California Postsecondary Education Commission

A Prospectus for Evaluating the Cost-Effectiveness of University Academic Preparation Programs

September 2004

It has been over eight years since the Commission authored a report on the effectiveness of university academic preparation programs. Since then, funding has been reduced, new programs have been added, and new philosophies have emerged. Given these occurrences, the time is right for the Commission to conduct a comprehensive study.

Contents

The Commission advises the Governor and Legislature on higher education policy and fiscal issues. Its primary focus is to ensure that the state's educational resources are used effectively to provide Californians with postsecondary education opportunities. More information about the Commission is available at www.cpec.ca.gov.

Commission Report 04-14

Introduction – The Public Policy Context of Program Evaluation

> *Like many poor people, evaluation in the United States has grown up in the projects—federal projects spawned by the Great Society legislation of the 1960s (Patton, 2000).*

Increasingly, public officials in California and across the nation want to know, or at least be reassured, that public investments in what had traditionally been called *university outreach programs* are having desired effects on student achievement and college preparedness among targeted student groups. For the most part, such programs are now commonly referred to as *academic preparation programs.* Given the Governor's 2004-05 budget proposal to reduce state support for these preparation programs, concerns about program effectiveness, as well as questions regarding which specific groups of students should be benefited by scarce public resources, have become even more important to lawmakers.

This *Prospectus* summarizes the results of recent outreach evaluative studies conducted independently since 1990 by the California Postsecondary Education Commission (CPEC), the Policy Analysis for California Education (PACE) research group, and the Strategic Review Panel on UC Educational Outreach. Following the summary, a mixed-methods evaluative framework is outlined, which the Commission believes could be used to make more in depth assessments of the short-term and long-term effects of preparation programs for public accountability purposes. The framework was selected because the Commission intends to collect and analyze both quantitative and qualitative data

and use various statistical procedures to measure program effectiveness. Implementing the framework will require a high degree of Commission collaboration with the higher education systems and the California Department of Education.

Public Accountability is referenced here to make clear a primary intent of the framework, which is to respond to a set of fundamental questions raised by public officials regarding the impact of academic preparation programs. Collecting more useful information will avoid a potential gap between generating evaluation findings and using those findings by public officials for intended purposes. It is quite likely that the proposed evaluation process would be of benefit to other states that have similar research interests and evaluation needs.

In developing this prospectus, the Commission was mindful of the general inability of evaluative studies to influence government decision making during the second half of the 20th Century. During the 1960s and 1970s, the United States had experienced a rapid increase in federal, state, and county programs that were funded to address a number of pressing social concerns, including poverty, disease, joblessness, mental instability, crime, hunger, school performance, and educational equity. As the dollar investment in those programs increased, public officials wanted some degree of assurance that public funds were being spent wisely and that social programs were effective and yielding desired consequences.

Largely in response to federal public accountability concerns, the practice of evaluation was born with the high expectation that such practices would (a) provide credible and useful data with which results could be judged, (b) discern the true nature of cause and effect relationships, (c) critically examine the cost-effectiveness of competing alternatives so that decision-makers could identify the most effective alternative, and (d) offer recommendations for improving program performance and strategic planning. Federal dollars for evaluation grew dramatically during this formative period. In 1977 alone, for example, federal agencies spent approximately $64 million on program evaluation studies (Abramson, 1978).

Display 1 below presents the three primary uses of evaluation.

DISPLAY 1 Three Primary Uses of Evaluation Findings

Uses	**Examples**
Judge merit or worth	Summative Evaluation Accountability Audits Quality Control Cost-effectiveness decisions Decide a program's future Accreditation/licensing
Improve programs	Formative evaluation Identification of strengths and weakness Continuous improvement Quality enhancement Being a learning organization Managing more effectively Adapting a model locally
Generate Knowledge	Generalizations about effectiveness Extrapolating principles about what works Theory building Synthesis of patterns across programs Scholarly publishing Policy making

Although evaluation practices have enhanced sound decision making in a few limited program areas, it is generally recognized that the majority of evaluative studies have been largely ignored when determining future policy directions. Not surprisingly, public dollars appropriated for evaluation have been on a spiraling decline since the early 1980s, and there has been a general shift from program evaluation to program inspection, auditing, and investigation. While the reasons are many and varied for the lack of evaluation, three reasons seem to loom large:

- First, the Federal General Accounting Office (GAO) in a 1995 report noted that evaluative information has not always been organized and communicated effectively, and that it either had not reached appropriate legislative committee members in a timely manner, or it had been received in a form too difficult to digest.
- Second, it has been observed that far too many evaluations have been directed towards general abstract users, rather than directing studies toward actual primary intended users and their explicit commitments to concrete, specific uses (Patton, 2000). In other words, is the evaluation to be used, as shown in Display 1, to (a) judge the merit or worth of a program, (b) improve a program, or (c) generate knowledge about a program? Rarely, is an evaluation intended to address all three areas simultaneously.
- A third reason for the lack of evaluation is that evaluators often are not attentive to what is referred to as the *personal factor*. That factor calls for evaluators to work with intended users in determining how an evaluation will be used. In other words, if an evaluation is not directed towards the needs of potential users, it is not likely that the results will have much utility. Further, no evaluation will garner much attention unless key decision makers understand, value, and care about evaluation.

To maximize utility and avoid the evaluation pitfalls just discussed, the Commission's evaluation framework proposes to embrace five key principles shown in Display 2. These principles will be discussed in the design section of this report.

DISPLAY 2 Guiding Principles of the Commission's Evaluation Framework

Focus of the Principle	**Actualizing the Principle**
Intended Users	Primary intended users will be identified to provide the focus of the evaluation and to generate a shared commitment to intended uses. At least some of the primary users will need to understand and value evaluation and will be enthusiastic, committed, and assertive.
Clarity	Every effort will be made to be clear about goals, purposes, what precisely what will be evaluated and judged, and how results will be used. Clarification also will involve determining what is worth doing and knowing.
Systematic Procedures	Work will be planned carefully and systematically, and the Commission will document what occurs at every stage of the decision-making and data collection process.
Design Utility and Statistical Procedures	A mix-methods approach will be used wherein the actual design methods (descriptive, quasi-experimental) and statistical procedures (e.g., correlation, regression, analysis of variance) will vary depending on the specific evaluation questions being addressed.
Program Effect Size	Where possible, program effects will be standardized and translated into cost-effect ratios so that policy makers can judge the cost-effectiveness of various types of outcomes.

Background—Understanding Academic Preparation Programs

Although public awareness of academic preparation programs has increased recently, many taxpayers and lawmakers are not fully informed about the breadth of programs offered and the wide range of goals, objectives, strategies and rationales, approaches, and foci associated with them. Some policy makers have mistakenly presumed that a single evaluation approach or research design could be used reliably to appraise a wide range of intended student outcomes and program purposes. This background section is included to enhance an understanding of academic preparation programs and their intended public value.

The term, *university outreach*, has been used to describe a wide range of short-term, intermediate-term, and long-term intervention programs and initiatives that have sought to assist students in overcoming educational and socio-economic disadvantages so that student learning, academic achievement, and college-going behaviors would be maximized. In the past, it has not always been understood that traditional outreach services have involved much more than the dissemination of information related to university admission and eligibility. Because of this misunderstanding, such programs are now referred to as academic preparation programs to emphasize that the principal goal of *outreach* is to improve the academic achievement of educationally disadvantaged and underrepresented student groups.

The term, *educationally disadvantaged*, refers to the high concentration of various demographic student groups in schools facing exceptional challenges, as measured by factors such as the school's academic performance index score (API), overall college-going rate, proportion of teachers who are not fully credentialed, and limited college preparatory course offerings. Since at least the 1960s, the California State University (CSU) and the University of California (UC) have recognized that with the growing number of poor and underrepresented student groups in the State, academic achievement could not be adequately addressed independent of the teaching and curricula in schools. Thus, the basic philosophy of academic improvement programs has always been that demonstrative progress can only be sustained by collaborative alliances that include the two public university systems, the K-12 system, private business, and philanthropic partners.

A recent example of a successful collaborative alliance is one that resulted in the development of the CSU Early Assessment Program. That program is designed to bridge the gap between high school standards and university expectations. Specifically, CSU placement standards in language arts and mathematics were incorporated earlier this year into the *California Standards Test* and the *Golden State Examinations*. Students who elect to participate in the program and take the tests during their eleventh grade of high school will receive a report indicating whether or not they are ready for CSU level work. If they are not ready, the test will identify areas where they would need additional preparation during their senior year to enter the CSU ready for college. CSU faculty members, high school teachers, and representatives from the State Board of Education and the California Department of Education all participate in this process.

Although some degree of overlap exists, outreach programs generally tend to have different goals and objectives, and they target various types of students at various points in the educational pipeline.

Examples of *short-term initiatives* include informational outreach that provides timely information to students, families, teachers, and counselors to improve planning and preparation for college. Those initiatives use a variety of media formats, such as publications, videos, and interactive digital protocols.

Examples of *intermediate initiatives* include the Mathematics, Science, and Engineering Program (MESA), which assists low-performing schools in preparing students for university instruction in mathematics- and science-based fields; and the Summer Bridge Program, which is intended to improve the college persistence of disadvantaged persons by assisting incoming fresh-

men through an intensive residential program during the summer immediately prior to college/university enrollment.

Perhaps most significant are *long-term*, *school-centered programs* of the University of California and the California State University that seek to address the underlying causes of low college/university eligibility and enrollment among disadvantaged students by working in regional partnerships to foster a culture that promotes academic success (UC Board of Regents Outreach Task Force, 1997). An example of a long-term intervention strategy is the University of California's School-University Partnership (SUP) program, in which UC campuses partner with individual public schools in providing a range of long-term interventions relating to such areas as professional development, curricular reform, academic planning, tutoring, and technology-mediated learning.

Recent External Evaluation Efforts

Findings of The Commission's 1992 and 1996 Studies

In 1992, CPEC assessed the effectiveness of collaborative academic development programs that had a collective goal of increasing college preparation among student groups that had low university eligibility rates and low participation rates. In this report, *Final Report on the Effectiveness of Intersegmental Student Preparation Programs*, using *participation rates* and *institutional collaboration* as the sole measures of effectiveness, the Commission concluded that programs had met their educational equity goals, and that they had enhanced collaboration between public schools and postsecondary institutions.

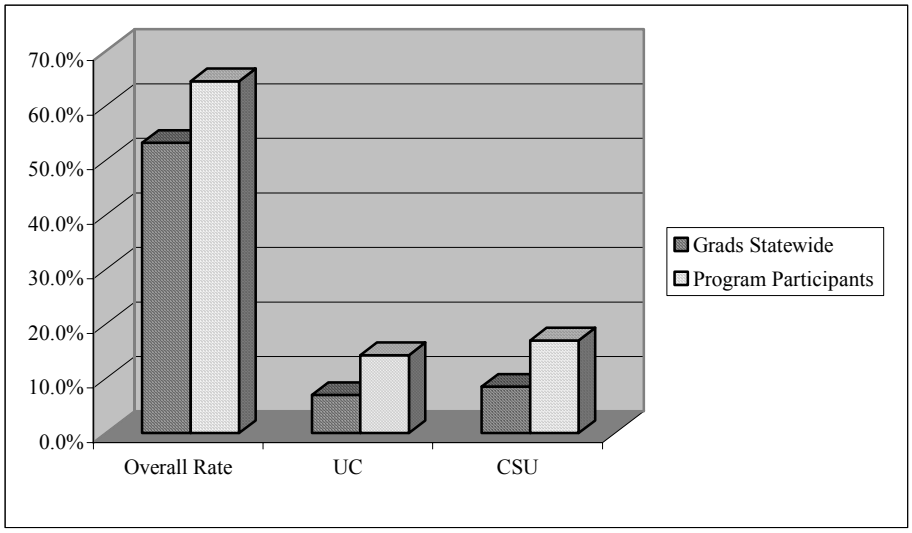
Although the evidence reviewed supported the Commission's general findings, there appears to have been at least four limitations of the 1992 study. First, the data were not disaggregated by gender or racial-ethnic group, so it was not possible to determine if the programs were equally effective for men and women and for various ethnic groups. Second, the study did not determine if underrepresented students that participated in the programs were better prepared for mathematics- and science-based fields than those students that did not participate in an academic preparation program. Third, no longitudinal data were collected, so it was not possible to determine the effect of programs over time. Fourth, the study did not control for the length of time a participant spent in a program, so it was not possible to determine the relationship between time spent in a program and academic preparedness.

In 1996, the Commission conducted a follow-up study to examine the increase in participation rates that might have occurred between 1992 and 1996. The nine collaborative academic development programs reviewed were:

- Advancement via Individual Determination (AVID)
- Alliance for Collaborative Change in Education in School Success (ACCESS)
- California Academic Partnership Program (CAPP)
- California Student Opportunity and Access Program (Cal SOAP)
- College Readiness Program (CRP)
- Early Academic Outreach Program (EAOP)
- Mathematics, Engineering, Science Achievement (MESA)
- Middle College (MC)
- Urban School-Community Collaborative (UCSCol)

The study design in the 1996 study was similar to the 1992 study in that no longitudinal data were collected; however, one-year snapshots, shown graphically in Display 3, proved useful.

DISPLAY 3 College and University Participation Rates of Public High School Academic Preparation Program Participants Compared with Graduates Statewide, Fall 1994



Overall, of the program participants that graduated from public high schools in 1994, 64.5% enrolled in postsecondary institutions as first-time freshmen following graduation. This compares favorably to a statewide rate of 53.2%, which translates to a program effect of more than 11 percentage points.

By university system, 14.3% of the program participants enrolled at the University of California in Fall 1994, compared with 7.0% of the public high school graduates statewide. Seventeen percent of the program participants enrolled at the California State University, in comparison to 8.5% of the graduates statewide. Display 4 shows similar data for selected preparation programs.

DISPLAY 4 Postsecondary Participation Rates for Graduates of Academic Programs Compared with the Participation of all California Public High School Graduates, Fall 1994, in Percents

	Comparison Group	**Selected Academic Preparation Programs**				
California Postsecondary Institutions	1994 State Graduates	1995 AVID	1994 ACCESS	1994 Cal-SOAP	1994 EAOP	1995 MESA
University of California	7.3	11.0	6.1	8.6	18.7	16.1
California State University	8.5	41.0	11.7	13.5	17.9	14.6
California Community Colleges	35.2	35.0	39.6	42.0	20.2	11.5
Total California Public Higher Education	**51.0**	**87.0**	**57.4**	**64.1**	**56.8**	**42.2**
Independent California Institutions	2.2	11.0	N/A	7.3	5.3	5.7

Note: Comparison Group with Low Eligibility consists of African American, Latino, and Native American students.

Again, although these data reflect only a one-year snapshot, they are consistent with the aggregate results just mentioned. For example, of the Early Academic Outreach Program participants, approximately 19% enrolled at the University of California as first-time freshmen in Fall 1994, and another 18% enrolled at the California State University. In contrast, about 7.3% of the public high school

graduates statewide enrolled at the University of California in Fall 1994, and another 8.5% enrolled at the California State University. Although these data demonstrate that many of the academic preparation programs have been effective in increasing college and university participation, the results should be treated with a degree of caution because of the four design limitations discussed previously. The Commission's mix-methods design, which will be discussed in the next section, proposes to correct for such limitations.

Findings of the 1997 PACE Study

In 1997, the Policy Analysis for California Education (PACE) research group was commissioned by the University of California to (a) synthesize existing outreach efforts, (b) identify effective practices, and (c) make recommendations to improve programs and evaluation practices. Based on the data and information reviewed, PACE observed that although outreach programs are beneficial to students, very little evidence existed to help policy makers make decisions about which programs are most effective and where scarce resources should be invested. It also was noted that few evaluations had been conducted rigorously. The following specific conclusions were reported:

- Long-term outcomes are rarely measured.
- Evaluations are not specific about how the various program components and performance indicators fit together.
- Virtually no attempts have been made to detect and consider systematic comparative data.

The 2002 Strategic Review Panel on UC Educational Outreach Study

A strategic review panel was established by the University of California in Fall 2002 to: (a) assess the effectiveness of the university's outreach programs, (b) recommend desirable changes, (c) set reasonable short-term and long-term goals, and (d) recommend a new working alliance with the state's K-12 educators and with California community colleges. Study teams were convened to conduct analyses related to four major elements of the University's outreach enterprise:

- Individual and collective program effectiveness
- Organization and governance
- Accountability systems
- Opportunities for collaboration

With regard to program effectiveness in 1997, UC set five-year eligibility goals for several academic preparation programs, including the *Early Academic Outreach Program (EAOP), Mathematics, Engineering (MESA)*, and the *Puente Project*. One major goal was to double the number of program participants that had met all UC admission requirements upon high school graduation, and to increase by half the number of program graduates that were competitively UC eligible. A general finding of the review panel was that student academic development programs were making progress towards these goals. For example, between 1998-99 and 2001-02, the number of UC-eligible participants increased 51% in EAOP, 85% in MESA, and 81% in the Puente Project. It was noted that progress varied across racial-ethnic groups. But even for those programs showing progress across ethnic groups, it is unclear whether improvement in eligibility was due to the programs, or was a result of the programs increasingly admitting more competitive students at the outset. To establish a true program effect, the panel could have provided or cited three important statistics: (1) the mean grade point average (GPA) for each student cohort at the time of acceptance into the program, (2) the mean change in GPA over time in college preparatory courses, and (3) the mean length of time spent in a program for each student cohort.

Commission's Mixed-Methods Evaluation Design

The Commission's evaluation design entails six preliminary steps that will be undertaken prior to analyzing data. These steps are:

- identify key potential users of the evaluation results,
- clarify what specifically public officials want to know about academic preparation programs,
- identify key uses of the evaluation results,
- determine data needs and requirements,
- select appropriate design methods and statistical procedures, and
- establish design principles to guide the evaluation.

These six preliminary processes are described in detail below.

Identifying key potential users of evaluation results

Although the California Legislature is responsible for making funding decisions related to public services and programs, some members have special interests in certain service and program areas. The first critical step the Commission intends to undertake is to identify those public officials and legislative aides that have both a keen or special interest in academic preparation programs and an interest in the effectiveness of these programs in meeting broad public goals. Once these public officials and legislative aides are identified, the Commission will work with them to develop a shared commitment to intended uses of the evaluation results.

Clarifying what public officials want to know about academic preparation programs

In reviewing the transcripts of recent legislative hearings on university outreach and academic preparation programs, it seems that policy makers would be most interested in obtaining in-depth knowledge about various programs so that they could make more informed judgments about the cost-effectiveness and merit of those programs.

At a minimum, the questions addressed in the Commission's 1992 and 1996 studies will be explored again, but in much greater detail. That is, trend data will be reported, and to the greatest extent possible, outcome data will be disaggregated by relevant demographic and socio-economic factors.

Determining data needs and requirements

Reporting on the effectiveness of academic preparation programs in depth as proposed in this prospectus will require collecting a rich mixture of qualitative and quantitative information from a variety of sources. The Commission's evaluation study team will determine what data elements are required to address various evaluative questions and how difficult and costly it would be to obtain those data. A major challenge of conducting longitudinal outreach studies is the difficulty and cost of tracking students from first entry into a preparation program through eventual graduation from a university or college. Despite this inherent difficulty, if public officials truly want know the long-term effects of academic preparation programs on baccalaureate attainment, as well as their impact on California's economy, part of the Commission's study will require such longitudinal data.

Selecting appropriate design methods and statistical procedures

Several methods and statistical procedures will be used to address the full range of anticipated questions. Some will involve a quantitative approach and some will require a more qualitative approach, such as the use of focus groups. In general, most designs can be classified as *descriptive*, *experimental*, or *quasi-experimental*. Many social scientists prefer experimental studies whenever the environmental conditions are suitable, because the results emanating from such studies will usually have a higher de-

gree of validity and reliability. This is because key statistical and environmental controls are built into those designs.

Because students are not randomly assigned to academic preparation programs, and because most school and university settings are not completely suitable for experimental studies, the Commission anticipates that the designs used in deriving outcome data will be descriptive and quasi-experimental in nature. Practically all research studies conducted by the U.S. Census Bureau and other social research agencies are categorized as either descriptive or quasi-experimental.

Pending the availability of funds, and the availability of requisite data, the Commission intends to collect both longitudinal and cross-sectional data. Longitudinal analysis would involve tracking a cohort of students over time, whereas cross-sectional analysis would entail tracking outcome data over time that would not necessarily involve the same cohort of students. Both analyses would yield useful data and would contribute significantly to the Commission's work.

Establishing design principles to guide the evaluation study

The design principles shown in Display 5 were developed following an extensive review of the literature on program evaluation at the outset of this prospectus. They are restated and discussed here because they are crucial to the success of this endeavor.

One might assume that all scholarly evaluation efforts would embrace these principles to some degree; however, program evaluations have not always been focused properly on the needs of the primary users. They have often been vague on why something is worth knowing, procedures have not always been planned carefully and systematically, and rarely have they involved a healthy discussion of the cost-effectiveness of program outcomes. The end result has been that such evaluations have rarely informed decision-making. The Commission believes that strict adherence to these guiding principles will reduce the likelihood of a potential gap between generating evaluation findings and the use of those findings by public officials for intended purposes.

DISPLAY 5 Guiding Principles of the Commission's Evaluation Framework

Principle	Actualizing the Principle
Intended Evaluation Users	Primary intended users will be identified to provide the focus of the evaluation and to generate a shared commitment to intended uses. At least some of the primary users will need to understand and value evaluation and be enthusiastic, committed, competent, and assertive.
Clarity	Every effort will be made to be clear about goals, purposes, what precisely what will be evaluated and judged, and how results will be used. Clarification also will involve determining what is worth doing and knowing.
Systematic Procedures	Work will be planned carefully and systematically, and the Commission will document what occurs at every stage of the decision-making and data collection process.
Design Utility and Statistical Procedures	A mix-methods approach will be used wherein the actual design methods (descriptive, quasi-experimental) and statistical procedures (e.g., correlation, regression, analysis of variance) will vary depending on the specific evaluation questions being addressed.
Program Effect Size	Where possible, program effects will be standardized and translated into cost-effect ratios so that policy makers can judge the cost-effectiveness of various types of outcomes.

Note: Same as Display 2.

Next Steps

Over the next several months, staff will discuss this prospectus with the public higher education systems and with other agencies that have periodically examined the effectiveness of preparation programs. Those research agencies include the Policy Analysis for California Education (PACE), the College Board, the Achievement Council of Los Angeles, Supporters of Educational Equality and Diversity (SEED), and the All Campus Consortium for Research for Diversity of the University of California (UC ACCORD).