, postmodernism has inspired research that centralizes fragmentation, where “what are called ‘organizations’ ...are nothing more than islands of relatively stabilized relational orders in a sea of ceaseless change” (Chia, 2003, p.131). These three conceptions of organizational culture, as I have explained, constitute the bulk of research on organizational culture.

Each of the three approaches to organizational culture discussed here has been criticized, and it is worth covering some of this ground. The main criticism of functionalist researchers is that although these researchers know unanimous agreement on culture is unlikely, they nevertheless impose an organization-wide consensus, thus imposing one perspective—that of management—over all others (Bush, 2003; Martin, 1992). In reifying a view of culture as homogeneous and harmonious, they legitimate organizational practices that ignore, downplay, or exclude the ideas, opinions, and interests of those who individually or collectively deviate from the supposedly dominant view, which stands as the “objective truth” (Martin, 1992, 2002; Martin, Frost, & O’Neill, 2004). The main criticisms leveled against research inspired by the critical tradition takes a postmodern position: organizational culture inspired by the critical tradition recognizes complexity, disagreement, and conflict, but does so in oppositional

terms, with no appreciation of ambiguity. And like the functionalists, critical researchers assume an objectivist stance, claiming “truth” and a clear understanding about what people share or where they disagree (Martin, 2002). For our purposes, the most important critique of research inspired by postmodernism are the practical dangers of focusing on ambiguity. How can the endless ambiguities of this approach be reconciled with the clarity required for a commitment to action (based in human agency) in organizations? (Willmott, 1994). Ambiguity and fragmentation are important aspects of organizational culture, but in the same way that there are limits to management engineering of culture, there are limits to ambiguity; organizations must be successful in shaping even a moderate degree of common understandings on at least some issues to survive (Alvesson, 2002a).

An issue at the core of these three perceptions of organizational culture and those of their critics is the question of whether culture is a variable or a metaphor (Alvesson, 2002; Barley, Meyer, & Gash, 1988; Connell & Nord, 1996; Hatch, 1997; Jermier, 1991; Martin, 2002; Smircich, 1983). In a landmark article on this issue, Linda Smircich (1983) traces the ways culture has been developed as a critical variable and as a root metaphor in organizational studies. In reviewing the literature, she shows how these two different conceptions of organizational culture have led to some researchers prioritizing the principles of prediction, generalizability, control, and especially, causality (e.g., Ouchi, Deal and Kennedy, Peters and Waterman, and Schein) while others concerned themselves with, to them, the more fundamental issues of the meaning-making process. The former assumes an “objectivist” position and privileges ideas about causality where

culture can be molded, shaped, and changed in ways consistent with managerial interests and the latter assumes a “subjectivist” position and promotes a view of organizations in expressive, ideational, and symbolic terms, seeing them as manifestations of human consciousness (Smircich, 1983). For one, culture is something an organization *has* and for the other, culture is something an organization *is* (Smircich, 1983).

This study operates with the assumption that organizational culture is a variable, reflecting a deterministic philosophy in which causes probably determine effects or outcomes and, accordingly, knowledge is developed through empirical observation and measurement (Creswell, 2003). As an evaluation, which ultimately is explanatory research as it looks for causality, culture here is treated as a variable and data, evidence, and rational considerations are viewed as the shapers of knowledge. Therefore, despite its shortfalls and biases in interests, organizational culture inspired by the functionalist tradition offers the most promise for this study.

Evaluation Utilization

A discussion of theory goes hand in hand with a discussion of methodology, and it is to methodology that I now turn. Research, after all, is an empirical undertaking done systematically and logically; “Method is the attribute that distinguishes research activity from mere observation and speculation” (Shulman, 1997, p.7). Frameworks that theories provide must meet with methodology to construct an empirical research design. In building this design, moreover, care must be taken to deeply understand research problems, questions, and purposes, which in this case center on organizational

improvement. It is ultimately the underlying value of improvement that drives the following discussion of evaluation methodology, which ironically begins with the main theoretical traditions within this field.

To address the need to build a comparative theory of evaluation, or as William Shadish put it, to find the “fine line between evaluation as a United Nations and evaluation as a Tower of Babel” (1998, p.13), Marvin Alkin and Christina Christie developed the evaluation theory tree (2004), holding that the field can be divided into three traditions, each seeing the purpose of evaluation differently. One tradition is the “methods” tradition where evaluation serves to collect accurate and scientifically-sound evidence; a second tradition is the “valuing” tradition where evaluation is used to make value judgments. Within the third tradition, the “use” tradition, evaluation is conducted to gather information for decision making (Alkin & Christie, 2004; Patton, 1997). The remainder of this section will focus on this third evaluation tradition, evaluation utilization, as it is here where organizational research has been integrated to the fullest extent.

Theorists within the evaluation utilization tradition are concerned with the way in which evaluation information will be used and focus on those who will use the information (Alkin & Christie, 2004). This orientation was born from experience: Early evaluators recognized that their evaluations were not being used (see Alkin, Daillak, and White, 1979; Patton, 1978, 2002; Patton, Grimes, Guthrie, Brennan, French, & Blyth, 1977; Weiss, 1972, 1977) and this inspired conversations about and taught evaluators the value of pragmatism. The word “use”, for practical purposes, stands in for pragmatism, the philosophical position that the usefulness, workability, and practicality of ideas,

policies, and proposals are the criteria of their merit (Datta, 1997; Greene & Caracelli, 1997; Murphy, 1990; Patton, 2002; Pawson & Tilley, 1997). This stance forced evaluators to reconsider their (naïve) assumption that evaluation results would be used and to see evaluation as a political act in a context where power, ideology, and interests are paramount and influence decisions more than evaluative feedback. Therefore, they concluded, evaluators must take active steps to increase the use of their results (Shadish, Cook, & Leviton, 1991).

The evaluation utilization tradition began with “decision-oriented theories,” theories which hold that it is critical to conduct evaluations that are designed specifically to assist key program stakeholders (stakeholders who more often than not commissioned the evaluation) in program decision making (Alkin & Christie, 2004). Three evaluators early in the evaluation utilization tradition, Daniel Stufflebeam, Joseph Wholey, and Carol Weiss, for example, paid close attention to decision makers. Stufflebeam not only advocated for his “CIPP” model,¹⁵ one of the most well-known “decision-oriented” theories, but has also argued for engaging stakeholder panels (usually made up of those in decision-making positions) to define the evaluation questions, review evaluation plans and draft reports, and disseminate findings (Alkin & Christie, 2004; Stufflebeam, 2004). Similarly, Wholey has focused on managers and policymakers. Wholey’s central focus, in fact, has been the use of evaluation for the improvement of management; his work attends to the complexity of managing an organization, viewing evaluation as an important process to make management more effective (Alkin & Christie, 2004; Wholey,

¹⁵ Stufflebeam describes evaluation as the following process: identifying needs to decide upon program objectives (context), decisions on strategies and designs (input), identifying shortcomings in a current program to refine implementation (process), and measures for outcomes for decisions around continuation or refocus of the program (product) (Stufflebeam, 1991).

1983). Weiss was among the first evaluators to show that evaluations do not take place in a political vacuum; for Weiss, evaluation is not only a research activity but a political activity, heavily influenced by decision makers, as well (Alkin & Christie, 2004; Weiss, 1998, 2004). These evaluators all stress improvement and would agree with Stufflebeam that “the most important purpose of evaluation is not to prove, but to improve” (Stufflebeam, 1991, 2004). In this way, these early evaluation utilization theorists, along with others from this tradition, conceive of evaluation as a functional activity oriented to improving program performance (Wholey, 2004).

In the present day, at the center of the evaluation utilization tradition is Michael Patton, particularly his book *Utilization-Focused Evaluation* (1978, 1986, 1997). Patton defines utilization-focused evaluation as “a process for making decisions about [evaluative] issues in collaboration with an identified group of primary users focusing on their intended uses of the evaluation” (Patton, 2004, p.278). Patton has consistently asserted that a primary factor in achieving use is the “personal factor” (Alkin & Christie, 2004; Patton, 1997), which means that people, not organizations, use evaluation information and the likelihood of an evaluation being utilized is greatly enhanced with the identification of individuals or groups who personally care about the evaluation and its findings—i.e., those who will feel a sense of ownership (Patton, 1997, 2004). Thus, Patton emphasizes “intended primary users,” those who are likely to be real users of the evaluation. This group, for Patton, must be identified in the beginning of the evaluation and involved throughout as intended use by intended users is the primary consideration from beginning to end (Patton, 1997, 2004). The “personal factor” has also led to Patton’s assertion that evaluators be “active—reactive—adaptive” (Alkin & Christie,

2004; Patton, 1997). This means that evaluators need to be active in identifying intended users and focusing questions, reactive in continuing to learn about the evaluative situation as it unfolds, and adaptive in changing evaluation questions and designs as a result of intended users' increased understanding (Alkin & Christie, 2004). What this suggests is that the evaluator be a negotiator who is flexible and responsive in modifying the evaluation to better serve user needs (Patton, 1997).

As Stufflebeam, Wholey, Weiss, Patton, and others pushed this tradition forward, and as the importance of use continually emerged, discussions and investigations of the types of use and the factors and conditions that affect it intensified. As a result, the concept expanded; "a good theory of use comprehensively discusses possible kinds of use, time frames in which use occurs, and things evaluators do to facilitate use" (Shadish, Cook, & Leviton, 1991, p.54). For example, evaluators now distinguish between short- and long-term use and discuss promoting use by being cognizant of choice of issues addressed, the role the evaluator adopts toward potential users, and the communication channels chosen for reporting results (Shadish, 1998; Shadish, Cook, & Leviton, 1991). The expansion of types of use, however, has overshadowed other aspects of use.

A literature on types of use has grown to delineate five types of use: instrumental, conceptual, process, symbolic, and misuse.¹⁶ Instrumental use refers to the direct application of what has been learned from the evaluation, the tangible and observable use of findings (Henry & Mark, 2003; Mark & Henry, 2004; Russ-Eft & Preskill, 2001; Shadish, Cook, & Leviton, 1991). Conceptual use occurs when evaluation findings are used to clarify an individual's or group's conceptualization or perception of that which is

¹⁶ Conceptual use is also called "knowledge use," "enlightenment," or "demystification" (Henry & Mark, 2003; Mark & Henry, 2004; Russ-Eft & Preskill, 2001; Shadish, Cook, & Leviton, 1991). Symbolic use is also called "political" or "persuasive" use (Russ-Eft & Preskill, 2001).

being evaluated. The individual's or group's thinking changes as a result of learning the findings (Henry & Mark, 2003; Mark & Henry, 2004; Russ-Eft & Preskill, 2001; Shadish, Cook, & Leviton, 1991). Process use refers to individual, program, or organizational changes that occur as a result of the learning that transpires during the evaluation process (Cousins & Lee, 2004; Cousins & Whitmore, 1998; Patton, 1997, 2002; Preskill & Torres, 2000). It is stimulated not by the findings, but by participation in the evaluation process. Symbolic use "walks a fine line between legitimate use and misuse" by use inspired for symbolic (and/or political) reasons (Bolman & Deal, 2003; Russ-Eft & Preskill, 2001, p.29). Finally, there is the misuse of the evaluation process or findings, which is usually an issue of human relations and political pressure (Stevens & Dail, 1994).

Evaluation Utilization and the Organizational Context

Developments within the evaluation utilization tradition, such as elaborations on the idea of "use", and evaluation approaches that appreciate the internal nature of organizations, such as attention to organizational decision makers and intended users, have allowed for connections to be made between evaluation utilization and organizational research. Two ideas in particular of the evaluation utilization tradition have become the foundation of work that integrates evaluation and organizational research: (1) process use, or changes in stakeholders or the organization as a result of participation in the evaluation, and (2) conceptual use, or learning that takes place as a result of the evaluation. Building off Patton's utilization-focused evaluation, what both of these ideas centralize is the *relationship* between evaluator and users; Torres and Preskill explain, "While better understanding of issues surrounding evaluation use was

helpful, it did not necessarily translate to significantly enhancing use. As a result, a large number of evaluation researchers turned to exploring the relationship between evaluations and users” (2001, p.388). Attention to this relationship, grounded in the ideas of process and conceptual use, spurred research on how the degree of internal involvement, engagement, and ownership amongst organizational participants in the evaluation will affect the nature and degree of impact on the organization (Patton, 1997). Theorists whose work begins with attention to the relationship between evaluations and users can be separated into groups, two of the most prominent being (1) those who expound participatory and empowerment evaluations and (2) those who link their work to organizational learning and development (generally referred to as developmental evaluation) (Patton, 1997). Key theorists who advocate participatory and empowerment evaluation are J. Bradley Cousins and David Fetterman; key theorists who advocate developmental evaluations are Hallie Preskill and John Owen. These two groups of theorists, both of whom have extended the concepts of utilization and evaluation impact beyond the particular program being evaluated to include ideas from organizational research, are discussed below.

Participatory evaluation assumes that stakeholder participation in evaluation will enhance its relevance and create a sense of ownership, thus securing utilization (Cousins & Whitmore, 1998). The results of the evaluation should be constructed by program participants, not the evaluator; “the locus of judgment about program merit, worth, and significance should lie with nonevaluator stakeholders located within the program community” (Cousins, 2004, p.321). In their *Participatory Evaluation in Education* (1995), Cousins and Earl write that the evaluator here acts as a collaborator, an equal

partner, who fosters informed judgment making (Cousins, 2004; Cousins & Earl, 1995). Like other models, stakeholders are engaged in all phases of the evaluation, but what distinguishes participatory evaluation is its recognition of the organization in grouping together intended users. Cousins writes, “much of our research on participatory evaluation has pointed to conditions and factors associated with organizational context and culture as being key in explaining the extent to which the approach is effective in leading to desirable practical consequences” (Cousins, 2004, p.327-8). Cousins’ concern with the organizational context recently led him to explore the problem of integrating evaluative inquiry into the organizational culture (Cousins, Goh, Clark, & Lee, 2004). In a literature review that connects evaluation utilization, evaluation capacity building, and organizational learning, the authors argue that the capacity to do effective evaluation is linked to the kinds of capacity-building outcomes of evaluation that are generally associated with evaluation utilization and that evaluation capacity building is tightly connected to organizational development and learning (Cousins, Goh, Clark, & Lee, 2004).

Cousins is careful to separate what he calls “practical” participatory evaluation from “transformative” participatory evaluation (Cousins & Whitmore, 1998). The former is pragmatic in that it supports organizational decision making and problem solving and is linked to organizational learning whereas the latter is aligned with the principles of emancipation and social justice (Cousins & Whitmore, 1998). Cousins and Whitmore (1998) explain that the two have different goals, functions, and historical and ideological roots, and that Cousins and Whitmore’s aim is practical problem solving (Cousins & Earl, 1995). David Fetterman’s empowerment evaluation, on the other hand, is an overtly

political process. Fetterman infuses principles of social justice into evaluation work, describing a process in which the use of evaluation concepts, techniques, and findings foster self-determination (Fetterman, 2003, 2004; Fetterman, Kaftarian, & Wandersman, 1996). Fetterman encourages program participants to in essence conduct their own evaluation, with the outside evaluator serving as a coach or “critical friend;” the goal is self-determination rather than dependency, to “help people help themselves” (Fetterman, 2004, p.305). The end point, for Fetterman, is not the assessment of a program’s worth, but, to return to process use, program participants and the organization as a whole traversing through an emancipatory process.

The second group under consideration here has focused on the role of evaluation in organizational learning and development and employs the idea of conceptual use to argue that when individuals participate in an evaluation process that is collaborative, learning occurs not only at the individual level but also at the group and organizational levels (Preskill & Torres, 2000). Preskill’s contention is that evaluation is a learning process, a catalyst for individual, group, and organizational learning (Preskill, 2004; Preskill & Torres, 1999, 2000). As such, it is a change-oriented tool for “transformative learning”: a process where information is identified, acquired, examined, questioned, validated, and revised in order for individuals, teams, and organizations to meet their goals (Alkin & Christie, 2004; Preskill, 2004; Preskill & Torres, 2000). To realize this process, evaluation is conducted using collaborative, participatory, and learning-oriented approaches where the evaluator becomes a facilitator who helps guide the leaning process occurring during the evaluation (Preskill, 2004). The evaluator operates with the goal of

creating a seamless blend of program work, research, evaluation, and organizational development (Alkin & Christie, 2004; Preskill & Torres, 2000).

Owen is also interested in organizational development and change and, like Preskill, he sees utilization as paramount in planning, negotiating, designing, and conducting evaluations. Unlike Preskill, however, who emphasizes planned and purposeful evaluations, Owen stresses the need for evaluators to be adaptable to stakeholder needs and concerns (Alkin & Christie, 2004). For Owen, this adaptability is most important during the planning and negotiation stage, before any evaluative data has been collected. Stakeholder participation is critical at this juncture as it is here when the way the rest of the evaluation will be undertaken is determined (Alkin & Christie, 2004; Owen, 2004; Owen & Rogers, 1999). The rationale for this has to do with an explicit understanding of systems thinking:

...the work of most evaluators has been directed towards bringing about incremental improvement in individual program interventions within an organization rather than contributing directly to thinking about the underlying policies or structures within that organization...but systems thinkers reject the notion that incremental improvement discussed above is sufficient or is the only type of change possible. What systems thinking tells us is that, while incremental changes can be useful, organizations must confront the reality that change of a more fundamental kind must be made if they wish to address long-term problems that face them. This implies that evaluation findings should be used to inform leaders so that they can accurately conceptualize the dynamic between individual

programs, and the policies and structures which support them. (Owen & Lambert, 1995, p.242)

This argument redefines the relationship between evaluator and organizational leadership in terms of the nature of the information provided: Leaders need to be well informed on the strategic aspects of their work and evaluators ought to consider their unique position to inform the mental models of leaders, thus influencing decision-making and organizational learning (Owen & Lambert, 1995, 1998). It is for this reason that Owen advocates for a blurring of the roles between evaluator and organizational development consultant (Owen & Lambert, 1998). And it is for this reason that others have called this approach “developmental evaluation,” where the evaluator becomes part of the design team helping to shape both organizational processes and outcomes in an environment of interaction, feedback, and change (Patton, 1997).

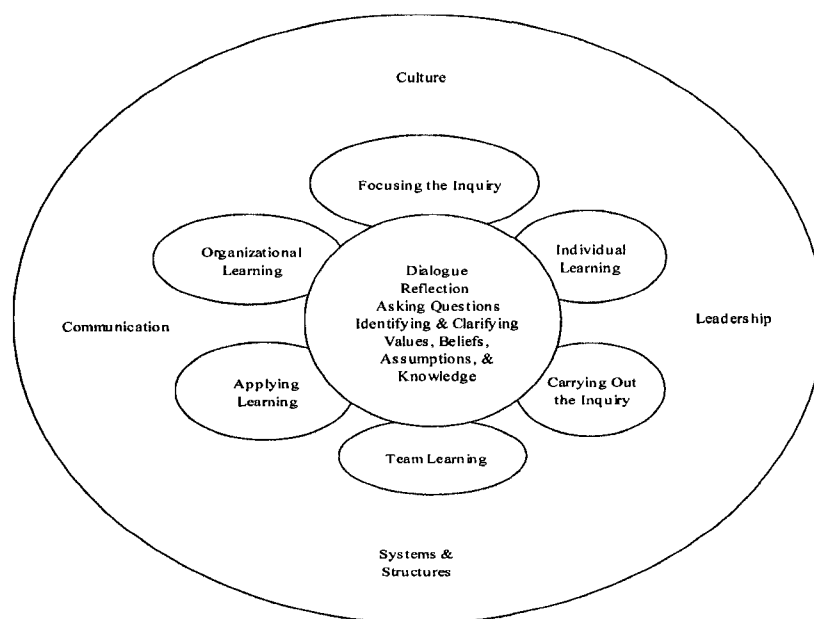
These four evaluators are certainly not the only individuals to connect evaluation utilization and organizational research, nor are they the only proponents of participatory, empowerment, and development evaluation (Alkin & Christie, 2004; Patton, 2002; Russ-Eft & Preskill, 2001; Seiden, 2000). They are, however, the most prominent within these fields. The more important point here has less to do with these individual evaluators than a trend within evaluation utilization: “The past fifteen years have marked a shift in utilization research from the investigation and classification of factors influencing utilization to a more complex examination of use that incorporates a sensitivity to organizational context” (Seiden, 2000, p.24). The fields of evaluation utilization, organizational learning, and organization development and change—all fundamental to

participatory, empowerment, and development evaluation—emphasize the importance and influence of organizational context (Seiden, 2000).

To illustrate this point, take Preskill and Torres' *Evaluative Inquiry for Learning Organizations* (1999), where a theory of developmental evaluation with three components is developed. First, if organizations are to benefit from evaluative inquiry, they must understand how learning occurs at the individual, group (team), and organizational level. Second, there are four learning processes that facilitate evaluative inquiry: (1) at the center, implying throughout the evaluation, there is dialogue, reflection, asking questions, and identifying and clarifying values, beliefs, assumptions, and knowledge; (2) focusing the inquiry; (3) carrying out the inquiry; and (4) applying learning from the evaluative inquiry. The diagram below shows these two components.

Figure 1

Preskill and Torres' Theory of Developmental Evaluation



Note. From *Evaluative inquiry for learning organizations*, by H. Preskill and R. T.

Torres, 1999, Thousand Oaks, CA: Sage.

The diagram also shows the third component, the organization's infrastructure. The critical elements of this infrastructure are culture, leadership, communication, and systems and structures. For evaluative inquiry to be successful, the organization's infrastructure must support and facilitate it: "How an organization is structured, how individuals' jobs are designed, how supportive the organization's culture is of evaluative inquiry, the extent to which the leadership models and supports learning and inquiry, and how information is distributed and accessed strongly influence the extent and the ways in which evaluation is successful" (Preskill, 2004, p.347). They advocate diagnosis of this infrastructure to determine readiness for evaluation, to frame their findings, and to set realistic goals and expectations for evaluation use (Alkin & Christie, 2004; Preskill & Torres, 2000).¹⁷

This diagnosis, then, is grounded in an understanding of organizational research. The problem, however, is this understanding is not as fully developed as it ought to be. Preskill and Torres' model above reflects the state of the integration of organizational research and evaluation utilization research. Their model, in a way absent of much precision, groups together culture, leadership, systems and structures, and communication as "the organizational infrastructure;" what is needed is a more detailed approach that separates out these components of the organizational context, to learn more about each of these pieces. This lack of a precision is not surprising considering the research base upon which Preskill and Torres and other evaluators must work. In the aforementioned literature review that connects evaluation utilization, evaluation capacity building, and organizational learning, Cousins et al. (2004) found the knowledge base linking

¹⁷ As an appendix in *Evaluation in Organizations* (2001), Russ-Eft and Preskill offers evaluators a list of questions to assist in determining an organization's readiness for evaluation.

evaluative inquiry to organizational development and learning to be at a relatively early stage of maturity, as evidenced by few empirical studies:

Methodologically, we conclude that research on evaluation as an organizational learning system is highly underdeveloped and that confidence in knowledge in this domain depends quite significantly on the extent to which more, and different, research occurs. Many of the studies that we located, some would argue, are not empirical studies at all.

(Cousins, Goh, Clark, & Lee, 2004, p.131)

The point here is that while there has been progress in the integration of organizational research and evaluation utilization research, the level of integration has not significantly deepened, as evidenced by the lack of empirically-based research—the research base upon which the integration of organizational research and evaluation utilization research rests is underdeveloped. One consequence of this is a lack of clarity about the specific roles of the parts within “the organizational infrastructure,” such as organizational culture.

Evaluation and Quality

The current period of expansion of the field of evaluation in the U.S. began with the election of Ronald Reagan as President, who, after a period of social democratic liberalism, ushered in conservative neoliberal economic policies, such as lower taxation, the dismantling of the public sector, privatization, and most importantly for our purposes, fiscal restraint (Hursh, 2006; Krieger, 1986). Underlying these policies was a primacy placed on the values of external accountability and internal quality, issues which immediately rose to the fore of public discourse (Patton, 1997; Russ-Eft & Preskill,

2001). For example, the belief that the Great Society and War on Poverty programs of the 1960s were ineffective and wasteful—failures, in short—became widespread and fueled support for neoliberal policies. Continuing concerns about federal budget deficits and runaway entitlement costs continued into the 1990s, intensifying the debate about the effectiveness of government programs, particularly the relationship between program funding and program impacts (Patton, 1997). “Both conservatives and liberals were faced with public demands to know what had been achieved by all the programs created and all the money spent. The call for greater accountability became a watershed at every level—national, state, and local; public sector, nonprofit agencies, and the private sector” (Patton, 1997, p.13). It becomes clear that the call for accountability not only spurred the assessment movement in higher education, but also the current growth and popularity of evaluation. This external call for accountability led to a flood of concern with internal quality and effectiveness, a concern that manifested itself as a concern with internal evaluation (Russ-Eft & Preskill, 2001).

The quality improvement campaigns begun by most major corporations in the 1980s were followed by quality campaigns in the health care, government, and education sectors in the 1990s (Detert, Seashore Louis, & Schroeder, 2001; Evans & Dean, 2004; Patton, 2002). In the 1990s, postsecondary education became the focus of quality improvement (after a wave of public concern with quality in education at the elementary and secondary levels in the mid-1980s inspired by *A Nation at Risk*) (Alstete, 1996). In 1993, for example, the American Association for Higher Education initiated the Continuous Quality Improvement (CQI) Project and, that same year, a *Business Week* survey found 61 percent of college presidents averring involvement in Total Quality

(Ewell, 1993b). One result of this was higher education borrowing new management-based improvement tools and strategies, such as continuous improvement, breakthrough improvement, reengineering, benchmarking, and perhaps the most popular, Total Quality Management, from the corporate sector (Ewell, 1999). There is, in fact, a sizable literature on the quality movement in higher education (American Association for Higher Education, 1993, 1994; Gaither, 1998; Krueger, 1993; Massey, 2003; Seymour, 1992, 1994; Sheer & Teeter, 1991; Teeter & Lozier, 1993).¹⁸ My task here is not to review this literature in detail but to make the point that these management innovations linked to external accountability and internal quality led to greater use of evaluation processes and results in higher education. The concern with the quality of higher education institutions, academic programs, courses, and classrooms has been tied to evaluation, and as calls continue for results-oriented, accountable programming rooted in the demand for quality, evaluators will be challenged to increase the use and effectiveness of evaluations, thus making the connection between evaluation and quality tighter (Patton, 1997).

Total Quality and Evaluation

My focus here is the connection between evaluation and Total Quality, a philosophy of management aimed at comprehensive, organization-wide improvement in the quality of products and services as well as a collection of tools and strategies for implementing this philosophy (Evans & Dean, 2003). Before this link is explicated, some background on Total Quality is necessary. Interestingly, the popularity of research on “strong” organizational cultures and Total Quality has the same source—answering the question why the U.S. was economically “losing” to Japan during the late 1970s and

¹⁸ Similar to the literature on support for assessment at the institutional level, the literature on Total Quality Management (TQM) in higher education mostly consists of descriptions of its adoption and prescriptions for “how to” implement TQM techniques (Entin, 1993).

early 1980s. In this context of many U.S. businesses losing significant market share to global competitors, Japan in particular, the National Broadcasting Corporation (NBC) aired a program in 1980—"If Japan Can... Why Can't We?"—featuring Edwards Deming, one of the foundational thinkers of the quality movement, which introduced him to the U.S. business mainstream (Evans & Dean, 2003). Soon after, Joseph Juran, Philip Crosby, Walter Shewhart, and Armand Feigenbaum, other Total Quality foundational thinkers, also gained national attention. Using their philosophies, most major companies embarked on quality improvement campaigns in the 1980s, implementing "quality" in all sort of ways. Consequentially, the need for frameworks and/or standards for quality emerged. Today, organizations can turn to the two most prominent frameworks for quality, the International Organization for Standardization (ISO) 9000 and the Malcolm Baldrige National Quality Award, both central to setting standards and criteria for quality.

There are six basic elements to the innumerable approaches to Total Quality: (1) customer focus, (2) a process orientation, (3) empowerment and teamwork, (4) continuous improvement and learning, (5) management by "fact" (evidence), and (6) leadership and strategic planning (Chang, 1996; Detert, Bauerly Kopel, Mauriel, & Jenni, 2000; Detert & Mauriel, 1997; Evans & Dean, 2003; Massey, 2003). The first element, a customer focus, means the customer judges quality, implying understanding their needs, listening to and learning from their input, and building satisfying relationships. In education, the discussion around "student centeredness" (e.g., Hurtado, Milem, Clayton-Pedersen, & Allen, 1999) and defining education quality in terms of outcomes (Massey, 2003) builds off of this principle. Second, a process orientation pushes organizations to

value working not only hierarchically but horizontally and cross-functionally as well, thereby creating a holistic understanding of the organization based on appreciating the linkages among all activities and units, “those spaces between the boxes on an organization chart” (Evans & Dean, 2003, p.18). A process orientation has led to viewing organizations as “systems” where interaction, alignment, and cooperation among units is crucial for overall organizational performance (Massey, 2003; Senge, 1990). In higher education, this might mean a focus on the process of teaching, learning, and student assessment or curricular coherence (Massey, 2003). Third, “empowerment,” similar to participatory, empowerment, and developmental evaluation, means the sharing of authority among all employees and involving them in decision making. From this, vertical, horizontal, and interorganizational teamwork and collaboration is fostered (Evans & Dean, 2003).

Fourth, continuous improvement, of both the products and services created and the systems that create them, and learning is part of the management of all systems and processes, and refers to both incremental and “breakthrough” improvement. The vibrant discussion about identifying and learning from “best practices” in higher education reflects the principle of continuous improvement and learning. As will be explained below, the drive to understand the links between key outcomes and the processes that influence them in the interest of improvement has led to the use of evaluation. Fifth, data and information should support analysis at all organizational levels. Organizations require performance measures that draw on “fact” or evidence and decision making that is data-based (Dowd, 2005). Finally, leadership for quality is the responsibility of senior administrators. “Senior leaders must set directions; create a customer orientation, clear

quality values, and high expectations that address the needs of all stakeholders; and build them into the way the company operates” (Evans & Dean, 2003, p.26). That is, senior administrators are responsible for the infusion of each of these principles into the organization; senior administrators must make their commitment to quality a priority, build support for the initiative, and model their expectations (Detert, Bauerly Kopel, Mauriel, & Jenni, 2000). They must also implement their commitment in a manner that is planned, which is to say, senior administrators must engage the organization in a process by which the members of the organization envision its future and develop the strategies to move toward that vision (Evans & Dean, 2003).

Of these six principles, continuous improvement and learning has been especially important to the field of evaluation. In fact, Michael Patton finds evaluation to serve three primary purposes that parallel the three traditions of Alkin & Christie’s “evaluation theory tree” (1997). One is “rendering judgments” (valuing) and another is “generating knowledge” (methods). For the third, Patton does not use the word “use” but instead the phrase “facilitating improvements.” He argues that the improvement purpose is informed by a developmental perspective and includes formative evaluations, responsive evaluations, humanistic evaluations, and evaluations rooted in organizational learning, quality enhancement, and Total Quality Management, among others (Patton, 1997). All of these approaches use information systems to monitor program efforts and outcomes related to improvement, and to do so regularly over time to provide feedback for further improving the program (Patton, 1997). By assuming that “social problem solving can be improved by incremental improvements in existing programs, better design of new programs, or terminating bad programs and replacing them with better one” (Shadish,

Cook, & Leviton, 1991, p.37), *program evaluation can be considered a tool for continuous improvement*, a tool for incrementally improving quality in particular (Evans & Dean, 2003; Konidari & Abernot, 2006; Light, 2001; Mark, 2003; Russ-Eft & Preskill, 2001).¹⁹

Of the different evaluation approaches, formative evaluations are especially linked to the idea of continuous improvement. The intent to improve and formative evaluation have been wedded to the extent that the two terms—“formative evaluation” and “continuous improvement”—are oftentimes used interchangeably (Donaldson & Gooler, 2003; Ewell, 1993b; Fitzpatrick, 2002; Gray, 1997; Konidari & Abernot, 2006; Scriven, 1991). Patton explains, “Formative evaluation typically connotes collecting data for a specific period of time, usually during the start-up or pilot phase of a project, to improve implementation...” (1997, p.69). By focusing on whether a program is being implemented as intended, formative evaluation seeks to improve the implementation of an intervention and the intervention itself, and by looking at what is and is not working and why, improved interventions can be developed (Kemmerer, 1994). This way,

Information is thus fed back to programme managers early in the implementation process so that necessary adjustments in procedures or content can be made and fed forward into the development of new regulations, new materials, etc., to improve the implementation process and the effectiveness of the intervention. (Kemmerer, 1994, p.64)

¹⁹ One of the best illustrations of this merging of evaluation and quality improvement is the “Shewhart cycle,” a systematic approach to drive continuous improvement in programs. The cycle consists of four never-ending stages: (1) plan, (2) do, (3) study, and (4) act (Evans & Dean, 2003). For more on the Shewhart cycle in educational settings, see Detert, Bauerly Kopel, Mauriel, and Jenni (2000), Detert and Mauriel (1997), and Mullin (2001).

It is the link between formative evaluation and continuous improvement that more often than not shapes the data produced, as well as the intended audience for this data—program managers.

Evaluation results to improve a program (formative evaluations) are fundamentally different from results that render judgment about merit, worth, or significance (summative evaluations). One typically (but not always) aims to provide information needed for program management and improvement and is commonly conducted by internal evaluators whereas the other typically aims to provide information needed to prove accountability and is commonly conducted by external evaluators (Donaldson & Gooler, 2003; Russ-Eft & Preskill, 2001). Therefore, formative evaluations tend to be oriented toward an organization's internal audience while summative evaluations tend to be oriented toward an external audience; "Program management, staff, and sometimes participants tend to be the primary users of improvement-oriented findings, while funders and external decision makers tend to use judgmental evaluation" (Patton, 1997, p.68). What Patton means here is the "intended primary users" of most formative evaluations are program administrators.

Total Quality, Evaluation, and Organizational Culture

Most of the literature that explores the relationship between Total Quality and organizational culture begins with the functionalist assumptions that a Total Quality-based culture (1) would be a unitary phenomenon and something an organization "has," (2) the essence of this culture lies at its "deepest" levels (a la Schein's "underlying assumptions"), and (3) this culture can be changed and controlled by management. Given the predominance of the functionalist approach in the literature on Total Quality-based

organizational cultures, organizational culture has been viewed as a powerful explanatory variable in Total Quality implementation (Chang, 1996; Kezar & Eckel, 2002). Unless the culture of organization is supportive of the principles of Total Quality, the campaign for its integration will fail (Detert, Schroeder, & Mauriel, 2000; Evans & Dean, 2003; Massey, 2003). Sentences like, “If the TQ effort is inconsistent with the organizational culture, it will be undermined” (Evans & Dean, 2003, p.200) or “Culture is the ‘hidden agenda of TQM.’ It is underestimated and frequently overlooked. Culture is subtle yet powerful. It is difficult to measure, but if it is not positive toward TQM then it is surely negative and will exact a costly toll” (Westbrook, 1993, p.3) are easy to find in the literature. Total Quality Management, then, stresses the need to change organizational culture so that behavioral change, based on a change in values and attitudes, is driven by the internalization of the principles of Total Quality by all employees (Bright & Cooper, 1993).

The question then becomes, what are the elements of a Total Quality organizational culture? In “A Framework for Linking Culture and Improvement Initiatives in Organizations,” Detert, Schroeder, and Mauriel (2000) identify the specific values and beliefs underlying Total Quality Management practice. But more importantly, after a qualitative content analysis of the literature and verification by an expert panel, they also synthesize the general dimensions (the values, beliefs, and underlying assumptions) of organizational culture used most commonly in the literature and show how these general dimensions correspond to the implementation of systemic improvement initiatives, Total Quality Management in particular:

Table 2

Model of Total Quality Management Values and Beliefs

A Proposed Model of TQM Values and Beliefs	
(Values and Beliefs Essential to TQM—Overlaid onto Organizational Culture Dimensions)	
Organizational Culture Dimension	TQM Value
1. The basis of truth and rationality in the organization	Decision making should rely on factual information and the scientific method.
2. The nature of time and time horizon	Improvement requires a long-term orientation and a strategic approach to management.
3. Motivation	Quality problems are caused by poor systems—not the employees. Employees are intrinsically motivated to do quality work if the system supports their efforts.
4. Stability versus change/innovation/personal growth	Quality improvement is continuous and neverending. Quality can be improved with existing resources.
5. Orientation to work, task, and coworkers	The main purpose of the organization is to achieve results that its stakeholders consider important. Results are achieved through internal process improvement, prevention of defects, and customer focus.

- | | |
|---|---|
| 6. Isolation versus collaboration/cooperation | Cooperation and collaboration (internal and external) are necessary for a successful organization. |
| 7. Control, coordination, and responsibility | A shared vision and shared goals are necessary for organizational success. All employees should be involved in decision making and in supporting the shared vision. |
| 8. Orientation and focus—internal and/or external | An organization should be customer driven. Financial results will follow. |

Note. From “A framework for linking culture and improvement initiatives in organizations,” by J. R. Detert, R. Schroeder, & J. Mauriel, 2000, *Academy of Management Review*, 25, p. 854.

In building this table, the authors have not only constructed a comprehensive framework for defining and measuring organizational cultures but have illustrated how the dimensions of organizational culture relate to the “ideal culture profile” for Total Quality Management implementation (Chang, 1996; Detert, Schroeder, & Mauriel, 2000). Having constructed this “ideal type,” they suggest that “future research and theory developments should be aimed at understanding the gaps between the culture that is espoused by certain organizational members and the one that actually describes the artifacts and behaviors visible throughout the organization” (Detert, Schroeder, & Mauriel, 2000, p.858). Indeed, the article does not operationalize its theoretical framework, much less provide empirical

evidence for it. Fortunately, this framework and the “culture gap” approach were the foundation for two further articles, both of which integrate organizational culture, Total Quality, and, to give this discussion a focus, educational reform.

In the first article, Detert, Seashore Louis, and Schroeder (2001) use the framework to examine the relationship between organizational culture and the ability to successfully implement quality management programs in schools by focusing explicitly on defining the cultural values underlying Quality Management in U.S. secondary schools. Using the same method as above—a literature review followed by verification by an expert panel—they present several theoretical propositions, based on the value congruence approach in the table above, which represent the underlying core of the quality management philosophy applied to the culture in schools:

Table 3

Total Quality Management Values and their Opposite

TQM Value	Opposite
Value 1: A shared vision and shared goals among faculty, staff and administrators are critical for school success.	Successful schools respect the right of individuals to establish their own vision and goals without regard for higher levels of goals which are often ambiguous and difficult to interpret.

Value 2: Educational needs should be determined primarily by parents, community groups, students, and other stakeholders.	Educational experts should make the important educational decisions.
Value 3: Improving education requires a long-term commitment.	Present pressures—the students in a school and the immediate external demands—are most important; the future is too uncertain to plan for or worry about.
Value 4: A school should strive to make continuous changes to improve education.	Schools should be conservative about making changes.
Value 5: Teachers should be active in improving the overall school operation.	Overall school operations should be left to administrators and a few teacher-leaders.
Value 6: Collaboration is necessary for an effective school.	Professional autonomy with a minimum of cooperation is key to school effectiveness.
Value 7: Decision-making should rely on factual information.	Decision-making should rely on personal, professional experience.
Value 8: Quality problems are caused by poor systems and processes, not by teachers.	The cause of most problems is human error.

Value 9: Quality can be improved with the existing resources. We are doing the best we can with existing resources.

Note. From “A cultural framework for education: Defining quality values and their impact in U.S. high schools,” by J. R. Detert, K. Seashore Louis, & R. Schroeder, 2001, *School Effectiveness and School Improvement*, 12, p. 191-193.

These “value gaps,” with the ideal on one end and the least desirable on the other, begin to operationalize Detert, Schroeder, and Mauriel’s (2000) framework; they are a step forward in proving the claim that an organization’s culture is a key predictor of the organization’s propensity to use Total Quality using empirical evidence. And while the authors are confident in having created a framework for measuring the degree of implementation and institutionalization of Total Quality initiatives—“the context of our research is U.S. high schools, but the general principles could be applied to other educational settings as well” (Detert, Seashore Louis, & Schroeder, 2001, p.183)—the fact is they offer no empirical evidence for their theoretical propositions.

For this reason, in 2003, Detert, Schroeder, and Cudeck reported on the development and validation of a survey instrument for measuring the culture of quality management in K-12 educational settings. In developing *The School Quality Management Culture Survey*, which investigated the relationship between school cultures and the success (or failure) of TQM implementation, multiple methods were used, including theory review, qualitative data analysis, practitioner input, and both exploratory

and confirmatory factor analytic techniques.²⁰ After piloting the survey twice and then revising it, the third version of the survey was given to licensed teachers (N=1,740) in 36 schools in seven districts in five states. Reliability and validity evidence, as well as measures of fit and interpretability, support their claim that the survey provides verified scales for evaluating multiple aspects of a school's TQM culture.

The survey's measures were shared vision, customer focus, long-term focus, continuous improvement, teacher involvement, collaboration, data-based decision-making, systems focus, and, finally, quality at the same cost—all basic elements of TQM. "Value gaps" in perception between espoused values (the "should") and values-in-use (the "is") were constructed around these measures. The following steps were taken in data analysis: (1) in order to do a confirmatory analysis, a "split sample" approach was taken, (2) a restricted factor analysis on the "is" column was built, (3) the output of the restricted factor analysis for the "is" column was interpreted, (4) a restricted factor analysis of the final model on the "should" column was run, and (5) a confirmatory factor analysis was conducted on both "is" and "should" columns. Summary statistics show that, first, the number of residuals with absolute value >0.10 is 10% or less in all cases, second, the root mean square error of approximation is 0.054 or less in all cases, and third, the goodness of fit is 0.90 or greater in all cases.

The survey, designed to "help schools evaluate their cultural alignment with quality management concepts, practices, and principles" employs the value congruence approach (Detert, Schroeder, & Cudeck, 2003, p.308). Instead of the "espoused values"

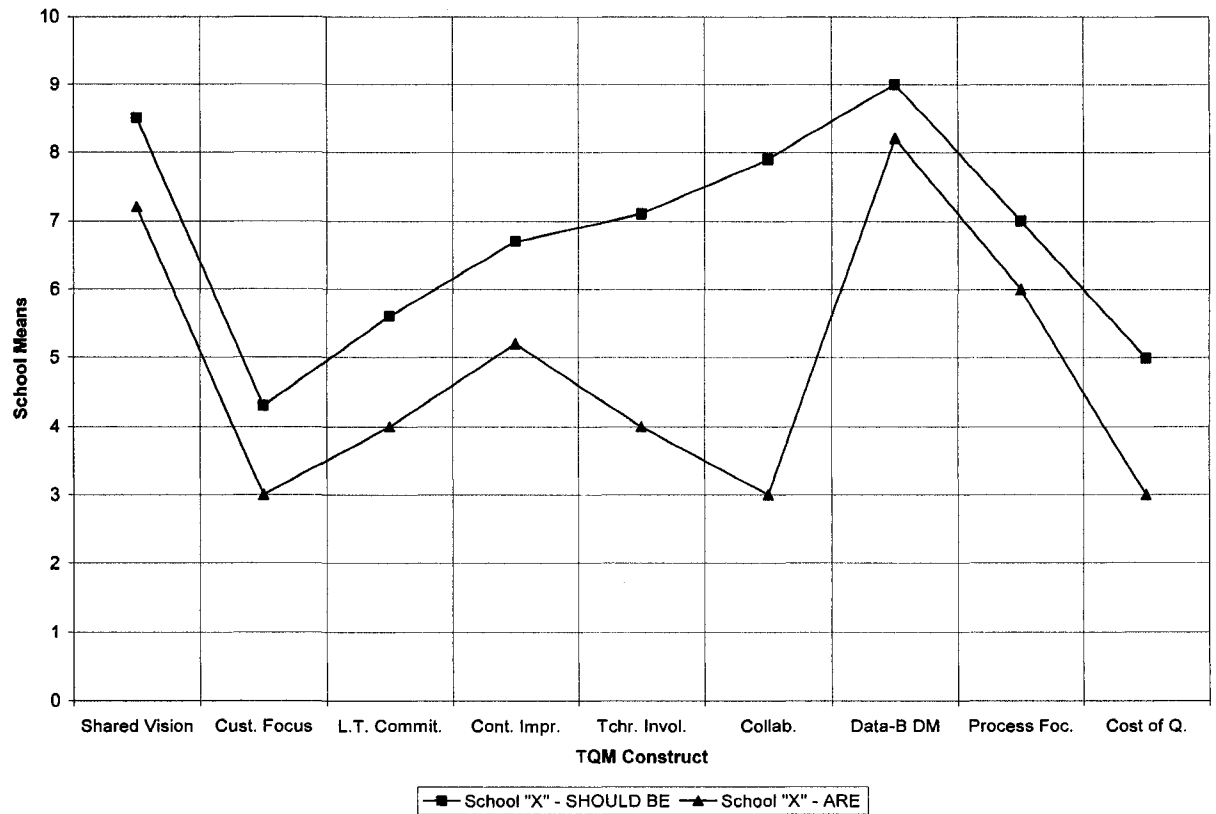
²⁰ Although there are a host of survey instruments that measure school culture and climate (Maslowski, 2006; Roach & Kratochwill, 2004), there are only a few survey instruments that focus on quality improvement (Roach & Kratochwill, 2004).

versus “values-in-use” language, however, it slightly tweaks this language towards the is-ought distinction:

...respondents were asked to indicate their level of agreement with each question twice—first indicating how things in their school actually ARE as regards the question, and second indicating how they think things SHOULD BE as regards that question, even if quite different from how things actually are at present in their school. Responses in the ARE column were aimed at assessing the behavioral norms present in a school, whereas the SHOULD BE column responses were intended to get at underlying values or beliefs. (Detert, Schroeder, & Cudeck, 2003, p.311)

Exploration of the size of the gaps between what “is” (values-in-use) and what “should be” (espoused) in employees’ minds regarding key cultural values enabled the development of “gap charts”:

Figure 2

Gap Chart Example

Note. From “The measurement of quality management culture in schools: Development and validation of the SQMCS,” by J. R. Detert, R. Schroeder, & R. Cudeck, 2003, *Journal of Operations Management*, 21, p. 312.

With this, the authors offer three tools related to Total Quality, evaluation, and organizational culture: (1) well-developed theoretical frameworks for measuring the culture of Total Quality in educational organizations, (2) a valid and reliable survey instrument rooted in these theoretical frameworks, and tailored to educational

organizations, and (3) a “gap chart” technique that can visually capture evaluative data. In total, these authors offer good methodological beginning points for further research.

Before moving on, let me summarize this discussion related to evaluation utilization thus far. This section began with background on evaluation utilization as, of the three evaluation traditions, it is this tradition that has integrated organizational research to the fullest extent. The evaluation utilization tradition is defined by its concern with the way in which evaluation information will be used and its focus on those who will use the information (e.g., “intended primary users”); the tradition rests on pragmatic tenets. As the tradition has evolved, so have its tools. Expansion of the concept of “use” led to the ideas of conceptual and process use, which became fundamental to participatory, empowerment, and developmental evaluation, all of which draw on developments within organizational research, such as organizational learning, development, and change. Despite these advances, work that integrates organizational research and the evaluation utilization tradition is underdeveloped, as evidenced by few empirical studies and a lack of clarity and specificity about the role of elements, such as organizational culture, that make up an organization’s “infrastructure.”

Evaluation’s current period of growth in the U.S. is tied to the “quality movement,” which was sparked by neoliberal policies, emanating from all levels of government beginning in 1980, that stressed external accountability and internal quality. By the early 1990s, quality improvement campaigns within institutions of higher education were taking place, and new management-based improvement tools and strategies, such as Total Quality Management, were introduced. The Total Quality philosophy, especially the principle of continuous improvement, supported the use of

program evaluation as a tool for improvement; of the types of program evaluation, formative evaluation has linked well with the principle of continuous improvement. At the organizational level, evaluation of Total Quality-based culture has been supported by the development of theoretical frameworks that construct an “ideal culture profile,” allowing for a “value gaps” approach, both generally and within educational contexts. While this work provides a methodological foundation to build upon in tying together Total Quality, evaluation, and organizational culture, little empirical evidence exists for these frameworks, and the single valid survey that has been developed is tailored for K-12 school environments.

There are a couple of problems, then, with the literature as it stands now. First, as I have already argued, while there has been progress in integrating organizational research into the evaluation utilization literature, this integration is not deep, as evidenced by the lack of empirical research. Consequently, there is a weak empirical base on which conclusions about evaluations rooted in organizational research can be made. Second, research that connects evaluation utilization, organizational culture, and the principles of Total Quality is also underdeveloped. One of the more obvious points to be made here is that the frameworks and survey research developed thus far have focused on K-12 school environments, leaving room for research on higher education environments. *However, what simplifies resolving these problems is the fact that examining an organizational culture supportive of Total Quality would be an excellent spring board, or jumping off point, to studying a student learning assessment-supportive organizational culture in higher education.*

There are two main reasons for this claim. First, again, researchers have provided well-developed theoretical frameworks for measuring the culture of Total Quality in educational organizations, a valid and reliable survey instrument rooted in these theoretical frameworks and tailored to educational organizations, and techniques that can visually capture evaluative data—the tools have been created. Second, at a fundamental level, the principles of Total Quality and the principles of the collegiate assessment movement are nearly identical, *which would allow for the use of these tools*. To demonstrate how these two sets of principles are nearly identical, below is a comparison between the principles of good practice for assessing student learning (American Association for Higher Education, 1992; Banta, 1997; Banta, Lund, Black, & Oblander, 1996) and the elements of Total Quality (Chang, 1996; Detert, Bauerly Kopel, Mauriel, & Jenni, 2000; Detert & Mauriel, 1997; Detert, Schroeder, & Cudeck, 2003; Evans & Dean, 2003; Massey, 2003):

Table 4

Principles of Good Practice for Assessing Student Learning Compared to the Basic Elements of Total Quality

Principles of Good Practice for Assessing	
Student Learning	Basic Elements of Total Quality
1. The assessment of student learning begins with educational values.	Shared vision; customer focus; leadership and strategic planning

2. Assessment is most effective when it reflects an understanding of learning as multidimensional, integrated, and revealed in performance over time.	Long-term focus; process orientation; systems focus; continuous improvement and learning
3. Assessment works best when the programs it seeks to improve have clear, explicitly stated purposes.	Leadership and strategic planning; shared vision
4. Assessment requires attention to outcomes but also and equally to the experiences that lead to those outcomes.	Process orientation; systems focus; management by fact (evidence)/data-based decision-making
5. Assessment works best when it is ongoing not episodic.	Continuous improvement and learning; long-term focus
6. Assessment fosters wider improvement when representatives from across the educational community are involved.	Empowerment and teamwork; stakeholder involvement; collaboration
7. Assessment makes a difference when it begins with issues of use and illuminates questions that people really care about.	Customer focus; empowerment and teamwork; collaboration; leadership and strategic planning
8. Assessment is most likely to lead to improvement when it is part of a larger set of conditions that promote change.	Leadership and strategic planning; systems focus; shared vision; process orientation; continuous improvement and learning

9. Through assessment, educators meet responsibilities to students and to the public.	Management by fact (evidence)/data-based decision-making; quality at the same cost; customer focus
10. Assessment is most effective when undertaken in an environment that is receptive, supportive, and enabling.	Empowerment and teamwork; leadership and strategic planning; shared vision; stakeholder involvement; collaboration

The issue becomes one not about agreement among the principles, but one about language: “Whereas assessment could ultimately adopt the discourse of scholarship, Total Quality never shed its corporate flavor—especially in the eyes of skeptical faculty” (Ewell, 2002, p.20). Total Quality, in the end, has not fared well on campuses primarily due to its alien language (e.g., students as “customers”). Others have commented on the language of Total Quality either not achieving widespread acceptance or being resisted, by teachers and faculty in particular (Detert, Bauerly Kopel, Mauriel, & Jenni, 2000; Detert, Schroeder, & Cudeck, 2003; Detert, Seashore Louis, & Schroeder, 2001; Ewell, 1999; Guskin, 1996; Kells, 1995; Konidari & Abernot, 2006; Palomba & Banta, 1999). “In spite of these criticisms, however,” write Detert, Seashore Louis, and Schroeder (2001),

Quality Management in various guises *has* invaded the educational world. “QM” has been replaced in many educators’ language with terms such as “continuous improvement,” “quality assurance,” or “knowledge work supervision” (Duffy, 1997) to avoid the impression that it is an attempt to deny agency to teachers and students; Stringfield (1995) prefers the term

“high reliability schools.” Whatever it is called, the elements of QM as espoused by Deming, Juran, and others map well on to the current context of school policy. (p.184)

Educational policies that reflect the essence of Barr and Tagg’s “learning paradigm” (1995), for example, are promoting a shift toward student centeredness and individual student learning and development, a shift that in other contexts would be labeled “customer focus.” The increasing use of benchmarking, in looking at undergraduate retention and graduation rates for example (Carey, 2004, 2005), reflects Total Quality principles. The general calls for “educational effectiveness” (Western Association of Schools and Colleges, 2001), an “improvement-oriented ethos” (Kuh, Kinzie, Schuh, & Whitt, 2005a), a “culture of evidence” (Dowd, 2005; Ewell, 2004), or the popularity of the “learning organization” concept in higher education (Ingram, Seashore Louis, & Schroeder, 2004; Konidari & Abernot, 2006; Terziovski, Howell, Sohal, & Morrison, 2000) reflect the impact of Total Quality in higher education. Indeed, although its language may have been stripped and rejected, the principles of the quality movement have provided a much-needed foundation for the assessment movement in higher education (Kells, 1995; Marchese, 1994; Palomba & Banta, 1999).

Conclusion

Several summative statements can be made about the collegiate student learning assessment, organizational culture, and evaluation utilization literatures. First, the collegiate assessment literature is divided by two purposes, external accountability and internal improvement. Assessment in the interest of internal improvement originated from within the academy among those who insisted on the primacy of student learning,

the mission of higher education. As the literature on improving student learning developed, it grew to become stratified; “assessment research” blossomed at the classroom, course, program (including general education), institutional, and university system levels (Miller & Leskes, 2005). At the institutional level, the particularly important role of institutional administrators has been highlighted, as this group holds “legitimate power” and controls the internal resource allocation process. Surprisingly then, there exists little research on the specific role of the provost, even while it is known that assessment activities and the provost’s office are tightly connected. Complicating this problem of a lack of research is the fact that most “research” on institutional administrators building support for assessment is descriptive and prescriptive. A gap in the literature exists in research rooted in empirical evidence and theoretical frameworks on the role of the provost in building support for assessment.

Second, in the conversation about building support for assessment at the institutional level, organizational culture has received the most attention (among other theoretical approaches at the organizational level) as, ultimately, cultural changes are perceived to be the only path to institutionalizing assessment initiatives. Central to these changes, it is claimed, are institutional administrators, who can shape campus culture. But is this “shaping” possible? How effectively can administrators shape organizational culture and, if they can effectively do so, whose interests are served? Answers to the questions vary according to paradigmatic stances within the organizational culture literature; this literature is clustered into research inspired by functionalism, critical theory, and postmodernism. This study operates with the assumption that organizational culture is a variable, and is most inspired by the functionalist tradition.

Third, of the three evaluation traditions, the evaluation utilization tradition has integrated organizational research to the fullest extent. Participatory, empowerment, and developmental evaluation are examples of this point. However, the synthesis of organizational research and the evaluation utilization research stands underdeveloped, as evidenced by the dearth of empirical studies integrating the two and by a lack of clarity and specificity about the role of elements within an organization, such as organizational culture, in the studies that do exist. Research based in the Total Quality philosophy would help resolve this state of underdevelopment. The underlying principles of Total Quality and higher education's assessment movement are nearly identical, allowing for the use of well-developed theoretical frameworks and methodological tools that tie together the Total Quality, evaluation, and organizational culture. Unfortunately, not enough empirical evidence backs these frameworks and tools, leaving a gap in the literature.

To address the need for empirically- and theoretically-based research on the role of the provost in building institutional support for assessment for the sake of internal improvement, to address the need for organizational culture research rooted in theory, to address the need for an evaluation rooted in the principle of continuous improvement and attention to administrators as "intended primary users," and to address the underdevelopment of the evaluation utilization tradition while using what applicable frameworks and tools that do exist, I propose a formative evaluation of provosts' effectiveness in building a student learning assessment-supportive organizational culture at their campuses. This study, to conclude, addresses gaps within three bodies of literature while simultaneously bringing areas of these literatures together in the interest of organizational improvement through education reform.

CHAPTER 3: RESEARCH DESIGN AND METHODOLOGY

Purpose

This study justified the initiation of a formative evaluation. It rested on the claim that program evaluation can be considered a tool for continuous improvement, a tool for incrementally improving quality in particular (Evans & Dean, 2003; Konidari & Abernot, 2006; Light, 2001; Mark, 2003; Russ-Eft & Preskill, 2001) and that, of the different evaluation techniques, formative evaluations are especially linked to the idea of continuous improvement. In fact, the intent to improve and formative evaluation have been wedded to such an extent that the two terms—“formative evaluation” and “continuous improvement”—are oftentimes used interchangeably (Donaldson & Gooler, 2003; Ewell, 1993b; Fitzpatrick, 2002; Gray, 1997; Konidari & Abernot, 2006; Scriven, 1991). Patton explains, “Formative evaluation typically connotes collecting data for a specific period of time, usually during the start-up or pilot phase of a project, to improve implementation...” (1997, p.69). By focusing on whether a program is being implemented as intended, formative evaluation seeks to improve the intervention itself and its implementation, and by looking at what is and is not working and why, improved interventions can be developed (Kemmerer, 1994). This way,

Information is thus fed back to programme managers early in the implementation process so that necessary adjustments in procedures or content can be made and fed forward into the development of new regulations, new materials, etc., to improve the implementation process and the effectiveness of the intervention. (Kemmerer, 1994, p.64)

The point here is that the link between formative evaluation and continuous improvement not only shapes the data produced, but also shapes the intended audience for this data—program managers. Formative evaluations tend to aim toward providing management with information needed for improvement.

Scope, Focus, and Units of Analysis

The scope of this evaluation was the 23 campuses of the California State University system, specifically their organizational cultures. The implications of this evaluation were for the entire CSU system as all CSU provosts were included. Below are the 23 names, campuses, and appointment dates of CSU provosts/vice presidents for academic affairs at the time of this study:

- Soraya M. Coley, California State University Bakersfield, appointed 2005;
- Theodore D. Lucas, California State University Channel Islands, appointed 2003;
- Sandra Flake, California State University Chico, appointed 2007;
- Sam Wiley, California State University Dominguez Hills, appointed 2007;
- Michael K. Mahoney, California State University East Bay, appointed 2007;
- Jeronima Echeverria, California State University Fresno, appointed 2004;
- Ephraim P. Smith, California State University Fullerton, appointed 1998;
- Robert Snyder, Humboldt State University, appointed 2007;
- Karen Gould, California State University Long Beach, appointed 2007;
- Desdemona Cardoza, California State University Los Angeles, appointed 2007;
- Don Zingale, California Maritime Academy, appointed 2004;
- Kathryn Cruz-Urbe, California State University Monterey Bay, appointed 2007;
- Harry Hellenbrand, California State University Northridge, appointed 2004;

- Herman D. Lujan, California State Polytechnic University Pomona, appointed 2007;
- Joseph Sheley, California State University Sacramento, appointed 2006;
- Louis A. Fernandez, California State University San Bernardino, appointed 1994;
- Nancy A. Marlin, San Diego State University, appointed 1998;
- John Gemello, San Francisco State University, appointed 2003;
- Carmen Sigler, San Jose State University, appointed 2005;
- William W. Durgin, California Polytechnic State University San Luis Obispo, appointed 2006;
- Emily F. Cutrer, California State University San Marcos, appointed 2006;
- Eduardo M. Ochoa, Sonoma State University, appointed 2003; and
- William A. Covino, California State University Stanislaus, appointed 2006.

This, then, was a “multisite evaluation,” in which provosts were evaluated at different geographical locations in a coordinated effort to address a core set of questions (Herrell & Straw, 2002; Turpin & Sinacore, 1991). A multisite evaluation, through involving multiple sites and cross-site evaluation activity, facilitates the generalizability of findings (Herrell & Straw, 2002; Turpin & Sinacore, 1991). Additionally, collecting data from multiple campuses increased overall sample size thus aiding in determining provost effectiveness. Finally, multisite evaluations should take into account the developmental stage of an intervention (Herrell & Straw, 2002) which in this case translated into appreciating the equal impact of CSU Chancellor-level policies and practices on campuses within the system.

The focus here was provosts' effectiveness in building a student learning assessment-supportive culture. Unfortunately, there are no empirically-tested instruments of a student learning assessment-supportive organizational culture. As a result, the focus here shifted to the principles of Total Quality, a philosophy of management aimed at comprehensive, organization-wide improvement in the quality of products and services, as well as a collection of tools and strategies for implementing this philosophy (Evans & Dean, 2003).²¹ I purport that examining an organizational culture supportive of Total Quality would be an excellent spring board, or jumping-off point, to studying a student learning assessment-supportive organizational culture.

The eight elements of a Total Quality culture measured in this evaluation were: (1) shared vision (e.g., "a shared vision and shared goals among faculty, staff, and administrators are critical for university success"); (2) long-term focus (e.g., "improving education requires a long-term commitment"); (3) continuous improvement (e.g., "a university should strive to make continuous changes to improve education"); (4) involvement (e.g., "faculty should be active in improving the overall university operation"); (5) collaboration (e.g., "collaboration is necessary for an effective university"); (6) data-based decision-making (e.g., "decision-making should rely on factual information"); (7) systems focus (e.g., "quality problems are caused by poor systems and processes, not by faculty"); and (8) quality at the same cost (e.g., "quality can be improved with the existing resources").

²¹ There are six basic elements to the innumerable approaches to Total Quality: (1) customer focus, (2) a process orientation, (3) empowerment and teamwork, (4) continuous improvement and learning, (5) management by "fact" (evidence), and (6) leadership and strategic planning (Chang, 1996; Detert, Bauerly Kopel, Mauriel, & Jenni, 2000; Detert & Mauriel, 1997; Evans & Dean, 2003; Massey, 2003).

There are two main reasons for focusing on a Total Quality culture. First, researchers have provided well-developed theoretical frameworks for measuring the culture of Total Quality in educational organizations (Detert, Seashore Louis, & Schroeder, 2001; Detert, Schroeder, & Mauriel, 2000), a valid and reliable survey instrument rooted in these theoretical frameworks and tailored to educational organizations (Detert, Schroeder, & Cudeck, 2003), and techniques that can visually capture evaluative data (Detert, Schroeder, & Cudeck, 2003). That is, the tools have been created. Second, as I have argued, the principles of Total Quality and the principles of the collegiate assessment movement are nearly identical at a fundamental level, which would allow for the use of these tools.

Finally, if the unit of analysis is the major entity under investigation, then the general unit of analysis here was organizational culture, a student learning assessment-supportive organizational culture specifically. However, there were different analyses in the evaluation which led to two well-defined units of analysis. First, analysis on campuses meant individual campuses were a unit; second, because averages across campuses—as opposed to individual campus data—were analyzed, the CSU system as a whole was a unit.

Sampling Strategies

In this evaluation, the focus was on a culture of assessment at the institutional level, which included the program (both the majors and general education) and institutional levels, and led to the selection of academic affairs administrators as the population to be sampled. A single survey of academic affairs administrators with both administrative and academic appointments who are involved with assessment policy and

practice at the campus level—presidents, provosts, associate provosts, deans, and associate deans—was conducted.

It is these individuals who were asked to provide administrative perceptions of gaps between provosts' values-in-use and ideal values regarding building a student learning assessment-supportive culture. This study, then, combined two purposeful sampling strategies: homogeneous sampling across campuses and stratified purposeful sampling both across and within campuses (Patton, 2002). Selection of these strategies was rooted in data triangulation, or collecting data from a variety of sources (Patton, 2002; Russ-Eft & Preskill, 2001). These strategies were also “purposive” in that they involved the selection of specific individuals because of their position, experience, knowledge, or attitudes (Russ-Eft & Preskill, 2001). Finally, it is important to note that “time sampling” was relevant here as “programs, organizations, and communities may function in different ways at different times during the year” (Patton, 2002, p. 229); therefore, the survey was conducted during the fall semester when the above administrators were more likely to be on campus.

Analytical Approach, Data Collection, and Data Management and Analysis

The survey conducted for this study was quantitative in nature; the questions precisely measured the “value gaps” in perception between provosts' ideal values and values-in-use. (See Appendix C for the survey instrument adapted from Detert, Schroeder, and Cudeck [2003]). All of the questions employed rating scales, with each statement followed by, “How effective is your provost/VP for academic affairs in building an organizational culture supportive of this statement?” Respondents were to select a number between 1 and 10 with 1 representing “minimally effective” and 10

representing “very effective.” Respondent data from an average of three statements were used to measure each Total Quality culture construct (see Appendix D). Only constructs with data collected on each of their items was included in the final analysis.

The survey was disseminated in both paper and electronic forms, with the electronic form relying on SurveyMonkey’s platform. Data from both forms was analyzed using Microsoft Excel and SPSS. Microsoft Excel was used to construct “gap charts,” line graph diagrams that showed the size of the gaps between what “is” (values-in-use) and what “should be” (the ideal) (Detert, Schroeder, & Cudeck, 2003). SPSS was used to report descriptive statistics, such as response rate, means, standard deviations, and ranges of scores and to run one-sample t-tests. T-tests examined if the means differed from the ideal; in this case, the means were the ranking of provost effectiveness by the academic affairs administrators and the population mean, or test value, was the ideal of effectiveness (a score of 10 represents “very effective”). This type of test was needed to ascertain if there was a gap between ideal provost effectiveness and actual provost effectiveness, as ranked by academic affairs administrators.

To explore the campus differences, seven groups were formed using the following institutional characteristics: Carnegie classification type, total enrollment, number of academic programs, length of provost tenure, number of academic affairs administrators, percent of graduate students, and age of the institution (Knight, 2003). Using SPSS, one-way, between-subjects analyses of variance were run to find significant group differences. The one-way analysis of variance was used to test the null hypothesis that several (more than two) population means are equal; it was appropriate in this case because each of the seven groups has more than two subcategories to be compared. I

chose to compare mean rankings of provost effectiveness across the seven groups to investigate which, if any, institutional characteristics had an effect on academic affairs administrator perceptions. Finally, multiple linear regression was used to evaluate how closely associated the selected institutional characteristics were with the Total Quality constructs.

Project Management, Logistics, and Survey Procedures

Survey research typically consists of six stages: (1) planning and survey development; (2) field pretesting; (3) final survey design and planning; (4) implementation of survey and data collection; (5) data coding and data-file construction; and (6) research and analysis of data (Scheuren, 2004).

Field pretesting took place summer of 2007 and consisted of the survey being sent to all University of San Diego School of Leadership and Education Sciences (SOLES) Leadership Studies faculty for improvement-oriented suggestions. Also in summer of 2007, I finalized survey design and planning by making revisions based on feedback from SOLES faculty in addition to engaging CSU provosts (the “intended users” of the evaluation). In fall of 2007, survey distribution and data collection commenced.

Implementation, the core of which is ensuring a high response rate, consisted of a mixed-mode administration (Porter, 2004). First, a paper-based cover letter and survey was mailed. Second, two follow-up e-mails were sent directing administrators to the web-based version of the survey and offering replacement questionnaires. Data coding and data-file construction, along with analysis of data, continued through spring of 2008.

Delimitations and Limitations

Research Methodology

Survey research has numerous limitations, the primary being it is not as effective as other methods in collecting many types of social science data, oftentimes not penetrating very deeply below the surface (Kerlinger, 1986). It offers little information on how contexts shape behavior or perceptions, is impersonal, not well suited for subtle or sensitive issues and presents limited opportunities for further probing, and carries the risks of nonresponse, errors in coverage, sampling, and measurement, biased or invalid answers, and overreliance on standard measures and models (Harrison, 2005; Salant & Dillman, 1994).

A specific limitation of this proposed survey that influences validity is social desirability response bias (Tuckman, 1999). The same reason why provosts deserve attention—their “legitimate power” that grants them authority—impacts the extent to which respondents want to show themselves (and their provost) in a good light and the extent to which respondents attempt to anticipate what others (again, their provost) want them to conclude. This evaluation, undoubtedly, can be considered an example of “studying up” (Nader, 1972), as it focused on the most powerful administrator in a university’s academic affairs division. As a result, it carried with it the most salient limitation of “elite studies”—accessibility (Hertz & Imber, 1995; Moyser & Wagstaffe, 1987). The same power that makes provosts a subject worthy of study also gives them the capacity to hire “organizational gatekeepers,” to erect barriers and hurdles that block whom they wish, or, in this case, to influence others or even encourage them not to complete the survey.

Theoretical Frameworks

The governing theoretical frameworks here are rooted in the three bodies of literature this study draws from: collegiate student learning assessment, organizational culture, and evaluation. Before an explication of this, a discussion of the importance of a “home perspective” (Martin, 1992, 2002) is needed. To resolve the potentiality of moving aimlessly moving from one “lens” to another, a “dominant frame” ought to be selected with related “supporting frames” playing a subsidiary role (Morgan, 2006; Patton 1997). Rubin and Rubin explain,

...especially when you feel lost, having a compass—a research philosophy—is useful because it provides guidance, suggests what to pay attention to, and alerts you to problems that may arise. In addition, a guiding philosophy provides legitimacy by helping to explain and justify why you have picked a particular set of research tools and designed a project in a particular manner. (2005, p.37)

Within the above three bodies of literature, all of which contain multiple sets of philosophical approaches within them, this study finds a “home perspective” in the internal improvement motive within the assessment literature, the functionalism-inspired research within the organizational culture literature, and the evaluation utilization tradition within the evaluation literature.

The first delimitation comes from within the collegiate student learning assessment literature; the delimitation here is research and theory focusing on assessment in the name of internal improvement (not external accountability). Seeing student learning and personal development as the mission of higher education, this perspective

argues for a shift away from faculty, traditional research, and instruction and toward students, scholarship, and learning (Angelo, 2002; Boggs, 1999; O'Banion, 1997, 1999; Tagg, 2003). The limitation of this focus on internal improvement stems from the fact that student assessment at any level, from the classroom to the university system level, can not be fully understood without attention to external accountability. This is evidenced by the finding that, at the institutional level, preparing for accreditation was the highest-rated institutional purpose for engaging in student assessment (Peterson & Vaughan, 2002) and that, in many cases, accrediting associations have replaced the state as the primary stimulus for assessment (Ewell, 1993a; Wright, 2002). The point here is the dialectic between the internal improvement and external accountability purposes have and continue to shape the student assessment movement; the movement, or any research related to it, can not be complete understood without acknowledging the tensions between these two purposes.

The second delimitation comes from within the organizational culture literature; the delimitation here is research inspired by functionalist theory. Organizational culture researchers inspired by functionalism typically do not separate the symbolic from the functionalist aspects of culture, defining culture as that which is shared and treating culture as a variable that can be manipulated for effectiveness. They have been informed by theories in the social sciences, in anthropology and sociology especially, that hold that all aspects of the social structure serve a purpose and are indispensable in that they are linked to the survival of society (Donaldson, 2003). To the claim that organizational culture is a unitary phenomenon, where culture is that which is shared throughout the organization, critics point to inconsistencies, disruptions, conflicts, and ambiguities in

organizational life (Martin, 2002), and to the approach's tendency to serve managerial interests. That is, although these researchers know unanimous agreement on culture is unlikely, they nevertheless impose an organization-wide consensus, thus imposing one perspective—that of management—over all others (Bush, 2003; Martin, 1992). In reifying a view of culture as homogeneous and harmonious, they legitimate organizational practices that ignore, downplay, or exclude the ideas, opinions, and interests of those who individually or collectively deviate from the supposedly dominant view, which stands as the “objective truth” (Martin, 1992, 2002; Martin, Frost, & O'Neill, 2004).

The third delimitation comes from within the evaluation literature; the delimitation here is evaluation utilization research and theory. Evaluation utilization sees the primary purpose as providing information for decision making, always asking how utilizable it is for decision makers. To Patton, an evaluator must seek out individuals who are likely to be real users (“intended primary users”) of the evaluation and success is measured by usefulness to and actual use by senior administrators for decision making and institutional improvement (Patton, 1997). The primary limitation of a “home perspective” in evaluation utilization stems from the threat of managerialism, where placing the interests of managers above all else can easily translate into an evaluation being a tool for this group (Alkin & Christie, 2004). In this case, this translates into management deciding on the value of the evaluation and deciding how to use the resulting information, oftentimes deciding that managerial views are generally shared by other members of the whole organization (Martin, 2002). This problem of managerialism is linked to the claim that evaluation utilization's pragmatic basis leads to operating

without a moral compass (Henry, 2000; Henry & Mark, 2003; House, 1995; Mark & Henry, 2004; Pawson & Tilley, 1997). In focusing on the acceptability of ideas, House finds utilization-focused evaluation to have a value-free doctrine at its base,

which amounts to the evaluator's not taking responsibility for ethical and moral judgments but rather substituting the values of clients, the powerful, or participants as the basis for the evaluation, or by ignoring the issue altogether through seeking refuge in methodology or the contract.

(1995, p.31).

One solution to this, to conducting research with integrity, meaningfulness, and coherence, is to find a "pragmatic anchor" in the set of ethics governing a university's Institutional Review Board (Greene & Caracelli, 1997). Accordingly, this study was approved by University of San Diego's Institutional Review Board.

CHAPTER 4: FINDINGS

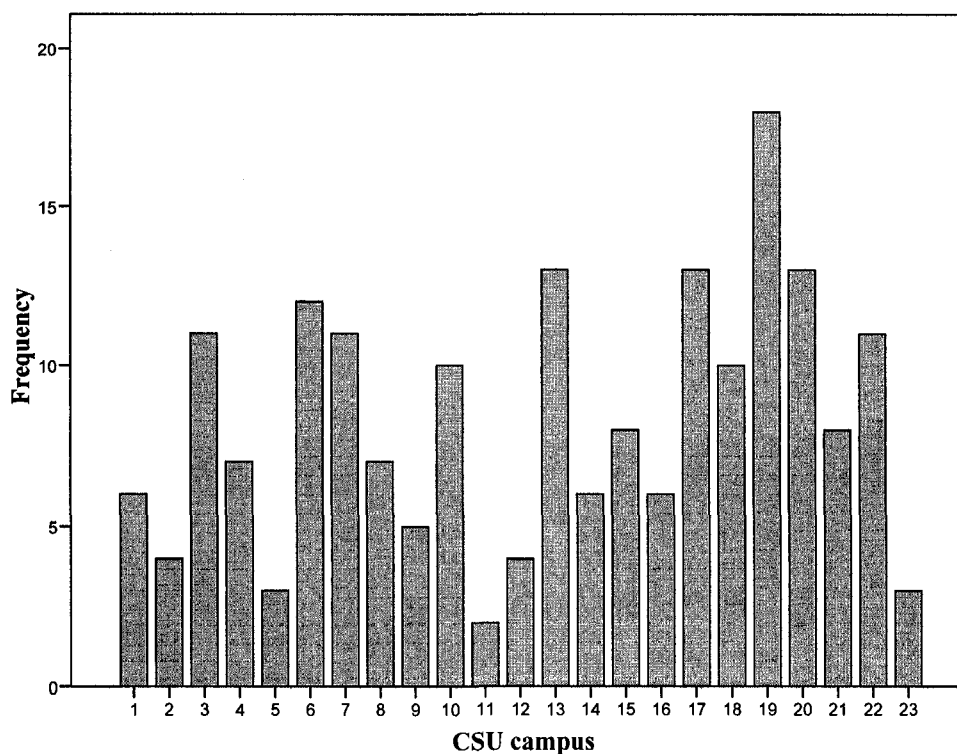
This chapter is divided into two sections, one that explored the survey data through descriptive statistics and a second that employed inferential statistics to make conclusive statements. In the first section, the mean, spread, and shape of the distribution are elaborated upon. Overall, it was found, provosts are more effective than not with a score of 7.11 where “10” represented “very effective” in building an assessment-supportive culture. The second section, named “Research Analysis,” used t-tests, analyses of variance, and regression techniques to find “value gaps” and differences along the lines of seven institutional characteristics. Among these institutional characteristics, total enrollment emerged as having significant power in explaining provost effectiveness in building an assessment-supportive culture.

Exploratory Data Analysis

A total of 453 administrators from all 23 CSU campuses were invited to participate in the survey and 195 responded, defined here as answering enough questions to compute an average for at least one construct.²² This translates into a 43% response rate. Administrators from every CSU campus were represented. An average of eight administrators responded from each campus (out of an average of approximately 20 administrators invited to participate per campus). The most respondents from one campus numbered 18; the lowest number of respondents from a campus was two. Only four (2%) administrators who responded did not indicate their campus. Below is a bar chart showing the distribution of responses among CSU campuses:

²² 224 administrators responded in some way to the survey, but 26 respondents were excluded for not providing enough information and three potential respondents chose not to provide consent.

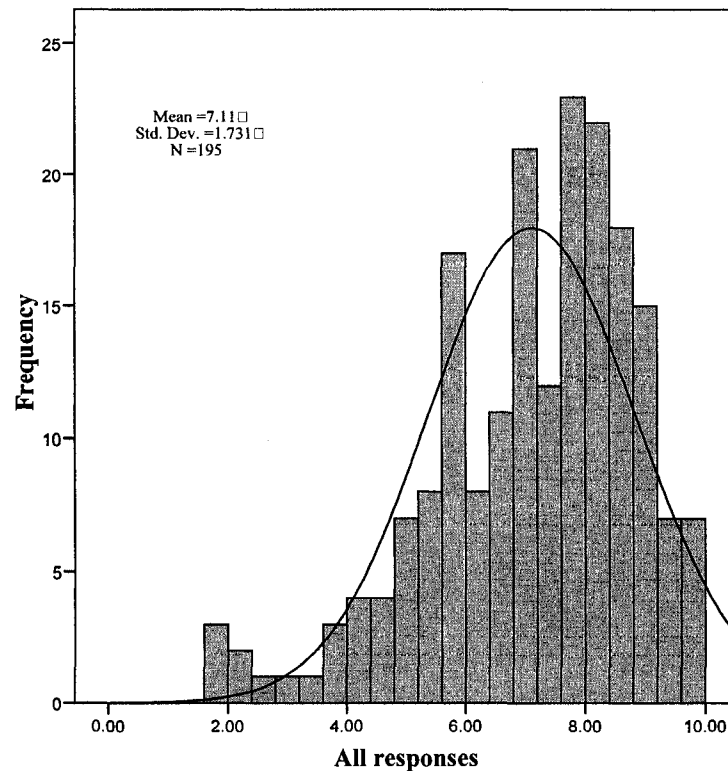
Figure 3

Bar Graph of Campus Responses to Survey

Considering that all administrators received a paper-based survey and two follow-up e-mails directing administrators to the web-based version of the survey, it is interesting to note that the majority of respondents (72%) chose to answer the paper-based questionnaire as opposed to the electronic-based version.

On a scale from 1 to 10, the overall mean for all responses was 7.11, with a standard deviation of 1.73. The distribution of responses was negatively skewed (skewness = $-.86$), meaning the tail of the distribution pointed toward the lower end of the score continuum, the mean was less than the median, and the distribution more peaked than the normal distribution (kurtosis = $.65$):

Figure 4

Histogram of All Responses

Recall that the survey examined eight constructs fundamental to Total Quality Management: shared vision, long-term focus, continuous improvement, involvement, collaboration, data-based decision-making, systems focus, and quality at the same cost. Means, standard deviations, and percent change compared to the overall mean for each of these constructs are below:

Table 5

Descriptive Statistics for Each Total Quality Construct

Construct	N	Mean	Standard Deviation	Percent change compared to overall mean (7.11)
Shared vision	186	7.63	1.99	+7.3%
Long-term focus	164	7.11	1.91	0%
Continuous improvement	184	7.13	1.87	+0.3%
Involvement	185	7.41	1.91	+4.2%
Collaboration	186	6.47	2.05	-9.8%
Data-based decision-making	178	7.12	1.98	+0.1%
Systems focus	178	7.04	1.92	-1.0%
Quality at the same cost	179	6.21	2.11	-12.7%

Notice that there are two constructs—collaboration and quality at the same cost—where the means are below the overall mean and standard deviations larger. The distribution of scores for these two constructs are shown in the histograms presented in Figures 5 and 6.

Figure 5

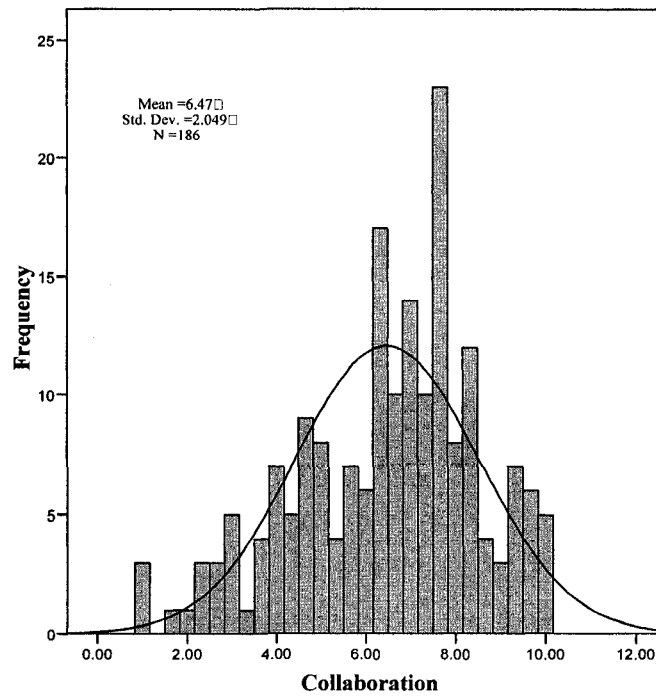
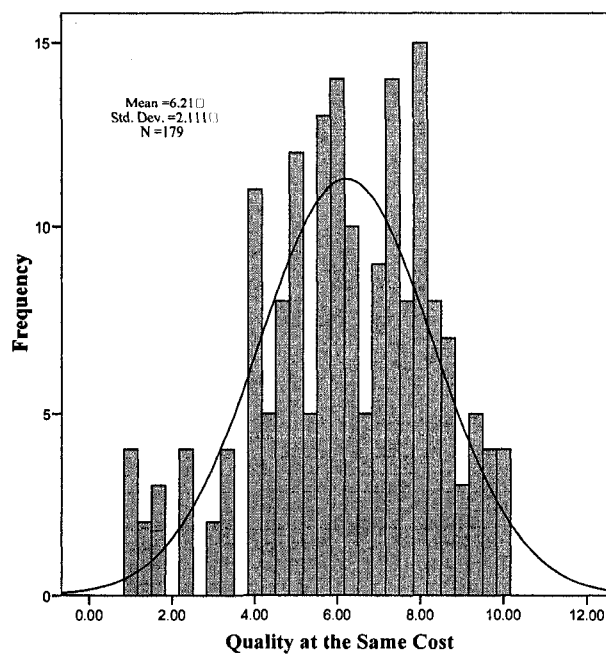
Histogram of Responses to Total Quality Construct—Collaboration

Figure 6

Histogram of Responses to Total Quality Construct—Quality at the Same Cost

When compared to the distributions of the other constructs, the distributions for these two constructs differ: their distributions suggest higher variability amongst scores.

To summarize, with a respectable response rate and responses from every campus, the survey offers a picture of perceptions of the effectiveness of provosts in building cultures of assessment across the California State University system. Descriptive statistics show that, overall, provost effectiveness was 7.11, with 1 representing “minimally effective” and 10 representing “very effective;” the distribution of scores leans more toward 10 than 1. The shape of the distribution for two constructs, collaboration and quality at the same cost, was different than the distribution for the other constructs.

Research Analysis

This study was guided by one overarching question: How can CSU provosts be more effective in building a student learning assessment-supportive culture at their campus? There were two subsidiary questions: What, if any, are the gaps between provosts’ values-in-use and ideal values regarding building a student learning assessment-supportive culture? What are the differences and similarities in “value gaps” among CSU campuses regarding provosts’ building a student learning assessment-supportive culture? One-sample t-tests were used to examine if means differed from the ideal; in this case, the means were the ranking of provost effectiveness by the academic affairs administrators and the ideal of effectiveness is represented by a score of “10”. This type of test was needed to ascertain if there were gaps between the ideal and actual level provost effectiveness, at least ranked by academic affairs administrators.

For all respondents, results indicated that the overall mean did differ significantly from the ideal of “10” ($t(194) = -23.33, p \leq .005$), meaning there is enough evidence to conclude that there is a gap between ideal and actual provost effectiveness. Additionally, a one-sample t-test was used to examine whether the construct means differed from the ideal, finding, in every case, significant differences:

Table 6

T-tests Results for Each Total Quality Construct

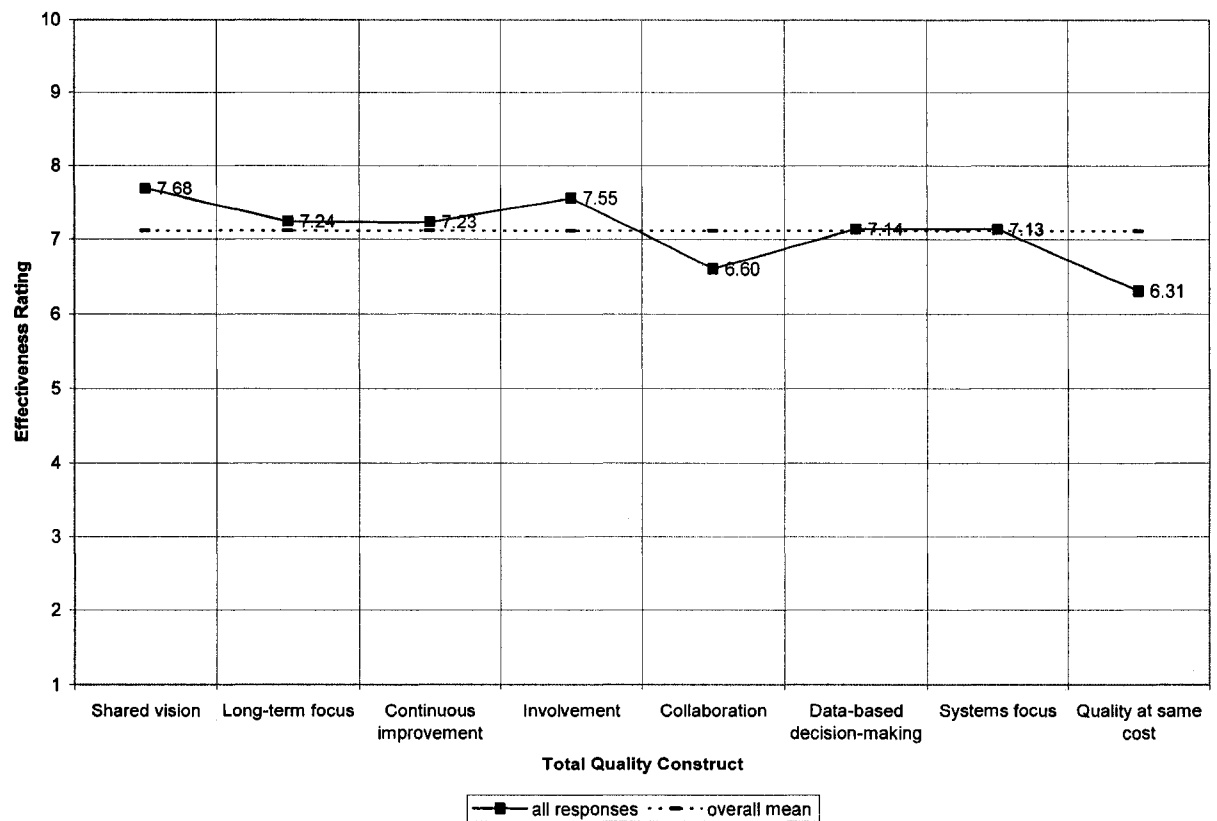
Construct	<i>t</i>	<i>Df</i>	<i>P</i>	Mean Difference
Shared vision	-16.22	185	$\leq .00$	-2.37
Long-term focus	-19.41	163	$\leq .00$	-2.89
Continuous improvement	-20.73	183	$\leq .00$	-2.86
Involvement	-18.41	184	$\leq .00$	-2.59
Collaboration	-23.46	185	$\leq .00$	-3.53
Data-based decision-making	-19.38	177	$\leq .00$	-2.88
Systems focus	-20.58	177	$\leq .00$	-2.96
Quality at the same cost	-24.04	178	$\leq .00$	-3.79

A pictorial way to examine these value gaps is through gap charts, line graph diagrams that show the size of the gaps between what “is” (values-in-use) and what

“should be” (ideal values) (Detert, Schroeder, & Cudeck, 2003). Below is the gap chart, with 10 representing the ideal, for all responses with regards to administrative perceptions of the effectiveness of CSU provosts in building a student learning assessment-supportive organizational culture:

Figure 7

Gap Chart of Provost Effectiveness—All Responses



At the CSU system level, then, provosts can be more effective by, first, improving quality at the same cost ($\bar{x}_1 = 6.31$) and collaboration ($\bar{x}_2 = 6.6$) as these means are below the overall average (7.11) and furthest from the ideal. Secondly, they ought to concentrate on a systems focus ($\bar{x}_3 = 7.13$) and data-based decision-making ($\bar{x}_4 = 7.14$), which are closest to the overall average. Third, attention should be given to continuous improvement ($\bar{x}_5 = 7.23$) and long-term focus ($\bar{x}_6 = 7.24$). Fourth, there are the areas

where provosts are the most effective, involvement ($\bar{x}_7 = 7.55$) and shared vision ($\bar{x}_8 = 7.68$).

Effectiveness by Institutional Characteristic

In order to explore differences and similarities in “value gaps” among CSU campuses, seven institutional characteristics were focused on: (1) Carnegie classification type, (2) total enrollment, (3) number of academic programs, (4) length of provost tenure, (5) number of academic affairs administrators, (6) percent of graduate students, and (7) age of institution.

One-way, between-subjects analyses of variance were conducted to find significant group differences. The one-way analysis of variance was used to test the null hypothesis that several (more than two) population means are equal; it was appropriate in this case because each of the seven groups has more than two subcategories to be compared. Mean rankings of provost effectiveness across the seven groups were compared to investigate which, if any, institutional characteristics have an effect on academic affairs administrator perceptions.

Carnegie Classification Type

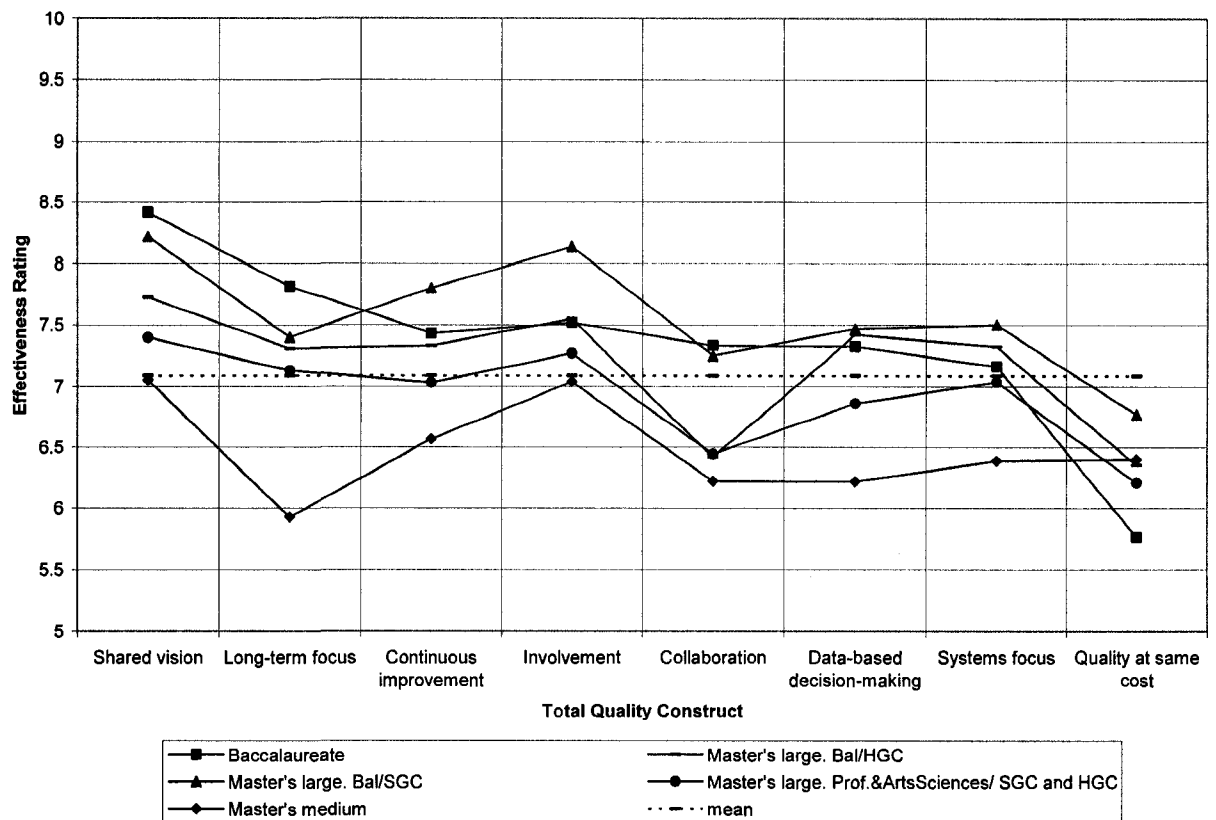
First, each of the CSU campuses with one exception was placed into one of five groups based on its Carnegie classification: (1) Baccalaureate colleges (3 campuses); (2) Master’s universities (larger programs) with undergraduate instructional programs characterized by balanced arts & sciences / professions and high graduate coexistence (7 campuses); (3) Master’s universities (larger programs) with undergraduate instructional programs characterized by balanced arts & sciences / professions and some graduate coexistence (3 campuses); (4) Master’s universities (larger programs) with undergraduate

instructional programs characterized by (a) arts & sciences plus professions and some graduate coexistence, (b) professions plus arts & sciences and some graduate coexistence, and (c) professions plus arts & sciences and high graduate coexistence (6 campuses); and (5) Master’s universities with medium-sized programs (3 campuses).²³

A one-sample t-test comparing mean scores for each of these five groups and the ideal of “10” found significant differences from the ideal ($p \leq .05$); evidence suggests there is “gap” between the means of these groups and the ideal. Below is the gap chart for these groups:

Figure 8

Gap Chart of Provost Effectiveness—Carnegie Classification Type Groups



²³ The one exception here is the CSU campus which was in its own category as a result of a graduate instructional program characterized by “doctoral, professions dominant.”

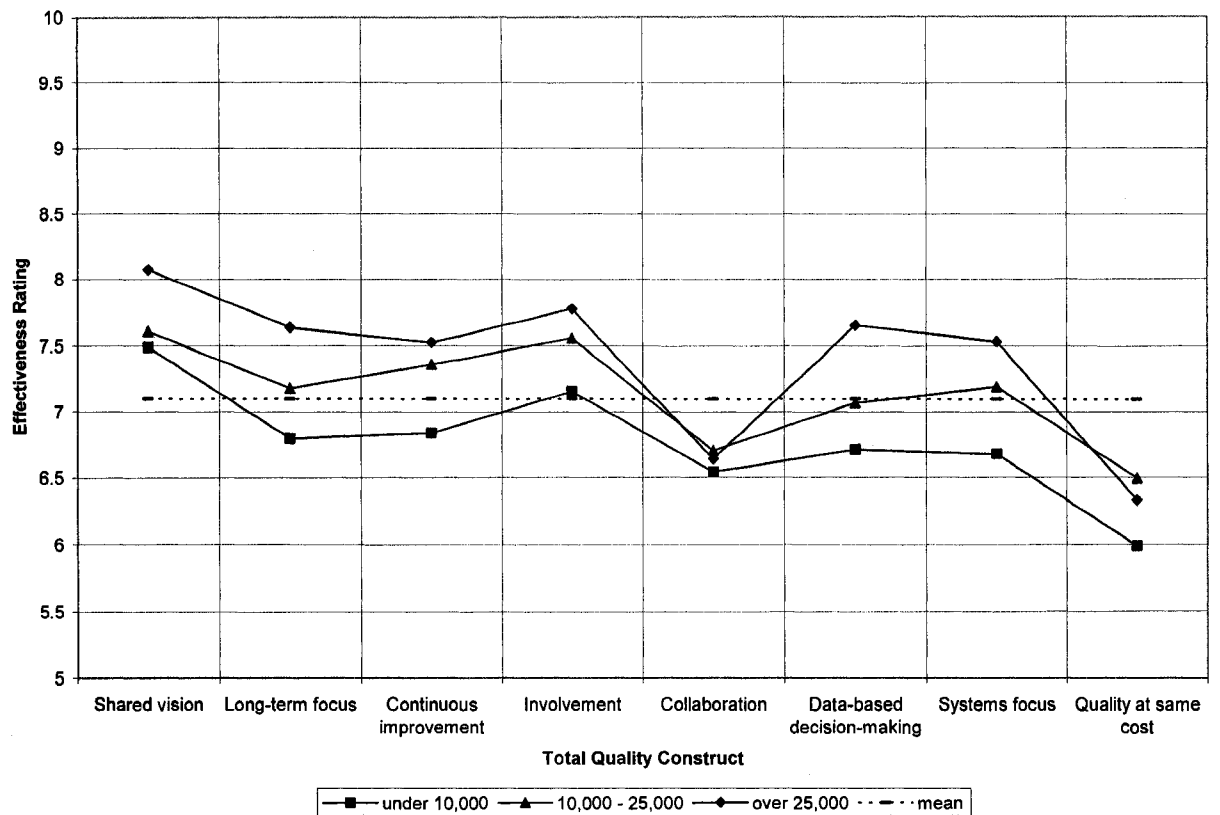
Generally, three groups can be formed here. First, there are provosts at master's medium campuses, who are furthest from the ideal with an average score ($\bar{a}_1 = 6.48$) below the overall average. Second, there campuses classified as master's universities characterized by balanced arts & sciences / professions and some graduate coexistence ($\bar{a}_2 = 6.92$), which are the second least effective. The remaining three types—master's universities characterized by balanced arts & sciences / professions and high graduate coexistence ($\bar{a}_3 = 7.18$), baccalaureate ($\bar{a}_4 = 7.35$), and master's universities characterized by arts & sciences, professions and some graduate coexistence / high graduate coexistence ($\bar{a}_5 = 7.57$)—are clustered together at the higher end of the scale above the overall mean. This third group is an interesting mix of institutions considering that the baccalaureate campuses are typically perceived as different than master's campuses, particularly large balanced ones with graduate programs.

However, a one-way analysis of variance found no significant group differences ($p = .332$) among the five groups, which implies that there is not enough evidence to show that these differences are real.

Total Enrollment

The second category used to group institutions was the total number of student enrollees. Three groups were formed: (1) campuses with less than 10,000 enrollees (8 campuses), (2) campuses with between 10,000 and 25,000 enrollees (8 campuses), and (3) campuses with more than 25,000 enrollees (7 campuses). A one-sample t-test comparing mean scores for each of these three groups and “10” found significant differences from the ideal ($p \leq .05$). Below is the gap chart for these groups:

Figure 9

Gap Chart of Provost Effectiveness—Total Enrollment Groups

The trend here is clear: campuses with a larger total enrollment were closer to the ideal, suggesting that provost effectiveness increases with enrollment. The mean for campuses with under 10,000 enrollees is 6.77, for campuses with between 10,000 and 25,000 enrollees, 7.14, and for campuses with over 25,000 enrollees, 7.4. Note that there are a couple of exceptions to this trend as found in the collaboration and quality at same cost constructs, where, as it has been pointed out, atypical distributions can be found.

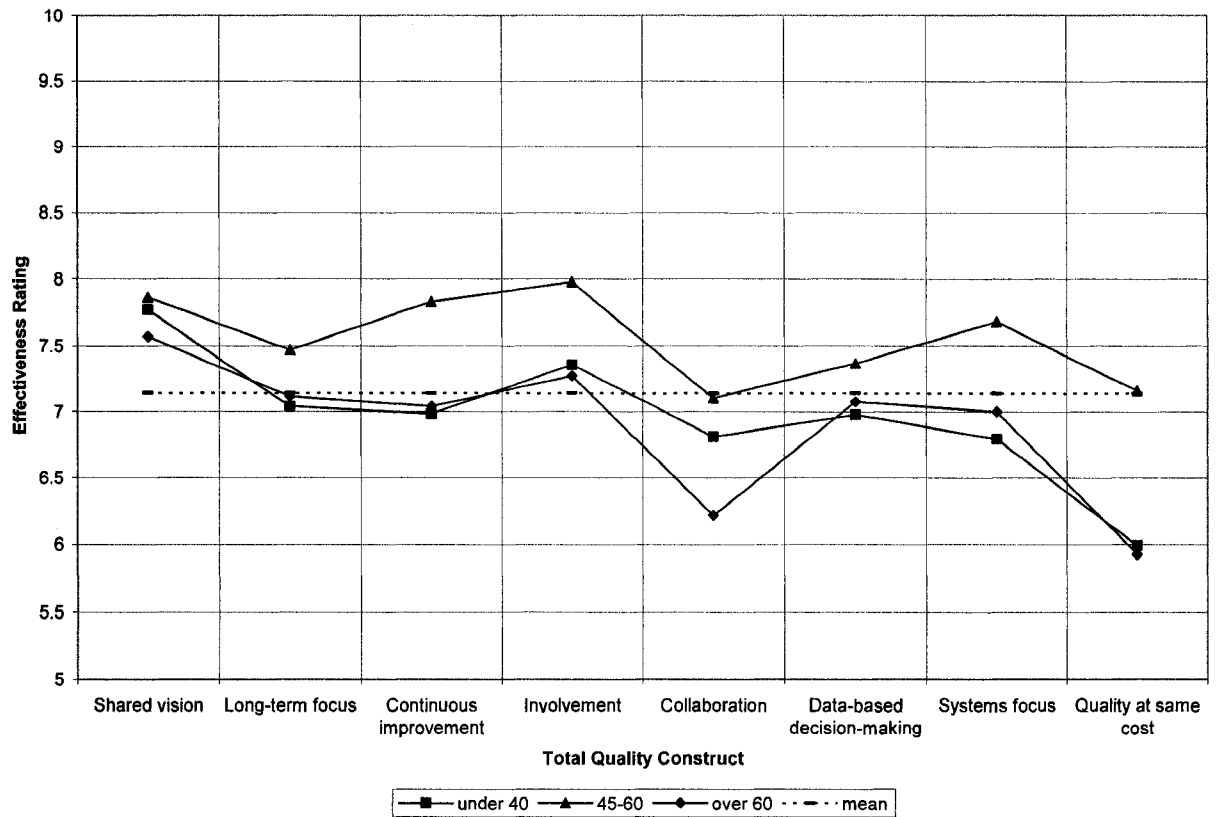
Levene's test confirmed that the homogeneity of variance assumption was met ($p = .00$). A one-way, between-subjects analysis of variance was then conducted to examine the influence of the three groups on all responses. The omnibus test revealed a significant overall effect of enrollment group, $F(2,188) = 5.11$, $p = .01$, partial $\eta^2 = .05$. Tukey's

Honestly Significant Difference tests were conducted to further examine mean differences between groups. Only one significant difference was found, that between campuses with under 10,000 enrollees and campuses with over 25,000 enrollees ($p = .01$).

Number of Academic Programs

The third institutional characteristic examined was number of academic programs (both undergraduate and graduate). There are also three groups here: (1) campuses with less than 40 academic programs (7 campuses), (2) campuses with between 40 and 60 programs (6 programs), and (3) campuses with more than 60 programs (10 campuses). A one-sample t-test comparing mean scores for each of these three groups and “10” found significant differences from the ideal ($p \leq .05$). Below is the gap chart for these groups:

Figure 10

Gap Chart of Provost Effectiveness—Number of Academic Programs Groups

Besides the score on the collaboration construct, the trends for campuses with over 60 academic programs ($\bar{c}_1 = 6.9$) and under 40 ($\bar{c}_2 = 6.97$) are nearly identical, both with means furthest from the ideal. Provosts at campuses with 45 to 60 programs are the most effective here, with a mean of 7.56. A one-way analysis of variance found, however, found no significant group differences ($p = .14$), which implies that there is not enough evidence to show that the number of academic programs makes a difference.

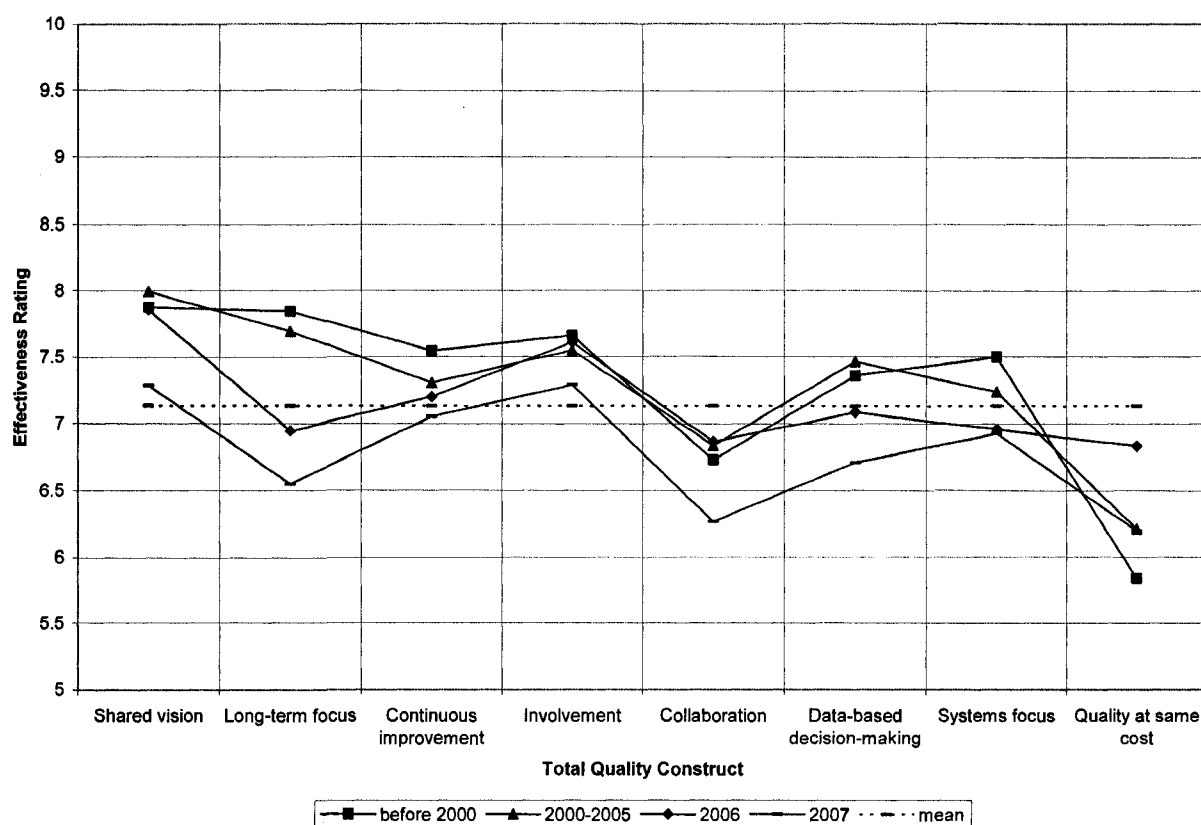
Length of Provost Tenure

The fourth category, length of provost tenure, points to an unexpected circumstance of this study—12 out of 23 provosts were appointed either in 2006 or 2007. Moreover, only 3 began before the year 2000; that is, 87% of provosts began between

2000 and 2007. On average, at the time of this study, provost tenure was about 2 and one-half years. A one-sample t-test comparing mean scores for each of these five groups and “10” found significant differences from the ideal ($p \leq .05$); evidence suggests there is “gap” between the means of these groups and the ideal. Below is the gap chart for provosts who began before 2000 (3 campuses), from 2000 to 2005 (8 campuses), in 2006 (4 campuses), and in 2007 (8 campuses):

Figure 11

Gap Chart of Provost Effectiveness—Length of Provost Tenure Groups



Generally, this chart shows the longer the provost’s tenure, the closer they are to the ideal; the longer the tenure, the more effective the provost in building an assessment-supportive culture. Provosts who began most recently (in 2007) had an average score of 6.79, provosts who began in 2006 had an average score of 7.17, and provosts who began

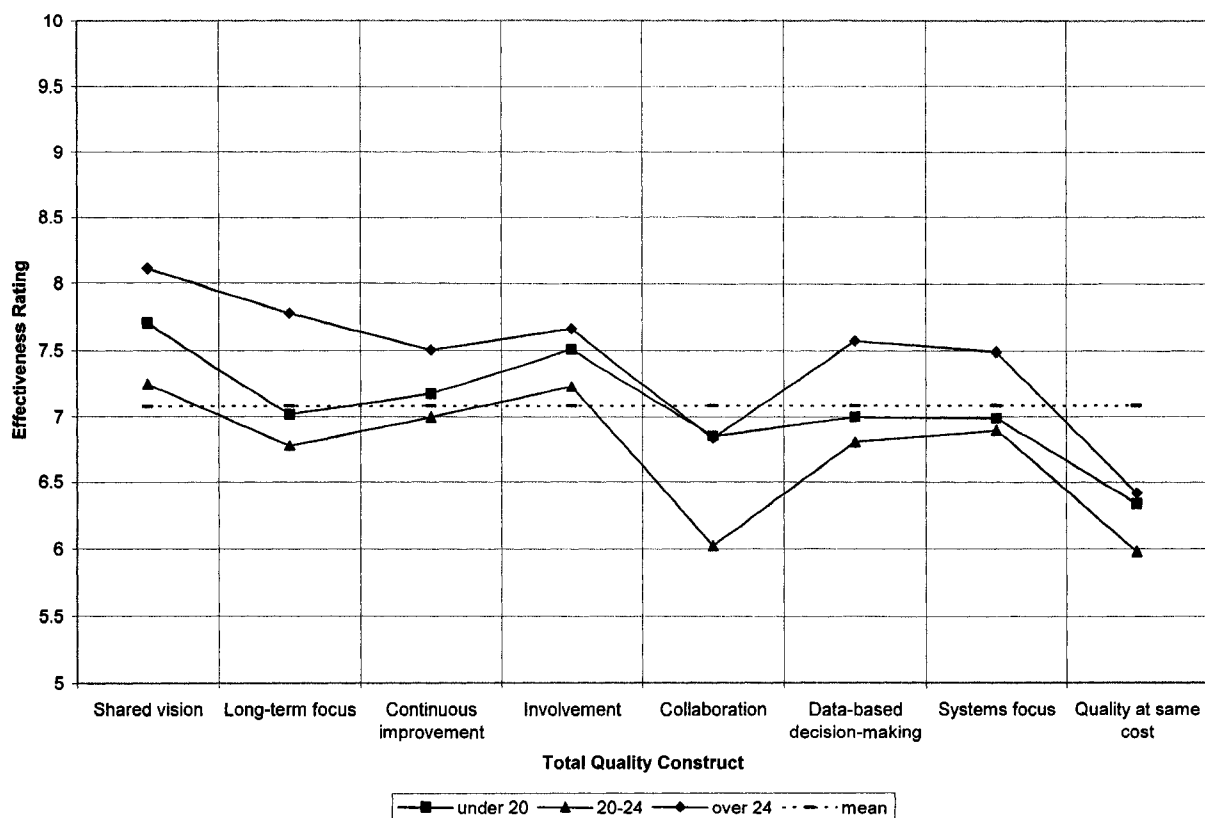
before 2000 and provosts appointed between 2000 and 2005 have an average score of 7.29. It is noteworthy to mention that a crosstabulation of CSU campuses and provost tenure found 53 (28%) respondents from campuses with provosts in their first year.

Levene's test confirmed that the homogeneity of variance assumption was met ($p = .01$). A one-way, between-subjects analysis of variance was then conducted to examine the influence of the four groups on all responses. The omnibus test revealed a significant overall effect of provost tenure group, $F(3,187) = 3.07, p = .03, \text{partial } \eta^2 = .05$. Tukey's Honestly Significant Difference tests were conducted to further examine mean differences between groups. Only one significant difference was found: between campuses with provosts who began in 2007 and campuses with provosts who began between 2000 and 2005 ($p = .02$).

Number of Academic Affairs Administrators

The fifth category is number of academic affairs administrators. The list used here is the same as the sample, thus totaling 453 administrators who were at the administrative strata of presidents, provosts, associate provosts, deans, and associate deans. Three groups were formed: (1) campuses with under 20 administrators (10 campuses), (2) between 20 and 24 administrators (6 campuses), and (3) over 24 administrators (7 campuses). A one-sample t-test comparing mean scores for each of these three groups and "10" found significant differences from the ideal ($p \leq .05$). Below is the gap chart for these groups:

Figure 12

Gap Chart of Provost Effectiveness—Number of Academic Affairs Administrators Groups

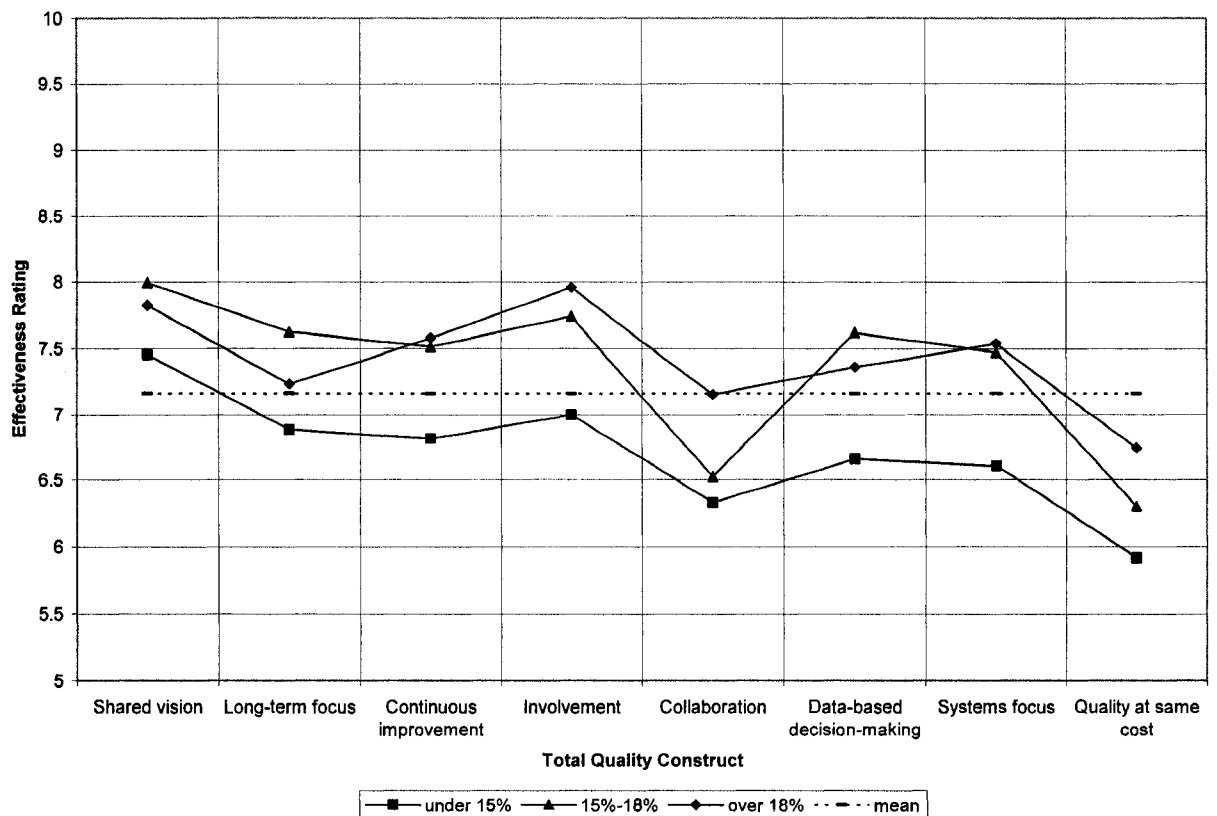
Here, medium-sized administrations are furthest from the ideal with an average score of 6.74, smaller administrations are second in being closest to the idea ($\bar{e}_2 = 7.07$) while larger administrations are closest to the ideal ($\bar{e}_3 = 7.42$). Levene's test confirmed that the homogeneity of variance assumption was met ($p = .00$). A one-way, between-subjects analysis of variance was then conducted to examine the influence of the three groups on all responses. The omnibus test revealed a significant overall effect of academic administrator group, $F(2,188) = 5.2, p = .01$, partial $\eta^2 = .05$. Tukey's Honestly Significant Difference tests were conducted to further examine mean differences between groups. Two significant differences were found: First, campuses with less than 20

administrators and with over 24 administrators ($p = .05$) and, second, campuses with between 20 to 24 administrators and campuses with over 24 administrators ($p = .01$).

Percent of Graduate Students

The sixth category was percent of graduate students. This characteristic was examined to explore the influence of graduate students who, it is assumed, are more like to have a research as opposed to a teaching orientation. Three groups were formed: campuses (1) with under 15% graduate students (10 campuses), (2) with 15% to 18% graduate students (6 campuses), and (3) with over 18% graduate students (7 campuses). A one-sample t-test comparing mean scores for each of these three groups and “10” found significant differences from the ideal ($p \leq .05$). Below is the gap chart for these groups:

Figure 13

Gap Chart of Provost Effectiveness—Percent of Graduate Students Groups

The trend here shows the lower the campuses percent of graduate students, the further from the ideal. Campuses with under 15% ($\bar{f}_1 = 6.71$) are below the average, campuses with between 15% and 18% graduate students had an average of 7.34, and campuses with over 18% graduate students had an average score of 7.42.

Levene's test found that the homogeneity of variance assumption was not met ($p = .07$). A one-way, between-subjects analysis of variance was then conducted to examine the influence of the three groups on all responses. The omnibus test revealed a significant overall effect of graduate student group, $F(2,188) = 3.78$, $p = .02$, partial $\eta^2 = .04$.

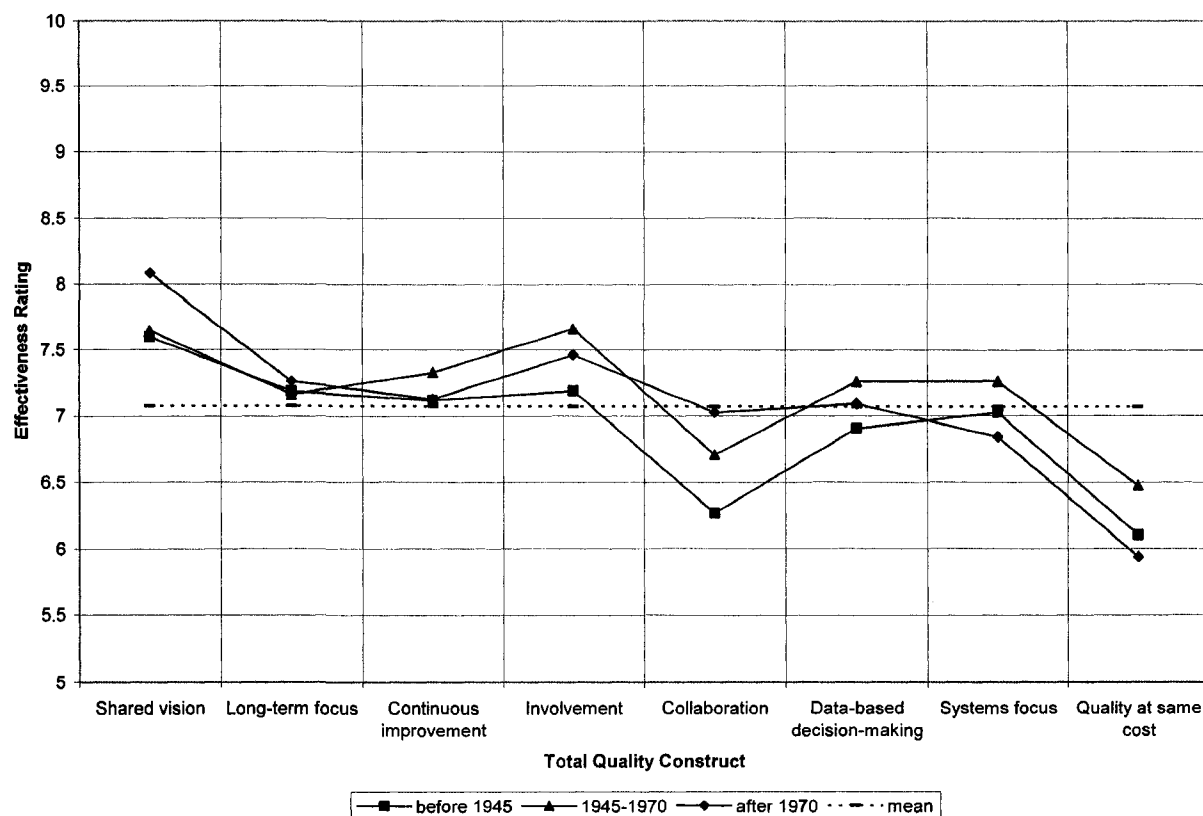
Tukey's Honestly Significant Difference tests were conducted to further examine mean

differences between groups. One significant difference was found: campuses with under 15% graduate students and campuses with between 15-18% graduate students ($p = .03$).

Age of Institution

The final category was age of institution, measured by year established. The CSU system is varied in this respect, with what began as San Jose State University established in 1862 to students beginning classes in Channel Islands, the location of the newest CSU campus, in 1999. Three groups were formed: (1) campuses established before 1945 (7 campuses), (2) campuses established between 1945 and 1970 (12 campuses), and (3) campuses established after 1970 (4 campuses). A one-sample t-test comparing mean scores for each of these three groups and “10” found significant differences from the ideal ($p \leq .05$). Below is the gap chart for these groups:

Figure 14

Gap Chart of Provost Effectiveness—Age of Institution Groups

On the whole, the trends here appear to not be different, with campuses established before 1945 having an average score of 6.92, campuses established after 1970 having an average score of 7.1., and campuses established between 1945 and 1970 having an average of 7.19. A one-way analysis of variance found no significant group differences ($p = .93$) implying that there is not enough evidence to show that campus age is important to building an assessment-supportive culture.

Effectiveness by Institutional Characteristic and Total Quality Construct

To move beyond the above discussion of significant group differences using all responses, one-way, between-subjects analyses of variance were conducted to find group differences within institutional characteristics by the Total Quality constructs. To provide

a more detailed picture of group differences within institutional characteristics, the table below shows significant ($p \leq .05$) scores by construct:

Table 7

Significance Scores by Institutional Characteristic and Total Quality Construct

Total Quality Construct	Institutional Characteristic						
	Carnegie type	Total enrollment	No. of academic programs	Length of provost tenure	No. of academic affairs administrators	Percent of graduate students	Age of institution
Shared vision		$p = .01$		$p = .03$	$p = .01$		
Long-term focus	$p = .02$	$p = .01$		$p = .00$	$p = .00$	$p = .04$	
Continuous improvement		$p = .02$	$p = .05$		$p = .03$	$p = .05$	
Involvement		$p = .03$				$p = .02$	
Collaboration				$p = .05$	$p = .04$		
Data-based decision-making		$p = .00$		$p = .01$	$p = .01$	$p = .01$	
Systems focus		$p = .01$			$p = .02$	$p = .01$	
Quality at same cost			$p = .02$	$p = .05$			

There are significant group differences among total enrollment groups (across 6 constructs), number of academic affairs administrators groups (across 6 constructs), length of provost tenure groups (5 constructs), and percent of graduate students groups (across 5 constructs). In contrast, Carnegie type (1 construct), number of academic programs (across 2 constructs), and age of institution (0 constructs) were less significant. Below are the details for the four significant categories:

Table 8

Significance Scores for Total Enrollment Groups by Total Quality Construct

Construct	Total Enrollment Group Differences
Shared vision ($p = .01$)	
Long-term focus ($p = .00$)	
Continuous improvement ($p = .01$)	Between campuses with less than 10,000 enrollees and campuses with more than 25,000 enrollees.
Involvement ($p = .03$)	
Data-based decision-making ($p = .00$)	
Systems focus ($p = .01$)	
Data-based decision-making ($p = .05$)	Between campuses with between 10,000 and 25,000 enrollees and more than 25,000 enrollees.

Table 9

*Significance Scores for Number of Academic Affairs Administrators Groups by Total Quality**Construct*

Number of Academic Affairs Administrators	
Construct	Group Differences
<hr/>	
Shared vision ($p = .03$)	
Long-term focus ($p = .00$)	Between campuses with under 20 administrators and
Data-based decision-making ($p = .03$)	campuses with over 24 administrators
Systems focus ($p = .04$)	
Shared vision ($p = .03$)	
Long-term focus ($p = .00$)	Between campuses with between 20 and 24
Collaboration ($p = .03$)	administrators and campuses over 24 administrators
Data-based decision-making ($p = .02$)	
Systems focus ($p = .05$)	
<hr/>	

Table 10

Significance Scores for Length of Provost Tenure Groups by Total Quality Construct

Construct	Length of Provost Tenure Group Differences
Shared vision ($p = .02$)	
Long-term focus ($p = .00$)	Between provosts who began from 2000 to 2005 and
Collaboration ($p = .04$)	provosts who began in 2007
Data-based decision-making ($p = .00$)	
Long-term focus ($p = .00$)	Between provosts who began before 2000 and provosts who began in 2007

Table 11

Significance Scores for Percent of Graduate Students Groups by Total Quality Construct

Construct	Percent of Graduate Students Group Differences
Long-term focus ($p = .03$)	
Continuous improvement ($p = .05$)	
Involvement ($p = .03$)	Between campuses with under 15% graduate students and campuses with 15% to 18% graduate students
Data-based decision-making ($p = .01$)	
Systems focus ($p = .02$)	
Systems focus ($p = .04$)	Between campuses with under 15% graduate students and campuses with over 18% graduate students

These results confirm analyses of variance conducted on institutional characteristic group differences:

- For total enrollment, the most common significant difference (6 constructs) was between campuses with less than 10,000 enrollees and campuses with more than 25,000 enrollees.
- For number of academic affairs administrators, the most common significant differences were between campuses with between 20 and 24 administrators and campuses with over 24 administrators (5 constructs) and between campuses with under 20 administrators and campuses with over 24 administrators (4 constructs).
- For provost tenure, the most common difference (4 constructs) was between provosts who began from 2000 to 2005 and provosts who began in 2007.
- For percent of graduate students, the most common differences (5 constructs) was between campuses with under 15% graduate students and campuses with 15% to 18% graduate students.

To further explore total enrollment, number of academic affairs administrators, length of provost tenure, and percent of graduate students, multiple linear regression was used to evaluate how closely associated these characteristics were with the Total Quality constructs, our criterion variables. Stepwise multiple regression with a $p = .05$ threshold was used and the results of these regressions are presented in Table 12.

Table 12

Regression Coefficients by Total Enrollment, Length of Provost Tenure, Number of Academic Affairs Administrators, and Percent of Graduate Students by Total Quality Construct and All Responses

Total Quality Construct	Total Enrollment	Length of Provost Tenure	No. of Academic Affairs Administrators	Percent of Graduate Student
Shared vision	3.33E-005*			
Long-term focus		.14***		
Continuous improvement	2.95E-005*			
Involvement				.05*
Collaboration				
Data-based decision-making	4.26E-005***			
Systems focus	3.54E-005**			
Quality at same cost				
All responses	3.03E-005**			

* = $p \leq .05$

** = $p \leq .01$

*** = $p \leq .00$

The obvious pattern here centers on total enrollment; total enrollment had a statistically significant relationship with four Total Quality constructs and all responses. The regression coefficient for all responses and total enrollment, for example, shows that for every additional 10,000 students, the score on all responses increased by .3. Length of provost tenure was also significant, but only in the case of long-term focus; for every additional year of provost tenure, the long-term focus score increased by .14. Percent of graduate students was significant, but only in the case of involvement; for every 10% increase of graduate students, the involvement score increased by .5. Number of academic affairs administrators was not significant.

Research Analysis Summary

First, t-tests revealed that all of the means, for all responses taken together, constructs, and for all institutional characteristic groups, were significantly different than the ideal; there were indeed gaps between ideal provost effectiveness and actual provost effectiveness, as ranked by academic affairs administrators. At the system level, approximately three scale points separated provost effectiveness from the ideal—provosts were ranked 7.11 on average, with “10” representing maximum effectiveness.

To explore the differences and similarities in these gaps among campuses, seven institutional characteristics were examined: (1) Carnegie classification type, (2) total enrollment, (3) number of academic programs, (4) length of provost tenure, (5) number of academic affairs administrators, (6) percent of graduate students, and (7) age of institution. Analyses of variance were conducted to search for significant differences within these characteristics, finding that not all groups were different from each other. Groups that were significantly different from each other are:

- Campuses with under 10,000 enrollees and campuses with over 25,000 enrollees;

- Campuses with less than 20 administrators and with over 24 administrators;
- Campuses with between 20 to 24 administrators and campuses with over 24 administrators;
- Campuses with provosts who began in 2007 and campuses with provosts who began between 2000 and 2005; and
- Campuses with under 15% graduate students and campuses with between 15-18% graduate students.

These groups point to the importance of campus size, as measured by total enrollment and number of academic affairs administrators, length of provost tenure, and percent of graduate students. To further explore these significant group differences across multiple constructs, multiple linear regression was used to evaluate how well these institutional characteristics were associated with the Total Quality constructs, finding total enrollment to account for a significant proportion of variance. This suggests that total enrollment has significant power to explain provost effectiveness in building an assessment-supportive culture compared to the other institutional characteristics.

CHAPTER 5: DISCUSSION

This chapter is divided into three sections. The first section addresses the main question, how can CSU provosts be more effective in building a student learning assessment-supportive culture at their campus?, along with the subsidiary questions, by drawing three major conclusions. The second section moves beyond the conclusions' implications for the institutional and university system levels to a broader context; it suggests the importance of this study to the collegiate student assessment movement. The final section draws from these conclusions and the an understanding of the broader context for student learning assessment to recommend future research on the study of how provosts can be more effective in building student learning assessment-supportive cultures.

Conclusions

First, there were indeed gaps between provosts' values-in-use and ideal values regarding building a student learning assessment-supportive culture. For all responses and constructs, and for all groups based on selected institutional characteristics, there were significant differences from the ideal. At the system level, however, scores were not excessively far from the ideal; the average score was 7.11 with "1" representing "minimally effective" and "10" representing "very effective." Moreover, there was not a great deal of variation; the lowest score for a construct was 6.31 and the highest for a construct, 7.68. *Overall, while provosts were not "ideally" effective, they were more effective than not in building a culture supportive of assessment.*

Second, the fact that there was some variation in scores among the eight Total Quality constructs suggested provosts should rank constructs by their mean, with lowest mean first and highest mean last, and begin becoming more effective in building an assessment-supportive culture by devoting attention to constructs with the lowest scores. An analysis of all CSU

campuses yielded this list, from lowest to highest score: quality at the same cost, collaboration, a systems focus, data-based decision-making, continuous improvement, a long-term focus, involvement, and shared vision. While the results suggest quality at the same cost should be addressed first, there are reasons why this should be reconsidered. First, in the study from which this survey was adapted, “The measurement of quality of management culture in schools: development and validation of the SQMCS” (Detert, Schroder, and Cudeck, 2002), the “quality at the same cost” construct was found to not be as reliable as the others.²⁴ Second, there were strong, negative reactions by some CSU administrators to the survey statements for this construct, expressing disagreement with the statements and pointing to an explanation of the atypical distribution of scores for this construct. *This suggests that collaboration, the next lowest-scoring construct after quality at the same cost, deserves the most attention in building an assessment-supportive culture as it may be more amenable to improvement.*²⁵

Third, there were differences in “value gaps” among CSU campus types. Results suggest that campus size, as measured by total enrollment and number of academic affairs administrators, length of provost tenure, and percent of graduate students matter in explaining group differences. The explanatory power of each of these three groups is not equal, however. Total enrollment accounted for a significant proportion of variance, suggesting that provosts at institutions with higher total enrollments are more effective at building an assessment-supportive culture and vice versa. Length of provost tenure and percent of graduate students also had an effect, albeit only on one construct each. As highlighted by many CSU administrators, however,

²⁴ In fact, having failed to write reliable items representing “quality at the same cost,” the authors suggest future work on this construct “to tap this elusive dimension of the Quality philosophy” (Detert, Schroder, and Cudeck, 2002, p.325).

²⁵ Recall that the collaboration construct was composed of statements focusing on ongoing collaborative work across administrative units, faculty frequently having conversations about student learning with colleagues from other academic units, and work time being structured to provide faculty with opportunities to work with other faculty members.

at the time of this survey about half of provosts were either in their first or second year. In fact, a handful of respondents noted that their responses reflected their perceptions of their immediately past provost as they decided their current provost was too new in the position for them to make evaluative judgments, a fact that negatively impacts the strength of these conclusions. *Taken as a whole, the comparative analysis of institutional characteristic groups suggests, first, provosts at institutions with smaller enrollments should be supported in building an assessment-supportive culture. Second, provosts with fewer years in the position should be supported in fostering a long-term focus and provosts at institutions with smaller percentages of graduate students should be supported in fostering involvement.*

Implications beyond the CSU

Beyond its significance at the institutional and system levels, this study was important in that it explored the extent to which the collegiate student assessment movement, one of the major reform movements in higher education in the last 25 years (El-Khawas, 2002), has been institutionalized.²⁶ If unsuccessful movements vanish after only a few years and if successful movements disappear as “movements” because their core values become part of the dominant culture and their practices fully institutionalized, then clearly the assessment movement is somewhere in between, in a state of limbo “both flourishing and in shambles” (Ewell, 2002; Lazerson, Wagner, & Shumanis, 2000, p.14). For this reason, Peter Ewell has described the present state of the movement as “broad but not deep” (2002, p.23). On this point, many others agree (e.g., Banta, 2002; Bok, 2006; Lazerson, Wagner, & Shumanis, 2000; Mullin, 2001; Palomba & Banta, 1999; Peterson & Vaughan, 2002; Wergin, 2005; Wright, 2002).

²⁶ As described by Levine (1980), reform movements within U.S. higher education have tended to follow a pattern: the initiation phase, when aims are set out; the implementation phase, when ideas are put into operation; and a final, institutionalization phase, when reforms try to achieve a stable, enduring form.

There is, in fact, evidence for this point. In a five-year study of student assessment at all U.S. public and private postsecondary institutions, the National Center for Postsecondary Improvement found “a pattern of partial institutionalization,” citing a process of student assessment that has yet to be fully developed and incorporated into the organizational and administrative processes of institutions (Peterson, Einarson, Augustine, & Vaughan, 1999; Peterson, Vaughan, & Perorazio, 2001). The issue of partial institutionalization is further illustrated by the fact that higher education is beyond the question of whether student assessment should exist (Erwin & Wise, 2002) and, accordingly, has developed an assessment toolbox full of approaches, methods, and instruments useful at multiple levels, but, nonetheless, cultures of assessment have yet to be institutionalized. This study, then, moves the conversation about the institutionalization of the collegiate student assessment movement forward by acknowledging the “partiality” of institutionalization and focusing on the pivotal issue of provost support in building an assessment-supportive culture, and patterns of administrative support for assessment more generally.

The role of organizational culture in the partial institutionalization of assessment led to this study’s focus on the “ideal culture profile” for Total Quality Management implementation. If examining an organizational culture supportive of Total Quality is an excellent spring board to studying a student learning assessment-supportive organizational culture, then this study also served to define the commonly-used phrase “a culture of assessment” by relying upon the principles of Total Quality. As a result, with the context of partial institutionalization in mind, it pushed forward the collegiate student assessment movement by adding dimensions of empiricism and theory to these types of popular catch phrases within the movement.

Recommendations for Further Research

Researchers further examining administrative support of assessment-supportive cultures, however, should be careful in deciding on their paradigmatic underpinnings. Several CSU respondents pointed to the assumed/imposed agreement with the functionalist assumptions of the study, questioning, for example, survey statements that assumed provosts have control over building an assessment-supportive culture, if provosts can build a value consensus, especially at large public universities with myriad of forces influencing it, and if the values of Total Quality represent the interests of those outside the administration. These critiques echo points discussed earlier in the review of the paradigmatic traditions driving the organizational culture literature. If Smircich and Calas (1987) are correct in that the field of organizational culture has become “dominant, but dead,” meaning that the initially innovative work that characterized the field has been dominated by the managerial interest and become a functionalist tool stripped of its vitality and promise, then further research should take a paradigmatic approach different than functionalism, if only to keep the field alive.

Examining perceptions among different strata of academic affairs administrators could assist this endeavor. A question that could have been asked in this study is, what are the differences and similarities in “value gaps” among different strata of administrators regarding provosts’ building a student learning assessment-supportive culture? That is, how do perceptions of provost effectiveness differ by administrative location in academic affairs? If organizational culture defined as an organization-wide consensus is a myth and culture is experienced differently at different levels or within different parts of the organization, then what is the nature of these differences? It is this empirically- and theoretically-based exploration of differentiation that could enliven the discussion of an assessment-supportive culture in higher education.

Finally, this study takes one large public university system as its subject, the California State University. Although the CSU has followed the general pattern of the initiation to codification to reaffirmation of codification of assessment policy and practice, it has been shaped by reports specific to the CSU, such as the 1989 *Student Outcomes Assessment in the California State University* and the 1998 *Cornerstones Report*; the regional accrediting agency, particularly WASC's 2001 *Handbook of Accreditation*; and California's history of organizing higher education, as exemplified by the state legislature's 1960 *A Master Plan for Higher Education in California* and the several master plans that followed (Commission for the Review of the Master Plan for Higher Education, 1987; Joint Committee to Develop a Master Plan for Education, 2002). Further research could explore the role of provosts in building an assessment-supportive culture within other large public university systems. For example, how could provosts within university systems comparable to the California State University, such as the State University of New York, University of Texas System, and the State University System of Florida, be more effective in building cultures supportive of assessment? What has been the role of assessment policy and practice within the governing bodies of these states, their regional accrediting agencies, and the university systems themselves? Taken together, these studies would begin to paint a national picture of the effectiveness of provosts in building a student learning assessment-supportive culture at their campuses.

REFERENCES

- Advisory Committee on Student Outcomes Assessment. (1989). *Student outcomes assessment in the California State University*. Long Beach, CA: California State University.
- Alkin, M. C., & Christie, C. A. (2004). An evaluation theory tree. In M.C. Alkin (Ed.), *Evaluation roots: Tracing theorists' views and influences* (pp. 12-65). Thousand Oaks: Sage.
- Alkin, M.C., Daillak, R., & White, B. (1979). *Using evaluations*. Beverly Hills, CA: Sage.
- Alstete, J. W. (1995). *Benchmarking in higher education: Adapting best practices to improve quality*. Washington, D.C.: The George Washington University Graduate School of Education and Human Development.
- Alvesson, M. (2002a). *Understanding organizational culture*. Thousand Oaks, CA: Sage.
- Alvesson, M. (2002b). *Postmodernism and social research*. Buckingham, England: Open University Press.
- American Association for Higher Education. (1992). *Principles of good practice for assessing student learning*. Washington, D.C.: American Association for Higher Education.
- American Association for Higher Education. (1993). TQM: Will it work on campus? *Change*, 25(3).
- American Association for Higher Education. (1994). *CQI 101: A first reader for higher education* (CQ9401). Washington, D.C.: American Association for Higher Education.
- Amiran, M., Schilling, K. M. & Schilling, K. L. (1993). Assessing outcomes of general education. In T. W. Banta (Ed.), *Making a difference: Outcomes of a decade of assessment in higher education* (pp. 71-86). San Francisco: Jossey-Bass.

- Andersen, C., Gardner, J. N., & Kuh, G. D. (2006, March). *Cultivating campus cultures that value student success*. Teleconference conducted by the National Resource Center for The First-Year Experience & Students in Transition.
- Anderson, L. A., & Anderson, D. (2001). *The change leader's roadmap: How to navigate your organization's transformation*. San Francisco: Pfeiffer.
- Angelo, T. A. (2002). Engaging and supporting faculty in the scholarship of assessment. In T. W. Banta (Ed.), *Building a scholarship of assessment* (pp. 185-200). San Francisco: Jossey-Bass.
- Angelo, T. A., & Cross, P. K. (1988). *Classroom assessment techniques: A handbook for college teachers*. San Francisco: Jossey-Bass.
- Angelo, T. A., & Cross, P. K. (1993). *Classroom assessment techniques: A handbook for college teachers* (2nd ed.). San Francisco: Jossey-Bass.
- Argyris, C., & Schön, D. A. (1974). *Theory in practice: Increasing professional effectiveness*. San Francisco: Jossey-Bass.
- Association of American Colleges and Universities. (2002). *Greater expectations: A new vision for learning as a nation goes to college*. Washington, D.C.: Association of American Colleges and Universities.
- Association of American Colleges and Universities. (2004). *Our student's best work: A framework for accountability worthy of our mission*. Washington, D.C.: Association of American Colleges and Universities.
- Association of American Colleges and Universities. (2005). *Liberal education outcomes: A preliminary report on student achievement in college*. Washington, D.C.: Association of American Colleges and Universities.

- Association of American Colleges and Universities. (2007). *College learning for the new global century*. Washington, D.C.: Association of American Colleges and Universities.
- Association of American Colleges. (1985). *Integrity in the college curriculum: A report to the academic community*. Washington, D.C.: Association of American Colleges.
- Astin, A. W. (1993). *Assessment for excellence: The philosophy and practice of assessment and evaluation in higher education*. Phoenix, AZ: Oryx Press.
- Bakhtin, M. (1981). *The dialogic imagination: Four essays* (C. Emerson & M. Holquist, Trans.). Austin, TX: University of Texas Press.
- Banta, T. W. (1993). *Making a difference: Outcomes of a decade of assessment in higher education*. San Francisco: Jossey-Bass.
- Banta, T. W. (1997). Moving assessment forward: Enabling conditions and stumbling blocks. *New Directions for Higher Education*, 100, 79-91.
- Banta, T. W. (2002). Characteristics of effective outcomes assessment: Foundations and examples. In T. W. Banta (Ed.), *Building a scholarship of assessment* (pp. 261-283). San Francisco: Jossey-Bass.
- Banta, T. W., Lund, J. P., Black, K. E., & Oblander, F. W. (1996). *Assessment in practice: Putting principles to work on college campuses*. San Francisco: Jossey-Bass.
- Barley, S., Meyer, G., & Gash, D. (1988). Cultures of culture: Academics, practitioners and the pragmatics of normative control. *Administrative Science Quarterly*, 33, 24-60.
- Barnard, C. (1938). *The functions of the executive*. Cambridge, MA: Harvard University Press.
- Barr, R. B., & Tagg, J. (1995). From teaching to learning—a new paradigm for undergraduate education. *Change*, 27(6), 12-25.

- Bartunek, J., & Moch, M. (1991). Multiple constituencies and the quality of working life: Intervention at FoodCom. In P. Frost, L. Moore, M. Louis, C. Lundberg, & J. Martin (Eds.), *Reframing organizational culture* (pp. 104-114). Newbury Park, CA: Sage.
- Baudrillard, J. (1988). *Selected writings* (M. Poster, Ed.). Palo Alto, CA: Stanford University Press.
- Belcher, A. R., Malmberg, M. A., & Parkman, A. W. (2002). "Learning your way": Awarding credits on the basis of competence. *Assessment Update*, 14(1), 6-8.
- Bennett, W. J. (1984). *To reclaim a legacy: A report on the humanities in higher education*. Washington, D.C.: National Endowment for the Humanities.
- Bennion, D. H., & Harris, M. (2005). Creating an assessment culture at Eastern Michigan University: A decade of progress. *Assessment Update*, 17(2), 7-9.
- Berger, P., & Luckmann, T. (1966). *The social construction of reality: A treatise in the sociology of knowledge*. Garden City, NY: Doubleday.
- Black, K. E., & Kline, K. A. (2002). Program review: A spectrum of perspectives and practices. In T. W. Banta (Ed.), *Building a scholarship of assessment* (pp. 223-239). San Francisco: Jossey-Bass.
- Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain*. New York: Longmans, Green.
- Boggs, G. R. (1999). What the learning paradigm means for faculty. *AAHE Bulletin*, 51, 3-5.
- Boisnier, A., & Chatman, J. (2003). The role of subcultures in agile organizations. In R. Peterson & E. Mannix (Eds.), *Leading and managing people in the dynamic organization* (pp. 87-117). Mahwah, NJ: Lawrence Earlbaum Associates.

- Bok, D. (2006). *Our underachieving colleges: A candid look at how much students learn and why they should be learning more*. Princeton, NJ: Princeton University Press.
- Bolman, L. G., & Deal, T. E. (2003). *Reframing organizations: Artistry, choice, and leadership* (3rd ed.). San Francisco: Jossey-Bass.
- Borden, V. M. H. (2002). Information Support for Assessment. In T. W. Banta (Ed.), *Building a scholarship of assessment* (pp. 167-181). San Francisco: Jossey-Bass.
- Boulding, K. (1956). General systems theory. *Management Science*, 2, 197-208.
- Bourdieu, P. (1977). *Outline of a theory of practice*. Cambridge: Cambridge University Press.
- Boyer Commission on Educating Undergraduates in the Research University. (1998). *Reinventing undergraduate education: A blueprint for America's research universities*. Stanford, CA: Carnegie Foundation for the Advancement of Teaching.
- Boyer, E. L. (1990). *Scholarship reconsidered: Priorities of the professoriate*. Princeton, NJ: Carnegie Foundation for the Advancement of Teaching.
- Bright, K., & Cooper, C. L. (1993). Organizational culture and the management of quality: Towards a new framework. *Journal of Managerial Psychology*, 8(6), 21-27.
- Buch, K., & Wetzell, D. K. (2001). Analyzing and realigning organizational culture. *Leadership & Organization Development Journal*, 22(1), 40-43.
- Burke, W. (2002). *Organization change: Theory and practice*. Thousand Oaks, CA: Sage.
- Burns, J. (1994). *A cultural case study of the implementation of Total Quality Management principles at a private religious university*. Unpublished doctoral dissertation, Texas Tech. University.
- Bush, T. (2003). *Theories of educational leadership and management* (3rd ed.). Thousand Oaks, CA: Sage.

- Business-Higher Education Forum. (2003). *Building a nation of learners: The need for changes in teaching and learning to meet global challenges*. Washington, D.C.: Business-Higher Education Forum.
- Business-Higher Education Forum. (2004). *Public accountability for student learning in higher education: Issues and options*. Washington, D.C.: Business-Higher Education Forum.
- California State University. (2007). *The CSU Accountability Process*. Retrieved March 1, 2007 from <http://www.calstate.edu/AcadAff/accountability/index.shtml>
- Carey, K. (2004). *A matter of degrees: Improving graduation rates in four-year colleges and universities*. Washington, D.C.: Education Trust.
- Carey, K. (2005). *One step from the finish line: Higher college graduation rates are within our reach*. Washington, D.C.: Education Trust.
- Chang, S. (1996). *Organizational culture and Total Quality Management*. Unpublished doctoral dissertation, University of Missouri-Rolla.
- Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Thousand Oaks, CA: Sage.
- Chia, R. (2003). Organization theory as a postmodern science. In H. Tsoukas & C. Knudsen (Eds.), *The Oxford handbook of organization theory* (pp. 113-139). Oxford: Oxford University Press.
- Chickering, A. W., & Gamson, Z. F. (1987). Seven principles for good practice in undergraduate education. *AAHE Bulletin*, 37(7), 3-7.
- Chun, M. (2002). Looking where the light is better: A review of the literature on assessing higher education quality. *AAC&U Peer Review*, Winter/Spring, 16-25.

- Cohen, M., March, J., & Olsen, J. (1972). A garbage can model of organizational choice. *Administrative Science Quarterly*, 17, 1-25.
- Collins, J. (2001). *Good to great: Why some companies make the leap...and others don't*. New York: HarperBusiness.
- Collins, J., & Porras, J. (1994). *Built to last: Successful habits of visionary companies*. New York: HarperBusiness.
- Commission for the Review of the Master Plan for Higher Education. (1987). *The master plan renewed: Unity, equity, quality, and efficiency in California postsecondary education*. Sacramento, CA: California Legislature.
- Connell, A., & Nord, W. (1996). The bloodless coup: The infiltration of organization science by uncertainty and values. *Journal of Applied Behavioral Science*, 32, 407-427.
- Council for Higher Education Accreditation. (2003). *Statement of mutual responsibilities for student learning outcomes: Accreditation, institutions, and programs*. Washington, D.C.: Council for Higher Education Accreditation.
- Cousins, J. B. (2004). Crossing the bridge: Toward understanding use through systematic inquiry. In Alkin, M. C. (Ed.), *Evaluation roots: Tracing theorists' views and influences* (pp. 319-330). Thousand Oaks, CA: Sage.
- Cousins, J. B., & Earl, L. M. (1995). *Participatory evaluation in education: Studies in evaluation use and organizational learning*. London: Falmer.
- Cousins, J. B., & Lee, L. E. (2004). Integrating evaluative inquiry into the organizational culture: A review and synthesis of the knowledge base. *The Canadian Journal of Program Evaluation*, 19(2), 99-141.

- Cousins, J. B., & Whitmore, E. (1998). Framing participatory evaluation. *New Directions for Evaluation*, 80, 5-23.
- Cousins, J. B., Goh, S. C., Clark, S., & Lee, L. E. (2004). Integrating evaluative inquiry into the organizational culture: A review and synthesis of the knowledge base. *The Canadian Journal of Program Evaluation*, 19(2), 99-141.
- Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches* (2nd ed.). Thousand Oaks, CA: Sage.
- Cross, P. K., & Steadman, M. H. (1996). *Classroom research: Implementing the scholarship of teaching*. San Francisco: Jossey-Bass.
- Cummings, T., & Worley, C. (2004). *Organization development and change* (8th ed.). Belmont, CA: South-Western.
- Darwin, C. (1964). *On the origin of species*. Cambridge, MA: Harvard University Press.
- Datta, L. (1997). A pragmatic basis for mixed-method designs. *New Directions for Evaluation*, 74, 33-46.
- Deal, T., & Kennedy, A. (1982). *Corporate cultures: The rites and rituals of corporate life*. Reading, MA: Addison-Wesley.
- Denison, D. (1990). *Corporate culture and organizational effectiveness*. New York: John Wiley & Sons.
- Denzin, N. K., & Y. S. Lincoln (Eds.), (1998). *The landscape of qualitative research: Theories and issues*. Thousand Oaks, CA: Sage.
- Derrida, J. (1978). *Writing and difference* (A. Bass, Trans.). Chicago: University of Chicago Press.

- Derrida, J. (1980). *Of grammatology* (G. C. Spivak, Trans.). Baltimore, MD: Johns Hopkins University Press.
- Detert, J. R., & Mauriel, J. J. (1997). *Using the lessons of organizational change and previous school reforms to predict innovation outcomes: Should we expect more from TQM?* Paper presented at the Annual Meeting of the American Educational Research Association.
- Detert, J. R., Bauerly Kopel, M. E., Mauriel, J. J., & Jenni, R. W. (2000). Quality management in U.S. high schools: Evidence from the field. *Journal of School Leadership*, 10, 158-187.
- Detert, J. R., Schroeder, R., & Cudeck, R. (2003). The measurement of quality management culture in schools: Development and validation of the SQMCS. *Journal of Operations Management*, 21, 307-328.
- Detert, J. R., Schroeder, R., & Mauriel, J. (2000). A framework for linking culture and improvement initiatives in organizations. *Academy of Management Review*, 25 (4), 850-863.
- Detert, J. R., Seashore Louis, K., & Schroeder, R. (2001). A cultural framework for education: Defining quality values and their impact in U.S. high schools. *School Effectiveness and School Improvement*, 12 (2), 183-212.
- Donaldson, L. (2003). Organization theory as a positive science. In H. Tsoukas & C. Knudsen (Eds.), *The Oxford handbook of organization theory* (pp. 88-112). Oxford: Oxford University Press.
- Donaldson, S. I., & Gooler, L. E. (2003). Theory-driven evaluation in action: Lessons from a \$20 million statewide Work and Health Initiative. *Evaluation and Program Planning*, 26, 355-366.
- Donaldson, S. I., & Scriven, M. (2003). *Evaluating social programs and problems: Visions for the new millennium*. Mahwah, NJ: Lawrence Erlbaum Associates.

- Dowd, A. C. (2005). *Data don't drive: Building a practitioner-driven culture of inquiry to assess community college performance*. Indianapolis, IN: Lumina Foundation for Education.
- Durkheim, E. (1984). *The division of labor in society* (W. D. Halls, Trans.). New York: Free Press.
- Education Commission of the States. (1986). *Transforming the state role in undergraduate education: Time for a different view*. Denver, CO: Education Commission of the States.
- Eisenberg, E. M., Murphy, A., & Andrews, L. (1998). Openness and decision making in the search for a university provost. *Communication Monographs*, 65, 1-23.
- El-Khawas, E. (1993). 1993 campus trends survey. Washington, D.C.: American Council on Education.
- El-Khawas, E. (2002). Reform initiatives in higher education. In *ERIC Digest*. Washington, D.C.: ERIC Clearinghouse on Higher Education.
- Entin, D. H. (1993). Less than meets the eye. *Change*, 25(3), 19-21.
- Erwin, T. D., & Wise, S. L. (2002). A scholar-practitioner model for assessment. In T. W. Banta (Ed.), *Building a scholarship of assessment* (pp. 67-81). San Francisco: Jossey-Bass.
- Evans, J. R., & Dean, J. W. (2003). *Total Quality: Management, organization, and strategy* (3rd ed.). Mason, OH: South-Western.
- Ewell, P. T. (1989). Institutional characteristics and faculty/administrator perceptions of outcomes: An exploratory analysis. *Research in Higher Education*, 30(2), 113-136.
- Ewell, P. T. (1993a). The role of states and accreditors in shaping assessment practice. In T. W. Banta (Ed.), *Making a difference: Outcomes of a decade of assessment in higher education* (pp. 339-356). San Francisco: Jossey-Bass.

- Ewell, P. T. (1993b). Total Quality & academic practice: The idea we've been waiting for? *Change*, 25(3), 37-43.
- Ewell, P. T. (1999). Imitation as art: Borrowed management techniques in higher education. *Change*, 31(6), 11-15.
- Ewell, P. T. (2002). An emerging scholarship: A brief history of assessment. In T. W. Banta (Ed.), *Building a scholarship of assessment* (pp. 3-25). San Francisco: Jossey-Bass.
- Ewell, P. T. (2004). *General education and the assessment reform agenda*. Washington, D.C.: Association of American Colleges and Universities.
- Feldman, M. (1991). The meanings of ambiguity: Learning from stories and metaphors. In P. Frost, L. Moore, M. Louis, C. Lundberg, & J. Martin (Eds.), *Reframing organizational culture* (pp. 145-156). Newbury Park, CA: Sage.
- Fetterman, D. M. (2003). Empowerment evaluation strikes a responsive cord. In S. I. Donaldson & M. Scriven (Eds.), *Evaluating social programs and problems: Visions for the new millennium* (pp. 63-76). Mahwah, NJ: Lawrence Erlbaum.
- Fetterman, D. M. (2004). Branching out or standing on a limb: Looking to our roots for insight. In M. C. Alkin (Ed.), *Evaluation roots: Tracing theorists' views and influences* (pp. 304-318). Thousand Oaks, CA: Sage.
- Fetterman, D. M., Kaftarian, S. J., & Wandersman, A. (1996). *Empowerment evaluation: Knowledge and tools for self-assessment and accountability*. Thousand Oaks, CA: Sage.
- Fitzpatrick, J. (2002). Dialogue with Stewart Donaldson. *American Journal of Evaluation*, 23(3), 347-365.
- Foucault, M. (1972). *The archeology of knowledge and the discourse on language*. New York: Pantheon Books.

- Foucault, M. (1973). *The order of things: An archaeology of human sciences*. New York: Vintage Books.
- French, J. R. P., & Raven, B. H. (1960). The bases of social power. In D. Cartwright and A. F. Zander (Eds.), *Group dynamics: Research and theory* (2nd ed.) (pp. 607-623). Evanston, IL: Row.
- Frost, P., Moore, L., Louis, M., Lundberg, C., & Martin, J. (Eds.) (1991). *Reframing organizational culture*. Newbury Park, CA: Sage.
- Fultz, M. L., & Wong, L. E. (2001). Creating and sustaining an assessment culture. *Assessment Update*, 13(2), 12-13.
- Gaither, G. H. (1998). Quality assurance in higher education: An international perspective. *New Directions for Institutional Research*, 99.
- Garvin, D. A. (1993). Building a Learning Organization. *Harvard Business Review*, 71(4) 78-91.
- Geertz, C. (1973). *The interpretation of cultures*. New York: Basic Books.
- Gerth, H. H., & Mills, C. W. (1946). *From Max Weber: Essays in sociology*. Oxford: Oxford University Press.
- Gordon, G. (1999). Creating effective internal processes for quality assessment, management, and development. *Assessment Update*, 11(6), 6-8.
- Grackin, J. A. (2005). Building a campuswide culture of assessment: Including nonacademic departments in the process. *Assessment Update*, 17(5), 1-13.
- Gray, P. J. (1997). Viewing assessment as an innovation: Leadership and the change process. *New Directions for Higher Education*, 100, 5-15.

- Gray, P. J. (2002). The roots of assessment: Tensions, solutions, and research directions. In T. W. Banta (Ed.), *Building a scholarship of assessment* (pp. 49-66). San Francisco: Jossey-Bass.
- Gray, P. J., & Banta, T. W. (Eds.). (1997). The campus-level impact of assessment: Progress, problems, and possibilities. *New Directions for Higher Education*, 100.
- Greene, J. C., & Caracelli, V. J. (1997). Defining and describing the paradigm issue in mixed-method evaluation. *New Directions for Evaluation*, 74, 5-17.
- Gregory, K. (1983). Native-view paradigms: Multiple cultures and culture conflicts in organizations. *Administrative Science Quarterly*, 28, 359-376.
- Guskin, A. E. (1996). Facing the future: The change process in restructuring universities. *Change*, 28(4), 27-37.
- Habermas, J. (1972). *Knowledge and human interests* (J. Shapiro, Trans.). Boston: Beacon.
- Harrison, M. I. (2005). *Diagnosing organizations: Methods, models, and processes* (3rd ed.). Thousand Oaks, CA: Sage.
- Hart, C. (1998). *Doing a literature review: Releasing the social science research imagination*. Thousand Oaks, CA: Sage.
- Hatch, M. J. (1993). The dynamics of organizational culture. *Academy of Management Review*, 18(4), 657-693.
- Hatch, M. J. (1995). [Review of the books *Organizational communication and cultural vision: Approaches for analysis*; *Organizational cultures: Types and transformations*; *Occupational subcultures in the workplace*]. *Contemporary Sociology*, 24, 249-252.
- Hatch, M. J. (1997). *Organization theory: Modern, symbolic, and postmodern perspectives*. Oxford: Oxford University Press.

- Hatch, M. J. (1999). Exploring the empty spaces of organizing: How improvisational jazz helps redescribe organizational structure. *Organization Studies*, 20, 75-99.
- Hatch, M. J., & Yanow, D. (2003). Organization theory as an interpretive science. In H. Tsoukas & C. Knudsen (Eds.), *The Oxford handbook of organization theory* (pp. 63-87). Oxford: Oxford University Press.
- Henry, G. T. (2000). Why not use? *New Directions for Evaluation*, 88, 85-98.
- Henry, G. T., & Mark, M. M. (2003). Beyond use: Understanding evaluation's influence on attitudes and actions. *American Journal of Evaluation*, 24 (3), 293-314.
- Herrell, J. M., & Straw, R. B. (Eds.). (2002). Conducting multiple site evaluations in real-world settings. *New Directions for Evaluation*, 94.
- Hertz, R., & Imber, J. B. (Eds.). (1995). *Studying elites using qualitative methods*. Thousand Oaks, CA: Sage.
- House, E. R. (1995). Principled evaluation: A critique of the AEA Guiding Principles. *New Directions for Program Evaluation*, 66, 27-34.
- Hursh, D. (2006). The crisis in urban education: Resisting neoliberal policies and forging democratic possibilities. *Educational Researcher*, 35(4), 19-25.
- Hurtado, S., Milem, J. F., Clayton-Pedersen, A., & Allen, W.A. (1999). *Enacting diverse learning environments: Improving the climate for racial/ethnic diversity in higher education*. Washington, D.C.: George Washington University Graduate School of Education and Human Development.
- Ingram, D. Seashore Louis, K., & Schroeder, R. G. (2004). Accountability policies and teacher decision making: Barriers to the use of data to improve practice. *Teachers College Record*, 106(6), 1258-1287.

- Jencks, C. (1977). *The language of post-modern architecture*. London: Academy.
- Jermier, J. (1991). Critical epistemology and the study of organizational culture: Reflections on *Street Corner Society*. In P. Frost, L. Moore, M. Louis, C. Lundberg, & J. Martin (Eds.), *Reframing organizational culture* (pp. 223-233). Newbury Park, CA: Sage.
- Jermier, J., Slocum, J., Jr., Fry, L., & Gaines, J. (1991). Organizational subcultures in a soft bureaucracy: Resistance behind the myth and façade of an official culture. *Organization Science*, 2, 170-194.
- Johnson, R., McCormick, R. D., Prus, J. S., & Rogers, J. S. (1993). Assessment options for the college major. In T. W. Banta (Ed.), *Making a difference: Outcomes of a decade of assessment in higher education* (pp. 151-167). San Francisco: Jossey-Bass.
- Johnson, R., Prus, J., Andersen, C. J., & El-Khawas, E. (1994). Assessing assessment: An in-depth status report on the higher education assessment movement in 1990. In J. S. Stark & A. Thomas (Eds.), *Assessment & Program Evaluation* (pp. 87-98). Boston: Simon & Schuster.
- Joint Committee to Develop a Master Plan for Education. (2002). *The California master plan for education*. Sacramento, CA: California Legislature.
- Kamuf, P. (Ed.). (1991). *A Derrida reader: Between the blinds*. New York: Columbia University Press.
- Kells, H. R. (1995). Creating a culture of evaluation and self-regulation in higher education organizations. *Total Quality Management*, 4(5 & 6), 457-467.
- Kemmerer, F. (1994). *Utilizing education and human resource sector analyses*. Paris: Imprimerie Gauthier-Villars.
- Kerlinger, F. N. (1986). *Foundations of behavioral research* (3rd ed.). Orlando, FL: Harcourt.

- Kerr, B. A. (2002). New kid on the block: Becoming a culture of assessment—a work in progress. *Assessment Update*, 14(4), 12-14.
- Kezar, A., & Eckel, P. D. (2002). The effect of institutional culture on change strategies in higher education. *The Journal of Higher Education*, 73(4), 435-460.
- Kilmann, R. (1985). Five steps for closing culture-gaps. In R. Kilmann, M. Saxton, R. Serpa, & Associates (Eds.), *Gaining control of the corporate culture* (pp. 351-369). San Francisco: Jossey-Bass.
- Kilmann, R., Saxton, M., Serpa, R., & Associates (1985). *Gaining control of the corporate culture*. San Francisco: Jossey-Bass.
- Kincheloe, J., & McLaren, P. (1998). Rethinking critical theory and qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *The landscape of qualitative research: Theories and issues* (pp. 260-299). Thousand Oaks, CA: Sage.
- Knight, W. E. (2003). *The primer for institutional research*. Tallahassee, FL: Association for Institutional Research.
- Konidari, V., & Abernot, Y. (2006). From TQM to learning organization: Another way for quality management in educational institutions. *International Journal of Quality & Reliability Management*, 23(1), 8-26.
- Kotter, J., & Heskett, J. (1992). *Corporate culture and performance*. New York: Free Press.
- Krieger, J. (1986). *Reagan, Thatcher and the politics of decline*. Oxford: Oxford University Press.
- Krueger, D. W. (1993). Total Quality Management. In T. W. Banta (Ed.), *Making a difference: Outcomes of a decade of assessment in higher education* (pp. 269-278). San Francisco: Jossey-Bass.

- Kuh, G. D., Kinzie, J., Schuh, J. H., & Whitt, E. J. (2005a). *Student success in college: Creating conditions that matter*. San Francisco: Jossey-Bass.
- Kuh, G. D., Kinzie, J., Schuh, J. H., & Whitt, E. J. (2005b). *Assessing conditions to enhance educational effectiveness: The inventory for student engagement and success*. San Francisco: Jossey-Bass.
- Lazerson, M., Wagener, U., & Shumanis, N. (2000). What makes a revolution: Teaching and learning in higher education, 1980-2000. *Change*, 32(3), 12-19.
- Levine, A. (1980). *Why innovation fails: The institutionalization and termination of innovation in higher education*. Albany, NY: SUNY Press.
- Light, R. J. (1990). *The Harvard Assessment Seminars: First report. Explorations with students and faculty about teaching, learning, and student life*. Cambridge, MA: Harvard University Press.
- Light, R. J. (1992). *The Harvard Assessment Seminars: Second report. Explorations with students and faculty about teaching, learning, and student life*. Cambridge, MA: Harvard University Press.
- Light, R. J. (2001). *Making the most of college: Students speak their mind*. Cambridge, MA: Harvard University Press.
- Love, A. J. (1991). *Internal evaluation: Building organizations from within*. Newbury, CA: Sage.
- Lyotard, J. (1984). *The postmodern condition: A report on knowledge*. Minneapolis, MN: University of Minnesota Press.

- Magruder, J., McManis, M. A. & Young, C. C. (1997). The right idea at the right time: Development of a transformational assessment culture. *New Directions for Higher Education*, 100, 17-29.
- Malinowski, B. (1935). *Coral gardens and their magic: A study of the methods of tiling the soil and of agricultural rites in the Trobriand Islands*. London: G. Allen & Unwin.
- March, J., & Olsen, J. (Eds.) (1976). *Ambiguity and choice in organizations*. Bergen, Norway: Universitetsforlaget.
- Marchese, T. J. (1994). Assessment, quality, and undergraduate improvement. *Assessment Update*, 6(3), 1-2, 12-14.
- Mark, M. M. (2003). Program evaluation. In Schinka, J. A. & Velicer, W. F. (Eds.), *Handbook of psychology: Volume 2: Research methods in psychology* (pp. 323-347). John Wiley & Sons.
- Mark, M. M., & Henry, G. T. (2004). The mechanisms and outcomes of evaluation influence. *Evaluation*, 10(1), 35-57.
- Martin, J. (1992). *Cultures in organizations: Three perspectives*. New York: Oxford University Press.
- Martin, J. (2002). *Organizational culture: Mapping the terrain*. Thousand Oaks, CA: Sage.
- Martin, J., & Meyerson, D. (1988). Organizational culture and the denial, channeling and acknowledgment of ambiguity. In L. Pondy, R. Boland, Jr., & H. Thomas (Eds.), *Managing ambiguity and change* (pp. 93-125). New York: John Wiley.
- Martin, J., & Siehl, C. (1983). Organizational culture and counterculture: An uneasy symbiosis. *Organizational Dynamics*, 52-64.

- Martin, J., Frost, P. J., & O'Neill, O. A. (2004). *Organizational culture: Beyond struggles for intellectual dominance*. Retrieved October 17, 2006, from Stanford University, Stanford Graduate School of Business Web site: <https://gsbapps.stanford.edu/researchpapers/library/RP1864.pdf>
- Maslowski, R. (2006). A review of inventories for diagnosing school culture. *Journal of Educational Administration*, 44(1), 6-35.
- Massey, W. F. (2003). *Honoring the trust: Quality and cost containment in higher education*. Boston: Anker.
- Master Plan Survey Team. (1960). *A master plan for higher education in California, 1960-1975*. Sacramento, CA: California State Department of Education.
- Merton, R. K. (1957). *Social theory and social structure* (2nd ed.). Glencoe, IL: Free Press.
- Meyerson, D. (1991). "Normal" ambiguity? A glimpse of an occupational culture. In P. Frost, L. Moore, M. Louis, C. Lundberg, & J. Martin (Eds.), *Reframing organizational culture* (pp. 131-144). Newbury Park, CA: Sage.
- Meyerson, D., & Martin, J. (1987). Cultural change: An integration of three different views. *Journal of Management Studies*, 24, 623-647.
- Miller, R., & Leskes, A. (2005). *Levels of assessment: From the student to the institution*. Washington, D.C.: Association of American Colleges and Universities.
- Morgan, G. (2006). *Images of organization*. Thousand Oaks, CA: Sage.
- Moyser, G., & Wagstaffe, M. (Eds.). (1987). *Research methods for elite studies*. London: Allen & Unwin.
- Mullin, R. (2001). The undergraduate revolution: Change the system or give incrementalism another 30 years? *Change*, 33(5), 54-58.

- Murphy, J. P. (1990). *Pragmatism: from Peirce to Davidson*. Boulder, CO: Westview Press.
- Nader, L. (1972). Up the anthropologist—perspectives gained from studying up. In D. H. Hymes (Ed.), *Reinventing anthropology* (pp. 284-311). New York: Pantheon.
- National Association of State Universities and Land-Grant Colleges. (2006). *Improving student learning in higher education through better accountability and assessment*. Washington, D.C.: National Association of State Universities and Land-Grant Colleges.
- National Center for Public Policy and Higher Education. (2005a). *Measuring up on college-level learning*. Washington, D.C.: National Center for Public Policy and Higher Education.
- National Center for Public Policy and Higher Education. (2005b). *Income of U.S. workers projected to decline in education doesn't improve*. National Center for Public Policy and Higher Education.
- National Center for Public Policy and Higher Education. (2005c). *State capacity for higher education policy*. National Center for Public Policy and Higher Education.
- National Governors' Association. (1986). *Time for results: The governors' report on education*. Washington, D.C.: National Governors' Association.
- Nelson, R. (1991). *Planning for assessment at UC Santa Cruz*. Retrieved October 16, 2006, from UC Santa Cruz, Office of Planning and Budget Web site: <http://planning.ucsc.edu/irps/office/nelson/scassess.htm>
- Norris, J. M. (2006). The why (and how) of student learning outcomes assessment in college FL education. *Modern Language Journal*, 90(4).
- O'Banion, T. (1997). *A learning college for the 21st century*. Phoenix, AZ: Oryx Press.
- O'Banion, T. (1999). *Launching a learning-centered college*. Mission Viejo, CA: League for Innovation in the Community College.

- Obler, S. S., Slark, J., & Umbdenstock, L. (1993). Classroom assessment. In T. W. Banta (Ed.), *Making a difference: Outcomes of a decade of assessment in higher education* (pp. 211-226). San Francisco: Jossey-Bass.
- Ouchi, W. (1981). *Theory Z: How American business can meet the Japanese challenge*. Boston: Addison-Wesley.
- Ouchi, W., & Wilkins, A. (1985). Organizational culture. *Annual Review of Sociology*, 11, 457-483.
- Owen, J. M. (2004). Evaluation forms: Toward an inclusive framework for evaluation practice. In M.C. Alkin (Ed.), *Evaluation roots: Tracing theorists' views and influences* (pp. 356-369). Thousand Oaks, CA: Sage.
- Owen, J. M., & Lambert, F. C. (1995). Roles for evaluation in learning organizations. *Evaluation*, 1(2), 237-250.
- Owen, J. M., & Lambert, F. C. (1998). Evaluation and the information needs of organizational leaders. *American Journal of Evaluation*, 19(3), 355-365.
- Owen, J. M., & Rogers, P. J. (1999). *Program evaluation: Forms and approaches*. London: Sage.
- Palomba, C. A. & Banta, T. W. (1999). *Assessment essentials: Planning, implementing, and improving assessment in higher education*. San Francisco: Jossey-Bass.
- Palomba, C. A. (1997). Assessment at Ball State University. *New Directions for Higher Education*, 100, 31-45.
- Palomba, C. A. (2002). Scholarly assessment of student learning in the major and general education. In T. W. Banta (Ed.), *Building a scholarship of assessment* (pp. 201-222). San Francisco: Jossey-Bass.

- Palomba, C. A., & Banta, T. W. (1999). *Assessment essentials: Planning, implementing, and improving assessment in higher education*. San Francisco: Jossey-Bass.
- Parsons, T. (1951). *The social system*. Glencoe, IL: Free Press.
- Pascale, R. T., & Sternin, J. (2005). Your company's secret change agents. *Harvard Business Review*, 83(5), 72-81.
- Pascale, R., & Athos, A. (1981). *The art of Japanese management: Applications for American executives*. New York: Simon and Schuster.
- Patton, M. Q. (1978). *Utilization-focused evaluation*. Beverly Hills, CA: Sage.
- Patton, M. Q. (1986). *Utilization-focused evaluation* (2nd ed.). Beverly Hills, CA: Sage.
- Patton, M. Q. (1997). *Utilization-focused evaluation: The new century text* (3rd ed.). Thousand Oaks, CA: Sage.
- Patton, M. Q. (2002). *Qualitative research & evaluation methods* (3rd ed.) Thousand Oaks, CA: Sage.
- Patton, M. Q. (2004). The roots of utilization-focused evaluation. In Alkin, M. C. (Ed.), *Evaluation roots: Tracing theorists' views and influences* (pp. 276-292). Thousand Oaks, CA: Sage.
- Patton, M. Q., Grimes, P. S., Guthrie, K. M., Brennan, N. J., French, B. D., & Blyth, D. A. (1977). In search of impact: An analysis of the utilization of federal health evaluation research. In C. H. Weiss (Ed.), *Using social research in public policy making* (pp. 141-163). Lexington, MA: Lexington Books.
- Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. Thousand Oaks, CA: Sage.
- Peters, T., & Waterman, R., Jr. (1982). *In search of excellence: Lessons from America's best-run companies*. New York: Harper & Row.

- Peterson, M. W., & Augustine, C. H. (2000). Organizational practices enhancing the influence of student assessment information in academic decisions. *Research in Higher Education*, 41(1), 21-52.
- Peterson, M. W., & Einarson, M. K. (1997). *Analytic framework of institutional support for student assessment*. Stanford, CA: Stanford University, National Center for Postsecondary Improvement.
- Peterson, M. W., & Vaughan, D. S. (2002). Promoting academic improvement: Organizational and administrative dynamics that support student assessment. In T. W. Banta (Ed.), *Building a scholarship of assessment* (pp. 26-46). San Francisco: Jossey-Bass.
- Peterson, M. W., Einarson, M. K., Augustine, C. H., & Vaughan, D. S. (1999). *Institutional support for student assessment: Methodology and results of a national survey*. Stanford, CA: Stanford University, National Center for Postsecondary Improvement.
- Peterson, M. W., Vaughan, D. S., & Perorazio, T. E. (2001). *Student assessment in higher education: A comparative study of seven institutions*. Stanford, CA: Stanford University, National Center for Postsecondary Improvement.
- Pettigrew, A. (1979). On studying organizational cultures. *Administrative Science Quarterly*, 24, 570-581.
- Pike, G. P. (2002). Measurement issues in outcomes assessment. In T. W. Banta (Ed.), *Building a scholarship of assessment* (pp. 131-147). San Francisco: Jossey-Bass.
- Pike, G. R. (2000). Assessment measures: Methodological issues in the assessment of learning communities. *Assessment Update*, 12(2), 14-15.
- Porter, S. R. (2004). Overcoming survey research problems. *New Directions for Institutional Research*, 121.

- Preskill, H. (2004). The transformational power of evaluation: Passion, purpose, and practice. In M. C. Alkin (Ed.), *Evaluation roots: Tracing theorists' views and influences* (pp. 343-355). Thousand Oaks, CA: Sage.
- Preskill, H., & Torres, R. T. (1999). *Evaluative inquiry for learning organizations*. Thousand Oaks, CA: Sage.
- Preskill, H., & Torres, R. T. (2000). The learning dimension of evaluation use. *New Directions for Evaluation*, 88, 25-37.
- Radcliffe-Brown, A. (1952). *Structure and function in primitive society*. London: Cohen & West.
- Riechers, A., & Schneider, A. (1990). Climate and culture: An evolution of constructs. In B. Schneider (Ed.), *Organizational climate and culture* (pp. 5-39). San Francisco: Jossey-Bass.
- Roach, A. T., & Kratochwill, T. R. (2004). Evaluating school climate and school culture. *Teaching Exception Children*, 37(1), 10-17.
- Roethlisberger, F., Dickson, W., Wright, H., Pforzheimer, C., & Western Electric Company (1939). *Management and the worker: An account of a research program conducted by the Western Electric Company, Hawthorne Works, Chicago*. Cambridge, MA: Harvard University Press.
- Rorty, R. (1989). *Contingency, irony, and solidarity*. Cambridge: Cambridge University Press.
- Rothwell, W., & Sullivan, R. (2005). *Practicing organization development: A guide for consultants*. San Francisco: Pfeiffer.
- Rubin, H., & Rubin, I. (2005). *Qualitative interviewing: The art of hearing data* (2nd ed.). Thousand Oaks, CA: Sage.
- Russ-Eft, D., & Preskill, H. (2001). *Evaluation in organizations: A systematic approach to enhancing learning, performance, and change*. Cambridge, MA: Perseus.

- Salancik, G. R., & Pfeffer, J. (1977). Who gets power—and how they hold on to it: A strategic-contingency model of power. *Organizational Dynamics*, 2-21.
- Salant, P. & Dillman, D. A. (1994). *How to conduct your own survey*. New York: John Wiley & Sons.
- Schein, E. (1985). *Organizational culture and leadership*. San Francisco: Jossey-Bass.
- Schein, E. (1991). What is culture? In P. Frost, L. Moore, M. Louis, C. Lundberg, & J. Martin (Eds.), *Reframing organizational culture* (pp. 243-253). Newbury Park, CA: Sage.
- Schein, E. (1992). *Organizational culture and leadership* (2nd ed.). San Francisco: Jossey-Bass.
- Schein, E. (1999). *The corporate culture survival guide: Sense and nonsense about cultural change*. San Francisco: Jossey-Bass.
- Schein, E. (2004). *Organizational culture and leadership* (3rd ed.). San Francisco: Jossey-Bass.
- Scheuren, F. (2004). *What is a survey*. Retrieved April 1, 2007, from <http://www.whatisasurvey.info>.
- Schultz, M., & Hatch, M. (1996). Living with multiple paradigms: The case of paradigm interplay in organizational culture studies. *Academy of Management Review*, 21, 529-557.
- Scriven, M. (1991). *Evaluation thesaurus* (4th ed.). Thousand Oaks, CA: Sage.
- Seiden, K. (2000). *Development and validation of the "Organizational Readiness for Evaluation" survey instrument*. Unpublished doctoral dissertation, University of Minnesota.
- Selznick, P. (1949). *TVA and the grass roots: A study in the sociology of formal organization*. Berkeley, CA: University of California Press.
- Senge, P. M. (1990). *The fifth discipline: The art and science of the learning organization*. New York: Currency/Doubleday.

- Seymour, D. T. (1992). *On q: Causing quality in higher education*. New York: Macmillian Publishing Company.
- Seymour, D. T. (1994). Total Quality Management on campus: Is it worth it? *New Directions for Higher Education*, 86.
- Shadish, W. R. (1998). Evaluation theory is who we are. *American Journal of Evaluation*, 19(1), 1-19.
- Shadish, W. R., Cook, T. D., Leviton, L. C. (1991). *Foundations of program evaluation: Theories of practice*. Newbury Park, CA: Sage.
- Sherr, L. A., & Teeter, D. J. (1991). Total Quality Management in higher education. *New Directions for Institutional Research*, 71.
- Shulman, L. S. (1997). Disciplines of inquiry in education: A new overview. In R. M. Jaeger (Ed.), *Complementary methods for research in education* (2nd ed.) (pp. 3-29). Washington, D.C.: American Educational Research Association.
- Smircich, L. (1983). Concepts of culture and organizational analysis. *Administrative Science Quarterly*, 28, 339-358.
- Smircich, L., & Calas, M. (1987). Organizational culture: A critical assessment. In F. Jablin, L. Putman, K. Roberts, & L. Porter (Eds.), *Handbook of organizational communication* (pp. 228-263). Newbury Park, CA: Sage.
- Smircich, L., & Morgan, G. (1982). Leadership: The management of meaning. *Journal of Applied Behavioral Science*, 18(3), 257-273.
- Southern Regional Education Board. (1985). *Access to quality undergraduate education*. Atlanta, GA: Southern Regional Education Board.

- State Higher Education Executive Officers. (2005). *Accountability for better results: A national imperative for higher education*. Boulder, CO: State Higher Education Executive Officers.
- Stevens, C. J., & Dial, M. (1994). Preventing the misuse of evaluation. *New directions for program evaluation*, 64.
- Study Group on the Conditions of Excellence in American Higher Education. (1984). *Involvement in learning: Realizing the potential of American higher education*. Washington, D.C.: U.S. Department of Education.
- Stufflebeam, D. L. (1991). The CIPP model for program evaluation. In G.F. Madaus, M. Scriven, & D. L. Stufflebeam (Eds.), *Evaluation models: Viewpoints on educational and human services evaluation* (pp. 117-141). Boston: Kluwer-Nijhoff.
- Stufflebeam, D. L. (2004). The 21st-century CIPP model: Origins, development, and use. In M.C. Alkin (Ed.), *Evaluation roots: Tracing theorists' views and influences* (pp. 245-266). Thousand Oaks: Sage.
- Tagg, J. (2003). *The learning paradigm college*. Bolton, MA: Anker.
- Task Force on Teaching and Career Development (2007). *A compact to enhance teaching and learning at Harvard*. Cambridge, MA: Harvard University.
- Taylor, J. H. (2002). Using assessment to build a culture of improvement. *The Presidency*, 5(2), 38-39.
- Teeter, D. J., & Lozier, G. G. (1993). Pursuit of quality in higher education: Case studies in Total Quality Management. *New Directions for Institutional Research*, 78.
- Terziovski, M., Howell, A., Sohal, A., & Morrison, M. (2000). Establishing mutual dependence between TQM and the learning organization: a multiple case study analysis. *The Learning Organization*, 7(1), 23-31.

- Torres, R. L., & Preskill, H. (2001). Evaluation and organizational learning: Past, present, and future. *American Journal of Evaluation*, 22(3), 387-395.
- Trice, H. M., & Beyer, J. M. (1984). Studying organizational culture through rites and ceremonies. *Academy of Management Review*, 9 (4), 653-669.
- Trice, H., & Beyer, J. (1993). *The cultures of work organizations*. Englewood Cliffs, NJ: Prentice Hall.
- Tsoukas, H. & Knudsen, C. (Eds.), (2003). *The Oxford handbook of organization theory*. Oxford: Oxford University Press.
- Tuckman, B. W. (1999). *Conducting educational research* (5th ed.). Orlando, FL: Harcourt.
- Turpin, R. S., & Sinacore, J. M. (Eds.). (1991). Multisite evaluations. *New Directions for Program Evaluation*, 50.
- Tushman, M., & O'Reilly, C. (1997). *Winning through innovation: A practical guide to leading organizational change and renewal*. Cambridge, MA: Harvard Business School Press.
- U.S. Department of Education, National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Washington, D.C.: U.S. Government Printing Office.
- U.S. General Accounting Office (1992). *Organizational culture: Techniques companies use to perpetuate or change beliefs and values* (GAO/NSLAD-92-105). Washington, DC: Government Printing Office.
- van Maanen, J. (1991). The smile factory: Work at Disneyland. In P. Frost, L. Moore, M. Louis, C. Lundberg, & J. Martin (Eds.), *Reframing organizational culture* (pp. 58-76). Newbury Park, CA: Sage.

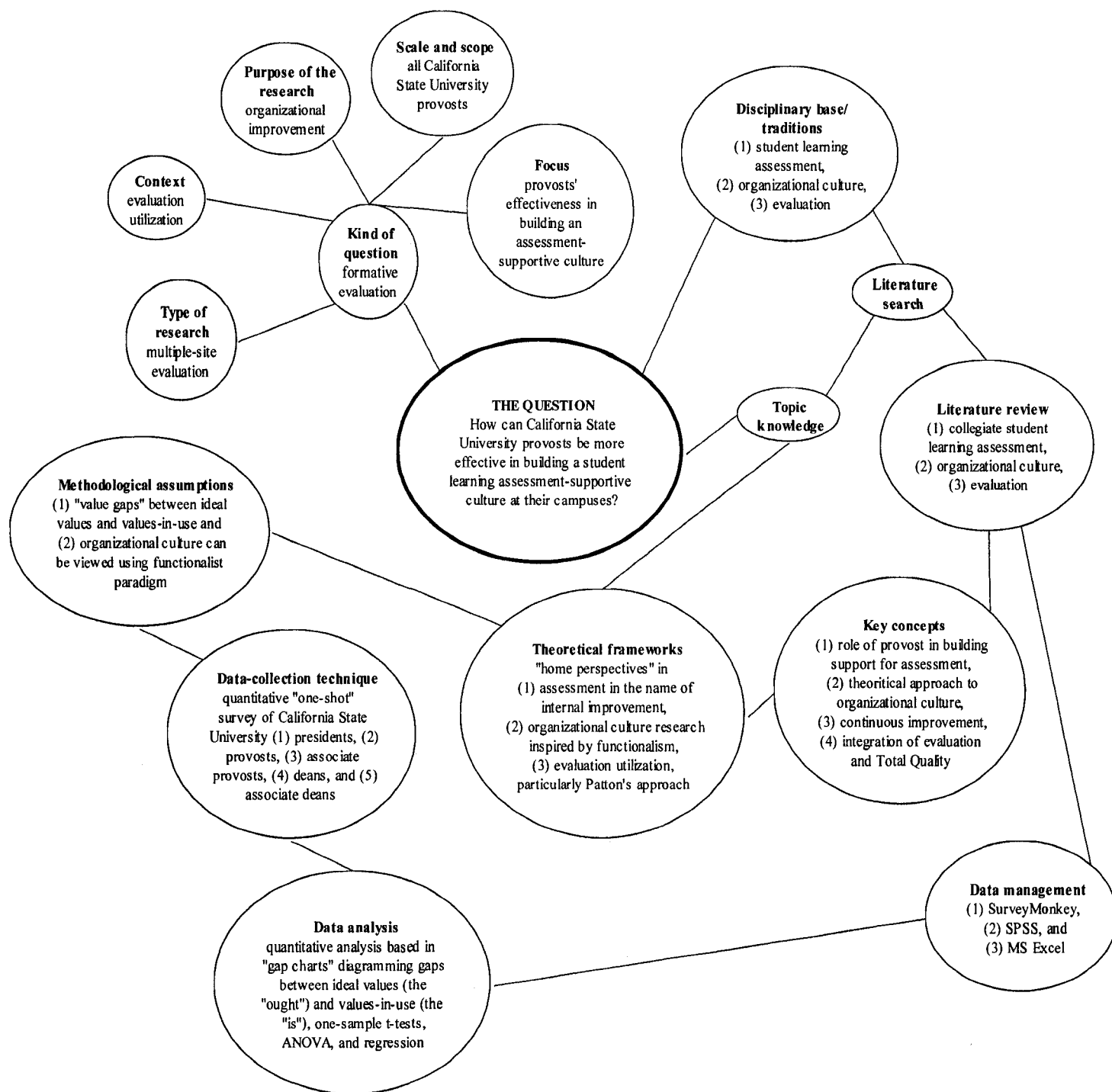
- Walvoord, B. E. (2004). *Assessment clear and simple: A practical guide for institutions, departments, and general education*. San Francisco: Jossey-Bass.
- Watt, J. H., Drennen, N. H., Rodrigues, R. J., Menelly, N., & Wiegel, E. K. (1993). Building assessment programs in large institutions. In T. W. Banta (Ed.), *Making a difference: Outcomes of a decade of assessment in higher education* (pp. 103-120). San Francisco: Jossey-Bass.
- Weber, M. (1964). *The theory of social and economic organization* (T. Parsons, Ed.). New York: Free Press.
- Weick, K. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, 21, 1-19.
- Weick, K. (1991). The vulnerable system: An analysis of the Tenerife air disaster. In P. Frost, L. Moore, M. Louis, C. Lundberg, & J. Martin (Eds.), *Reframing organizational culture* (pp. 117-130). Newbury Park, CA: Sage.
- Weiss, C. H. (1972). *Evaluation research: Methods for assessing program effectiveness*. Englewood Cliffs, NJ: Prentice-Hall.
- Weiss, C. H. (1977). *Using social research in public policy making*. Lexington, MA: Lexington Books.
- Weiss, C. H. (1998). *Evaluation: Methods for studying programs and policies* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.
- Weiss, C. H. (2004). Roots for evaluation: A cliff notes version of my work. In M.C. Alkin (Ed.), *Evaluation roots: Tracing theorists' views and influences* (pp. 153-168). Thousand Oaks: Sage.

- Welsh, J. F., & Metcalf, J. (2003). Cultivating faculty support for institutional effectiveness activities: Benchmarking best practices. *Assessment & Evaluation in Higher Education*, 28(1), 33-45.
- Wergin, J. F. (2005). Taking responsibility for student learning: The role of accreditation. *Change*, 37(1), 30-33.
- Westbrook, J. D. (1993). Organizational culture and its relationship to TQM. *Industrial Management*, 35(1), 1-3.
- Western Association of School and Colleges. (2001). *WASC handbook of accreditation*. Retrieved March 1, 2007 from http://www.wascenior.org/wasc/Doc_Lib/2001%20Handbook.pdf
- Western Association of Schools and Colleges. (2007). Retrieved February 1, 2007, from <http://www.wascenior.org/wasc>.
- Wholey, J. S. (1983). *Evaluation and effective public management*. Boston: Little, Brown.
- Wholey, J. S. (2004). Using evaluation to improve performance and support policy decision making. In Alkin, M. C. (Ed.), *Evaluation roots: Tracing theorists' views and influences* (pp. 267-275). Thousand Oaks, CA: Sage.
- Whyte, W. (1943). *Street corner society: The social structure of an Italian slum*. Chicago: University of Chicago Press.
- Williford, A. M. (1997). Ohio University's multidimensional institutional impact and assessment plan. *New Directions for Higher Education*, 100, 47-57.
- Willmott, H. (1994). Bringing agency (back) into organizational analysis: responding to the crises of (post)modernity. In J. Hassard & M. Parker (Eds.), *Toward a new theory of organizations* (pp. 87-130). London: Routledge.

- Willmott, H. (2003). Organization theory as a critical science? Forms of analysis and 'new organization forms.' In H. Tsoukas & C. Knudsen (Eds.), *The Oxford handbook of organization theory* (pp. 88-112). Oxford: Oxford University Press.
- Wilson, J. Q. (1989). *Bureaucracy: What government agencies do and why they do it*. New York: Basic Books.
- Wright, B. D. (2002). Accreditation and the scholarship of assessment. In T. W. Banta (Ed.), *Building a scholarship of assessment* (pp. 240-258). San Francisco: Jossey-Bass.
- Young, C. C., & Knight, M. E. (1993). Providing leadership for organizational change. In T. W. Banta (Ed.), *Making a difference: Outcomes of a decade of assessment in higher education* (pp. 25-39). San Francisco: Jossey-Bass.

Appendix A:
Dissertation Concept Map

Dissertation Concept Map*



*Concept map categories adapted from Hart (1998).

Appendix B:
Survey Cover Letter



Survey Cover Letter

October 8, 2007

Dear California State University administrator:

As a student who began college at San José State, worked at San Diego State for several years, and has had a longtime interest in the CSU system, I ask for your assistance. I am currently a doctoral candidate at the University of San Diego working to complete a dissertation about the role of CSU provosts in building a student learning assessment-supportive organizational culture.

Enclosed is a survey that evaluates the effectiveness of CSU provosts in building an assessment-supportive culture by measuring the “value gaps” between actual practice and the ideal. Your response to the 27 statements that make up the survey should take less than 10 minutes. The CSU Chancellor’s Office has reviewed this survey and all CSU provosts have been informed about it.

This survey guarantees your anonymity and I ask for no personal information. However, I do ask for the name of your CSU campus, which will not be identified by name in my study. If you are interested in seeing the results of this survey, or if you have any questions about this study, please contact me at cprocello@sandiego.edu. Thank you for your help.

Sincerely,

Chris Procello, Ed.M.

Enclosures 2

Chris Procello, Academic Assessment Coordinator
Office of the Vice President and Provost
5998 Alcalá Park, San Diego, CA 92110-2492
Ph: (619) 260-7976
Email: cprocello@sandiego.edu

Appendix C:
Survey Instrument

Survey Instrument*

I have read and understand the enclosed Research Participant Consent Page and give my consent for the information that I provide to be used in this study:

- Yes
 No

Please read each of the following statements and answer the related question about the effectiveness of your provost in building an organizational culture supportive of the statement. Answer the question by selecting a number that represents your evaluation, with 1 being “minimally effective” and 10 being “very effective.”

1) Our senior administration has similar goals from year to year.
 How effective is your provost/VP for academic affairs in building an organizational culture supportive of this statement?
 (“minimally effective”) 1 2 3 4 5 6 7 8 9 10 (“very effective”)

2) Each of our jobs includes working on projects aimed at university-wide improvement.
 How effective is your provost/VP for academic affairs in building an organizational culture supportive of this statement?
 (“minimally effective”) 1 2 3 4 5 6 7 8 9 10 (“very effective”)

3) Our administrators use feedback from faculty to improve their work.
 How effective is your provost/VP for academic affairs in building an organizational culture supportive of this statement?
 (“minimally effective”) 1 2 3 4 5 6 7 8 9 10 (“very effective”)

4) There is ongoing collaborative work across administrative units in our university.
 How effective is your provost/VP for academic affairs in building an organizational culture supportive of this statement?
 (“minimally effective”) 1 2 3 4 5 6 7 8 9 10 (“very effective”)

5) When educational effectiveness is less than optimal, we try to identify where our systems are failing.
 How effective is your provost/VP for academic affairs in building an organizational culture supportive of this statement?
 (“minimally effective”) 1 2 3 4 5 6 7 8 9 10 (“very effective”)

6) In setting educational outcomes, our administrators consider the overall vision and goals of the university.
 How effective is your provost/VP for academic affairs in building an organizational culture supportive of this statement?
 (“minimally effective”) 1 2 3 4 5 6 7 8 9 10 (“very effective”)

7) Our faculty use data to assess teaching and learning.
 How effective is your provost/VP for academic affairs in building an organizational culture supportive of this statement?
 (“minimally effective”) 1 2 3 4 5 6 7 8 9 10 (“very effective”)

8) We believe that improving the quality of education for students does not require more money.

How effective is your provost/VP for academic affairs in building an organizational culture supportive of this statement?

("minimally effective") 1 2 3 4 5 6 7 8 9 10 ("very effective")

9) Our long-term educational objectives do not change with the composition of the senior administrative team.
How effective is your provost/VP for academic affairs in building an organizational culture supportive of this statement?
("minimally effective") 1 2 3 4 5 6 7 8 9 10 ("very effective")

10) We understand our university mission as it applies to our work.

How effective is your provost/VP for academic affairs in building an organizational culture supportive of this statement?

("minimally effective") 1 2 3 4 5 6 7 8 9 10 ("very effective")

11) When changes are proposed, we use data to support the proposal.
How effective is your provost/VP for academic affairs in building an organizational culture supportive of this statement?
("minimally effective") 1 2 3 4 5 6 7 8 9 10 ("very effective")

12) Our administrators' approach to student learning is that learning can be improved without increasing the budget.

How effective is your provost/VP for academic affairs in building an organizational culture supportive of this statement?

("minimally effective") 1 2 3 4 5 6 7 8 9 10 ("very effective")

13) When someone demonstrates a better way, we tend to support it.
How effective is your provost/VP for academic affairs in building an organizational culture supportive of this statement?
("minimally effective") 1 2 3 4 5 6 7 8 9 10 ("very effective")

14) When something goes wrong, we typically look for the cause in our processes rather than in specific employees.

How effective is your provost/VP for academic affairs in building an organizational culture supportive of this statement?

("minimally effective") 1 2 3 4 5 6 7 8 9 10 ("very effective")

15) We use a breakdown of students' educational achievement to evaluate the university's educational effectiveness.
How effective is your provost/VP for academic affairs in building an organizational culture supportive of this statement?
("minimally effective") 1 2 3 4 5 6 7 8 9 10 ("very effective")

16) We assume some personal responsibility when our university improvement goals are not met.

How effective is your provost/VP for academic affairs in building an organizational culture supportive of this statement?

("minimally effective") 1 2 3 4 5 6 7 8 9 10("very effective")

17) Our faculty can make changes to improve student learning without additional resources.
How effective is your provost/VP for academic affairs in building an organizational culture supportive of this statement?
("minimally effective") 1 2 3 4 5 6 7 8 9 10("very effective")

18) The university's overall vision and goals guide our day-to-day work with regard to educational effectiveness.

How effective is your provost/VP for academic affairs in building an organizational culture supportive of this statement?

("minimally effective") 1 2 3 4 5 6 7 8 9 10("very effective")

19) We change things when they are not working.
How effective is your provost/VP for academic affairs in building an organizational culture supportive of this statement?
("minimally effective") 1 2 3 4 5 6 7 8 9 10("very effective")

20) We test our assumptions about the causes of student learning with data.

How effective is your provost/VP for academic affairs in building an organizational culture supportive of this statement?

("minimally effective") 1 2 3 4 5 6 7 8 9 10("very effective")

21) When there is a problem in the classroom, our administrators identify where the systems need to be improved to keep the problem from happening again.
How effective is your provost/VP for academic affairs in building an organizational culture supportive of this statement?
("minimally effective") 1 2 3 4 5 6 7 8 9 10("very effective")

22) Our faculty not only teach their classes, but constantly improve them.

How effective is your provost/VP for academic affairs in building an organizational culture supportive of this statement?

("minimally effective") 1 2 3 4 5 6 7 8 9 10("very effective")

23) Our faculty frequently have conversations about student learning with colleagues from other academic units.
How effective is your provost/VP for academic affairs in building an organizational culture supportive of this statement?
("minimally effective") 1 2 3 4 5 6 7 8 9 10("very effective")

24) The university's long-term objectives should not change every time we get a new senior administrator.

How effective is your provost/VP for academic affairs in building an organizational culture supportive of this statement?

("minimally effective") 1 2 3 4 5 6 7 8 9 10 ("very effective")

25) We are responsible for improving things within the university, even when these things do not directly involve our individual work.
 How effective is your provost/VP for academic affairs in building an organizational culture supportive of this statement?
 ("minimally effective") 1 2 3 4 5 6 7 8 9 10 ("very effective")

26) Work time is structured to provide our faculty with opportunities to work with other faculty members.

How effective is your provost/VP for academic affairs in building an organizational culture supportive of this statement?

("minimally effective") 1 2 3 4 5 6 7 8 9 10 ("very effective")

27) When we introduce a major university improvement program, we give it at least 3 years to show results.
 How effective is your provost/VP for academic affairs in building an organizational culture supportive of this statement?
 ("minimally effective") 1 2 3 4 5 6 7 8 9 10 ("very effective")

28) Please indicate your CSU campus:

- Bakersfield
- Channel Islands
- Chico
- Dominguez Hills
- East Bay
- Fresno
- Fullerton
- Humboldt
- Long Beach
- Los Angeles
- Maritime Academy
- Monterey Bay
- Northridge
- Pomona
- Sacramento
- San Bernardino
- San Diego
- San Francisco
- San Jose
- San Luis Obispo
- San Marcos
- Sonoma
- Stanislaus

Thank you for taking the time to complete this survey and, again, I can be reached at cprocello@sandiego.edu if you are interested in the survey's final results.

* Survey instrument adapted from Detert, Schroeder, and Cudeck (2003).

Appendix D:
List of Survey Statements as Connected to Total Quality Culture Construct

Survey Statements as Connected to Total Quality Culture Construct

Construct 1: Shared vision

- 6. In setting educational outcomes, our administrators consider the overall vision and goals of the university.
- 10. We understand our university mission as it applies to our work.
- 18. The university's overall vision and goals guide our day-to-day work with regard to educational effectiveness.

Construct 2: Long-term focus

- 1. Our senior administration has similar goals from year to year.
- 9. Our long-term educational objectives do not change with the composition of the senior administration.
- 24. The university's long-term objectives should not change every time we get a new senior administrator.
- 27. When we introduce a major university improvement program, we give it at least 3 years to show results.

Construct 3: Continuous improvement

- 3. Our administrators use feedback from faculty to improve their work.
- 13. When someone demonstrates a better way, we tend to support it.
- 19. We change things when they are not working.
- 22. Our faculty not only teach their classes, but constantly improve them.

Construct 4: Involvement

- 2. Each of our jobs includes working on projects aimed at university-wide improvement.
- 16. We assume some personal responsibility when our university improvement goals are not met.
- 25. We are responsible for improving things within the university, even when these things do not directly involve our individual work.

Construct 5: Collaboration

- 4. There is ongoing collaborative work across administrative units in our university.
- 23. Our faculty frequently have conversations about student learning with colleagues from other academic units.
- 26. Work time is structured to provide our faculty with opportunities to work with other faculty members.

Construct 6: Data-based decision-making

- 7. Our faculty use data to assess teaching and learning.
- 11. When changes are proposed, we use data to support the proposal.
- 15. We use a breakdown of students' educational achievement to evaluate the university's educational effectiveness.
- 20. We test our assumptions about the causes of student learning with data.

Construct 7: Systems focus

- 5. When educational effectiveness is less than optimal, we try to identify where our systems are failing.

14. When something goes wrong, we typically look for the cause in our processes rather than in specific employees.

21. When there is a problem in the classroom, our administrators identify where the systems need to be improved to keep the problem from happening again.

Construct 8: Quality at the same cost

8. We believe that improving the quality of education for students does not require more money.

12. Our administrators' approach to student learning is that learning can be improved without increasing the budget.

17. Our faculty can make changes to improve student learning without additional resources.