

Final Draft — July 1999

Prepared for the Vocational Education Unit, Educational Services and Economic
Development Division, State Chancellor's Office, California Community Colleges

PERFORMANCE MEASURES FOR IMPROVING
CALIFORNIA COMMUNITY COLLEGES:
ISSUES AND OPTIONS

W. Norton Grubb
David Gardner Chair in Higher Education

Norena Badway

Community College Cooperative and
National Center for Research in Vocational Education
School of Education
University of California, Berkeley

Acknowledgments

This report has been commissioned by the Vocational Education Unit of the State Chancellor's Office, the California Community Colleges, under the direction of Vikki Warner. We have been helped in preparing this report by meetings of the Vocational Education Research and Accountability Technical Advisory Committee (VERATAC) of the Chancellor's Office, chaired by Jack Friedlander of Santa Barbara City College; by a meeting with the board of the California Community College Association of Occupational Educators; by discussions with deans of occupational education in both the San Francisco area and the Los Angeles area; by meetings with various occupational educators from several states at the annual conference of the AACC; and by discussions with and comments from Ron Addy, Marilyn Jorgenson, Megan Juring, Larry Miller, Rick Moore, Jan Paulson, Harriett Robles, Rona Sherriff, Cheryl Stecher, Judy Walters, Vikki Warner, and Chuck Wisely.

PERFORMANCE MEASURES FOR IMPROVING
CALIFORNIA COMMUNITY COLLEGES:
ISSUES AND OPTIONS

Contents

Introduction: Our Principal Concerns 1

I. The Process of Establishing Performance Measures 7

II. Core Performance Measures and Standards for Occupational Education 12

III. Consistency with Other Performance Measures: 38
WIA, PBA, and PFE

IV. Other Issues Generated by Performance Measures: 58
Funding and Resources

V. A Summary of Options 65

References 70

PERFORMANCE MEASURES FOR IMPROVING CALIFORNIA COMMUNITY COLLEGES: ISSUES AND OPTIONS

Introduction: Our Principal Concerns

Accountability is surely here to stay. Community colleges in California are currently operating under multiple sets of performance measures. The direction — in this state and across the country — is to devise more performance measures, to use these measures in ways that have more important consequences for local providers of education and training, including — in some states and some programs — using performance to allocate funding. There's every reason to think that the current movement for accountability and the requirements for performance measures in both federal and state legislation will be permanent, rather than temporary reforms that can be safely ignored.

In this paper, we are concerned first and foremost with the accountability system and performance measures required by the Carl Perkins Vocational and Applied Technology Education Amendments of 1998, which provides federal support for occupational education. This is the first set of federal performance measures that California community colleges must meet, and this is a principal responsibility of the Vocational Education Unit that has commissioned this paper. However, we have also used this paper to explore various other systems of accountability, both federal and state, because the issues surrounding accountability are increasingly complex.

In devising performance measures to meet various federal and state requirements, we have a number of concerns that will influence our discussion and the alternatives we present:

- Performance measures should first and foremost be mechanisms that enable local program to improve their quality. If this cannot happen — if, for example, state-required measures do not provide the right kind of information to local programs to enable them to see what is going wrong, or to see how they can improve, or if there are no resources available for improvement or if the state establishes state-level performance measures without any influence on local programs — then performance measures will become ways of judging programs but without effects at the local level. In particular, we will often be concerned with the value of performance measures in improving the quality of the lowest-performing programs, on the assumption that this is where the greatest improvements can be made.

- Performance measures should be devised to minimize unwanted or unforeseen consequences. The history of performance standards under the Job Training Partnership Act (JTPA) — the forerunner of the Workforce Investment Act (WIA) — is full of unintended effects. For example, the requirement of meeting placement standards led many programs to "cream", or to accept only the most able applicants; the cost per placement standards, intended to reduce the costs of training, caused many programs to offer short, cheap programs with poor long-run consequences for employment. We will therefore try to identify the potential negative effects of various performance measures, and identify alternatives that might minimize such unintended effects.

- The process of setting performance measures should allow for improvement

over time. It's unreasonable to think that California can develop the "right" set of performance measures the first time around, particularly for institutions like community colleges that are so complex and multi-purpose. We think it's inevitable that the first measures developed will prove to have some flaws, some unintended consequences, and some gaps in their effects. However, if the process of setting performance measures allows for enough feedback from local colleges to the state, and for modification of initial performance measures, then over time this process should develop better measures — that is, indicators that measure performance in more reliable and valid ways. (Indeed, the Perkins Amendments requires a wide consultation process for the development of state plans, and this consultation process is one — though only one — such mechanism.) For the same reasons, throughout this paper we provide some options for long-run improvements in the development of performance measures and underlying data systems — since the data currently available and the performance measures initially proposed may not be adequate to the complexity of the underlying problems.

Alternatively, if the state does not create procedures within this kind of consultation and improvement, we fear that performance measures will come to be seen as illegitimate — and then local providers of education and training will comply with the letter but not the spirit of performance measures, or will try to fudge or cook the numbers they provide the state (as JTPA providers sometimes have). This is not a climate in which the improvement of local programs is possible.

- There should be some consistency in the performance measures required of local colleges — both consistency over time, so that colleges are not surprised by state requirements, and consistency across programs. In this paper we are particularly concerned with consistency across multiple sets of performance measures to which community colleges must respond, since it is possible that colleges could be subject to inconsistent measures. They include at a minimum the Carl Perkins Vocational and Applied Technology Education Amendments of 1998 (or VTEA, also called Perkins III by some), providing federal support for occupational education; the Workforce Investment Act of 1998 (or WIA), providing federal support for various workforce development efforts including some job training; the Partnership for Excellence (PFE), providing amounts ranging from \$100 million per year in 1998-99 in exchange for movement toward performance targets; and the state report card (SB645), which specifies a number of performance measures which several kinds of programs must report. There are of course other federal and state regulations and reporting requirements that affect local programs, but these four are the most important in developing specific performance measures and targets that California community colleges should meet.

- The process of setting performance measures should ideally be stable and predictable, so that local providers of education and training can plan for changes that are coming. In examining other states' efforts to create coherent workforce development systems, we have been impressed by the importance of stability: states that have stayed with the same policies over many years have often made substantial progress in making their workforce development systems more coherent, while those that have changed course — often because of changes in political leadership and direction, or changes in philosophy — have made almost no progress (Grubb et al., 1999). So too in the arena of accountability and performance measures, we suspect that it will be critical to establish performance measures, create a plan for their improvement over the long run, and stick to

this plan, rather than modifying the state's plan frequently or changing the overall direction of the state's approach.

- It's important to clarify the differences between state performance measures, and local performance measures which may be based on state measures. Federal legislation requiring performance measures typically requires the state to come up with a series of state measures — that is, measures reflecting how the state as a whole is doing, for purposes of comparison with other states or for examining a state's overall performance over time. In addition, federal legislation — both VTEA and WIA — specify that local measures be developed, applying to local institutions and corresponding to state measures. The underlying logic, evidently, is that if the state does not impose local performance measures on local programs, there is no way to improve the state measures that it must report to the federal government — and the system of performance measures becomes simply a mechanism of reporting, rather than one for improving the quality of programs.

Federal legislation also specifies the consequences for local programs of failing to meet local standards. In VTEA, for example, local programs that fail to meet local standards are required to develop mechanisms to enhance their performance in subsequent periods, in developing local improvement plans (Section 123(c)). However, the state always has the option of taking more drastic action than the federal legislation requires; and in cases like the Partnership for Excellence (PFE) and Performance-Based Accountability (PBA) where there are currently no local measures or standards, the state has the choice of various ways of encouraging or requiring local programs to meet performance standards, ranging from doing nothing and hoping that programs will voluntarily improve, to reporting local performance and hoping that the publicity forces them to improve, to reviewing local programs that fall below acceptable levels, to performance-based funding, to incentives for meeting or exceeding goals, to Draconian measures closing all programs that fail to meet certain standards. We review some of these options for occupational education in Section II, and for programs funded by WIA in Section III. Throughout, therefore, we will be concerned with the distinction between state performance measures and local indicators of performance, since they serve different purposes.

In this report, we first discuss the process of setting performance measures, in Section I. Our concerns in this process extend both to the communication between state government and local colleges, and to the communication across the four different programs that now set performance measures. Then in Section II we examine the core performance measures required under VTEA, examining those that the state has already devised and suggesting some options for both the short run and the long run.

In Section III we examine the performance required in other state and federal legislation. Our concern is principally with the consistency among these measures, as well as with the value of these alternative programs in stimulating local improvement.

In Section IV, we examine other issues generated by the development of performance measures, particularly cases where these measures are at odds with the purposes and incentives currently in state legislation. We also address the resource requirements for developing effective performance measures. Indeed, often there is a lack of resources for the analysis required by basic policy decisions. These resource problems will arise throughout the report, and we will clarify the implications both for state-level

expenditures and for resources available to local colleges. Our point is that, while developing performance measures can be done on the cheap, an effective approach to measuring performance may require additional resources for further analysis of existing data, for the development of additional sources of data, for methods of translating performance measures to the local level, and for methods of responding to this information.

I. The Process of Establishing Performance Measures

The process of establishing performance measures in California needs to provide enough consultation between state officials making decisions and local colleges so that the limitations of any one set of performance measures can be understood, and alternatives devised. Then the state's performance measures can, over time, develop into increasingly reliable and valid ways of measuring performance, increasingly useful to local programs in helping them diagnose and correct various problems. We stress, then, that while performance measures are sometimes required by the federal government, for federal purposes, the state of California and local colleges should always regard them as potential mechanisms for improving the performance of local programs — not merely as pesky federal requirements to be subverted whenever possible.

At the same time, a similar relationship exists between states like California and the federal government, with the same possibilities for viewing accountability either as a mechanisms of improvement or as a requirement to be evaded. If a particular federal requirement is not helpful in improving the quality of local programs, then the state must be ready to argue to federal officials why such a measure is inappropriate — or why its development must be postponed. In our interpretation, federal policy-makers are more interested in seeing states take performance seriously and move toward improved measures of outcomes, rather than in enforcing federal requirements in narrow and ultimately counter-productive ways.

The process of establishing the first set of performance measures required by VTEA was necessarily rushed, since California was required to submit a preliminary state plan by April 15, 1999. The initial submission is discussed more fully in Section II. It was developed by the Vocational Education Unit of the State Chancellor's Office, with the consultation of the Vocational Education Research and Accountability Technical Advisory Committee (VERATAC), headed by Jack Friedlander of Santa Barbara City College, with several institutional researchers and deans of occupational education. While this committee cannot represent all 107 colleges in the state, it is reasonably representative of those who are knowledgeable about existing data and about the occupational programs that will be subject to these performance measures. In the interests of consistency over time, the Chancellor's Office may want to continue using VERATAC to further develop and refine performance measures since it is a relatively small group, well-versed in both the issues of data and of practice. We suspect that it is impossible to make such changes in a large and politically-representative group — like the consultation process mentioned in VTEA (Section 122(b)(1), which includes "teachers, eligible recipients, parents, students, interested community members, representatives of special populations, representatives of business and industry, and representatives of labor organizations" — since these groups will not be familiar with the technical issues of data

systems and indicator construction. Such groups are still valuable for information about potential effects, inaccuracies, and the like.

To expand input to VERATAC, the Chancellor's Office should continue a broad process of consultation about the further development of its performance measures. (This paper is, of course, a small part of this process.) In particular, representatives could meet with institutional researchers of the colleges, at their annual meeting in April, for the purpose of both explaining existing measures and hearing suggestions for their modification, including data problems experienced on local campuses. Similar meetings should continue to take place with the deans of occupational education, again at annual as well as regional meetings, for the same purposes. In addition, consultation with Chief Instructional Officers and with presidents and chancellors, at their annual meetings, would probably help reinforce the importance of performance measures in general and would provide yet additional information for improving these measures. If this broader process of consultation is followed, there remains the problem of coordinating information. One alternative is to formalize the existing informal process, with representatives from each group nominated to serve on VERATAC, and to be the mechanisms of communication between the state and the local levels. While this kind of representation now takes place informally — since VERATAC has members who include deans of occupational education and institutional researchers — a formalization of this process might be warranted (see Option 1). This minor change would serve the interests of consistency and stability that we have highlighted above (see p. 3).

A different dimension of the process of establishing performance measures involves providing information to local colleges so that they can improve local programs. A number of local deans have complained that the data now provided by the state is not particularly user-friendly. The information comes in large tables of data, with no interpretation of what they mean, and especially for small programs the results are often either missing or meaningless. Statewide figures are provided for comparative purposes, but colleges are usually unable to determine whether they are performing well or badly, compared to other similar colleges. Some colleges with institutional researchers may be able to translate these results to deans and instructors, but colleges without institutional researchers and those whose institutional researchers are otherwise occupied — since institutional researchers tend to be scattered among many different assignments — report that they are unable to make good use of these data. One option, then, would be for the state to develop a pilot project with a small number of colleges — both those with strong offices of institutional research, and those without institutional researchers — in order to translate state data into formats that occupational deans and faculty can more readily understand, and to develop appropriate comparisons so that local colleges can see how they are doing relative to others (see option 2). This step is crucial to these data being understood at the local level, and therefore in being used to make decisions leading to improvement.

Another complaint from local deans and institutional researchers involves the timeliness of data reported to them. They have cited instances where results from the early 1990s, about programs that have since been extensively modified, have been reported as indicators of current performance. We are unclear precisely what the origin of this problem has been, though we note that it is necessary for the data on performance measures to be as current as possible. However, in many cases a lag is unavoidable. For

example, if a performance measure requires information about the employment of those leaving community colleges, it's necessary to wait a year to identify individuals who have left rather than re-enrolling. The availability of other data, particularly financial aid data used to identify economically disadvantaged students, creates its own delay. And when we propose shifting to even longer-run outcomes (like earnings three years after leaving college) or using truly longitudinal measures, the lags become even longer. These unavoidable delays mean that state policies and local program improvements should be based on stable pattern in performance measures, not fluctuations from year to year.

In addition to the process of establishing the performance measures required by VTEA, there are other mechanisms to establish the performance measures required by WIA, by the state report card (SB645) or Performance-Based Accounting (PBA), and by the Partnership for Excellence. Currently, the PBA measures are established by the PBA Committee, with members from a variety of education and training programs (including community colleges), with technical work provided under a contract with AMPG. The PFE measures are specified by a PFE committee within the State Chancellor's Office, using the state's data base on community colleges. And the process of establishing WIA performance measures is currently in a state of great uncertainty and high anxiety, because legislation establishing a State Workforce Investment Board has not yet passed and therefore there is no entity that can start to think about what these measures should be — and, as we clarify in Section III, the Department of Labor may specify in great detail how these performance measures are to be calculated.

We cannot say how well these processes work. We suspect that there is some friction in some of them, like the PBA Committee with its representatives from very different programs; we would anticipate similar friction in committees created to implement WIA. (Of course, disputes among different programs can be healthy as well as destructive.) Our point for the moment is that the processes for establishing performance measures in California are formally independent of one another. In practice, several individuals (including technical staff) sit on more than one of these committees, and they create informal linkages and channels of communication among the different accountability systems. But the potential of inconsistency among accountability systems is still present, particularly when the goals of various systems are somewhat different.

II. Core Performance Measures and Standards for Occupational Education

The Carl Perkins Vocational and Applied Technology Education Amendments (VTEA) require four core indicators of performance:

1. Student attainment of academic, vocational, and technical proficiencies;
2. Student completion of credentials.
3. Placement in employment, further education, or military service.
4. Participation in non-traditional occupational programs.

States may also develop other indicators of performance, and a few (e.g., South Carolina with 39 measures in its Performance Indicator Reporting System) have done so.

As we stressed in the introduction, VTEA requires state measures of performance, and they are to be used in two distinct ways: They form the basis for agreements between the Secretary of Education and state agencies, and presumably the Secretary of Education can cut off funds to states that fail to meet their agreed-upon performance targets —

though typically this ends up being a political rather than a technical issue. In addition, WIA includes incentive grants that are allocated to states on the basis of their performance measures in occupational education (both secondary and postsecondary), training programs supported by WIA, and adult education. Thus California community colleges contribute toward overall performance measures for WIA incentive grants, though the other components are not under the control of community colleges at all.

Because there are some consequences for California — however small and remote — of poor performance on these state indicators, we will be concerned in discussing the proposed measures with the potential changes in these measures over time, particularly as populations of students change and as economic conditions vary.

However, given that the possible consequences for federal funding of poor performance are still relatively remote under VTEA, it may be best for California to be motivated by its own sense of how best to improve local programs, using federally-required performance measures as a way to accomplish this. In this spirit we will be concerned with how best to develop a system of local indicators that the state can use to improve the performance of local programs — and that will, along the way, also serve to improve the state indicators.

This is the first place to confront the issue of creating local indicators and then developing incentives for local programs to improve. In subsequent sections, we will specify potential local indicators. However, there are currently no consequences for local colleges of performing poorly on any measures. The state has a number of options, for this and all other measures required by VTEA, WIA, and state accountability measures, though some of them would require legislation and others would require the state-local relationship to change substantially. The alternatives are myriad; a few examples include the following:

a. The state could provide information to every college and program about their performance, leaving it up to the local college to improve performance.

b. The state could require every college, or every program, falling below local performance indicators to provide the Chancellor's Office with both a diagnosis of the problem, and a series of steps to be taken to bring the college or program up to the state standard.

c. The state could subject every college or program falling below local standards to a state review and audit, with state officials developing the diagnoses and remedies. This would require increasing the staffing within the Chancellor's Office, and probably the creation of a unit with responsibility for these local audits.

d. The state could specify a period of one (or two, or three) years for a college or program falling below minimum standards to meet those standards, where all programs failing to do so within this period would cease to receive state funding. (For example, for some years Florida has required all occupational programs to achieve 70 percent related placement, and those failing to do so are terminated.)

e. As part of any state review or performance requirement, the state could also provide technical assistance to local programs, providing information about effective practices in other community colleges both inside and outside of California. Of course, this activity requires resources, as we recapitulate in Section IV.

f. The state could institute performance-based funding using the various local performance indicators to set funding levels. We note that, while performance-based

funding creates substantial incentives to improve, by paying bonuses to the best-performing colleges, it also reduces resources to programs with low indicators — and these may be precisely the programs most in need of resources. (Alternative d is also a form of performance-based funding.)

The decision among these and other possible local incentives is a political one, and should be discussed in VERATAC and other councils responsible for performance measures and accountability. We note, however, that purely voluntary goals have not worked well in one state that has adopted them, Oklahoma (Grubb et al., 1999). At the other extreme, the two performance-based funding alternatives may strike many Californians as somewhat premature: California may want to adopt performance-based funding at some point, but it should probably be evaluated as a long-run direction of state policy, rather than an immediate response to federal policy. This leaves alternatives (b) and (c), of which (b) is relatively feasible while (c) would require a very different relationship between the Chancellor's Office and local colleges — and many more state-level staff, competent to work closely with local colleges — than is now the case. This process of elimination suggests that requiring an improvement plan from local colleges and programs may be the best alternative, at least in the near term. This option is also consistent with the requirement in VTEA of requiring local improvement plans for colleges that fail to meet state performance standards.

We note that this option — generating local improvement plans — also provides the Chancellor's Office with information about what local colleges are doing to improve. This information, plus information about what other successful colleges are doing, could be the basis for an annual report on promising practices so that all colleges could learn from their peers, and for the kind of technical assistance outlined in (e) above. This would involve the Chancellor's Office in providing a kind of technical assistance to improve local programs that is complementary to the role of performance measures in creating incentives to improve. (See Option 3.)

Whatever decision the Chancellor's Office makes about the process of local improvement, the principal of consistency and predictability suggests that the same procedure be adopted for each of the core measures required by VTEA — and, by extension, other federally-mandated performance measures like those in WIA, and by state-required accountability systems like PBA and PFE.

Core Measure 1

The first core measure required by VTEA must measure the acquisition of academic, vocational, and technical skill proficiencies. The Chancellor's Office has proposed using the percent of students in vocational-technical courses with a grade of C or better (from the 1999-2000 California State Transition Plan, 3/11/99) as the measure of proficiencies attained. Under the assumption that course content meets state standards as specified in Title 5, course grades are appropriate measures of competencies acquired as measured by the tests, papers, projects, and other activities devised by local instructors.

The state's earlier submission to the federal Department of Education did not specify a state level of performance. These levels of performance must be developed by each state and, with the agreement of the Secretary of Education, they apply to the first two years of VTEA (1999-2000, and 2000-2001). VERATAC has suggested starting with a standard of 70 percent of students completing vocational-technical courses — defined

as SAM (Student Accountability Model) level A - C courses, which are respectively apprenticeship courses, advanced occupational courses, and clearly occupational courses — with grades of C or better. Seventy percent is an admittedly arbitrary figure but one that represents a preponderance of students — more than two thirds — attaining the competencies embedded in standard grading. In the 1997-98 data, 75.0% percent of all students in these courses received a C or better, indicating that the state would be above its own standard in the first year of VTEA. However, the 70 percent figure will, if embedded in local performance standards, provide strong incentives for local programs to improve the rates at which students receive grades of C or better, since 39 out of 171 program areas failed to meet this 70% standard in 1997-98.

An alternative measure would set the state standard according to the proportion of local programs where the percent of students with a C or better is 70 percent or above. This alternative measure would clarify that the state's principal concern is with the programs showing low levels of performance, where large fractions of students fail to master the material in SAM level A-C courses. In identifying the programs that are substandard, such a measure would also recognize that there are ceiling effects — that is, programs that are performing well may not be able to increase their performance levels. It is, for example unreasonable to expect local programs to have 100 percent of students receiving a C or better; this could happen only with grade inflation allowing every student, no matter how lackadaisical in attendance or performance, to receive a C, or through "creaming" to select only the best students — neither of which is a desirable outcome. In 1997-98 data, 132 programs out of 171 (or 77 percent) had successful completion of 70% or more. One possibility would therefore be to start in 1999-2000 with a state standard of 77 percent of programs meeting the standard of 70 percent C's or better. (See Option 4.) Improvement on this state standard would then require local programs to improve during 1999-2000 and 2000-2001.

If the state is to develop local performance indicators, the proportion of students with C's or better in SAM level A-C courses is a clear local performance measure, one that can be reported both for colleges and for occupational programs. The appropriate standard would be taken from the state standard; for example, the standards would be to have at least 70% of all students receiving C's or better. Then the Chancellor's Office could take any of the measures described above (alternatives a - e, or some other variant) using this measure of performance.

Note that this approach contains incentives for low-performing programs to improve, but it does not include any incentives for high-performing programs — e.g., those with more than 70 percent of students already receiving C's — to maintain their high standards. Under this approach to local performance indicators, one option would be to require that programs above the 70 percent figure maintain or improve their performance; but programs whose performance fell would be identified and then subject to any of the state's improvement mechanisms set out in (a) - (f).

Changes in State Performance Over Time: An issue to consider is how the state performance measure is likely to change over time, even in the absence of local program improvement. We note that this completion measure has been relatively stable over time: between 1992-93 and 1997-97 it varied between 74.2% and 75.4% statewide, variation that is quite trivial. However, two trends seem possible. One is that the proportion of underprepared students enrolling in community colleges is likely to continue increasing,

because of immigration and the lack of much progress in improving high schools — potentially depressing the proportion of students with C's or better. One way to handle this possibility is to compute the performance measure only for those students who have acceptable scores on diagnostic exams given at matriculation, or who have completed the appropriate remedial/ developmental courses. Currently, the data necessary to compute such a measure is not available at the state level, since colleges use different exams as students first enter colleges, do not report scores to the Chancellor's Office, and follow different practices about which students take diagnostic exams. While the state knows which students have enrolled in various remedial/developmental courses, it does not know which students should have enrolled in such courses (based on diagnostic exams) but did not. The Chancellor's Office might consider including such data into its system, both for this purpose and to monitor the performance of low-achieving students in general. If the state decides to take this approach, then it will need to confront the variation in the diagnostic tests and standards used from college to college. While it is possible to use unstandardized estimates of the proportion of students needing remediation in state data systems,

it would be better from the perspective of uniformity to shift to a statewide diagnostic process — though this would require agreement among 107 colleges on what process to use. (See Option 5.)

In addition, California (like the rest of the country) is now in a boom period; when a recession occurs — as it surely must sometime — then enrollments are likely to increase with individuals unable to find employment (Betts and McFarland, 1992). We suspect that students entering during a recession are likely to be more able — since currently there is likely to be a tendency for the least able students, who cannot find employment even in boom times, to be enrolled — though we aren't sure whether the magnitude of this effect is substantial or not. For the moment, it might be sufficient for VERATAC to be aware of this possible change, but to postpone any additional analysis until a recession starts to emerge.

However, in considering local performance measures, there are surely differences among local colleges in the preparation levels of students, from suburban colleges with middle-class populations to urban colleges with high proportions of students from low-quality high schools, of immigrants from countries with mediocre educational systems, and of low-income students. This suggests that the local performance standards should, over the long run, be adjusted to reflect the preparedness and composition of students in different colleges. This could be done, for example, with regression adjustments similar to those that have been used for JTPA, with independent variables including the preparedness of students (measured by diagnostic tests at matriculation, for example, or high school grades), language or immigrant background, and income measures like eligibility for student grants and loans. Such a study would be difficult to carry out at the moment, because some of these data are not available or are not uniformly collected across the state; therefore it may be necessary for the Chancellor's Office to engage in some pilot studies to examine the value and the feasibility of such regression adjustments (Option 6)..

Unintended Effects: There are two potential unintended effects from this indicator. One is that it provides incentives for "creaming", or accepting the most-able students and discouraging the least-able students from enrolling in occupational courses.

Even though community colleges are open-access institutions, there are many ways for colleges to change the mix of students, from aggressive counseling procedures to selective information and recruitment to informal "aptitude" tests. One way to detect this is to examine the composition of students entering various colleges and programs, and to see whether there are any dramatic changes in student composition over time.

A second unintended effect is that measuring competencies according to grades provides an incentive for grade inflation. Of course, grade inflation has long been part of American education, and there are many other incentives to inflate grades — ranging from federal grants and loans, to enrollment pressures, to instructors' sympathy with students, to the constant badgering of students themselves. Furthermore, there are relatively few antidotes to grade inflation. One of them — developing external standards for various occupational areas — is discussed below but requires long-run development. Others — particularly a consensus on standards among instructors and the internalization of standards in individual instructors and their approaches to teaching — require difficult improvements in the quality of teaching (Grubb and Associates, 1999). For the moment it may be sufficient for the Chancellor's Office to be aware of the possibility, and to be alert for any signs of suspiciously high rates of improvement in grades.

The incentives for creaming and grade inflation almost surely increase if local performance measures are set unreasonably high — that is, if it appears that there are no ways to improve performance except by "cheating". With this in mind, the state might therefore want consistently to set performance measures so that they can be reasonably attained through program improvement, rather than setting levels so much above current levels of performance that they seem impossible to attain. This is another reason to develop regression adjustments to reflect the differences among colleges in the preparation of students.

Long-run Developments: VTEA requires measuring academic as well as occupational and technical competencies. In subsequent revisions of performance measures, the state might consider grades on SAM level E courses — which are academic courses — to supplement information on SAM level A-C courses. It would be particularly important to use information on academic courses related to a student's area of study (Option 7).

Over the long run, the state might consider other ways to measure the "challenging state-established academic, and vocational and technical, skill proficiencies" called for in VTEA. Grades are, after all, measures of proficiency only if instructors have incorporated the relevant proficiencies in their courses, and if their grading policies reflect student attainment of these proficiencies — but they are not direct measures of the proficiencies themselves. On the other hand, the direct measure of proficiencies that vary among occupational areas — and also, in all likelihood, among regions in California as labor markets and required skills vary — is a complex task.

One set of proficiency measures currently available is embedded in licensing requirements and tests. The Chancellor's Office and VERATAC might therefore consider incorporating information on the rates of passing licensing exams in different programs — though first it would be necessary to know which occupational areas are covered by state-required licensing provisions. Other proficiency standards will be more difficult to measure, though they might be the basis for pilot projects. For example, many industry associations (like the National Automotive Technicians Educational Foundation and the

American Society of Welders) have voluntary standards, and some firms (like Novell, Microsoft, and Toyota) have other certification procedures applicable to their own products. For purposes of Core Measure 1, such "private" credentials are again difficult to incorporate because applying for such credentials is now voluntary and therefore uneven among programs, and because they apply to only some occupational areas. Still, there are other reasons to explore such "private" credentials, as we explain in conjunction with Core Measure 2, and so a pilot project might be appropriate.

The only way we know of to develop complete and external measures of proficiencies would be for every occupational area within community colleges to convene, agree about appropriate standards, and develop the assessments to measure such standards. For occupational areas, these would surely include new kinds of performance-based and other "authentic" assessments, not simply pen-and-paper tests of knowledge. One possibility would be to explore the development of such standards in occupational areas without licensing requirements; then proficiencies statewide could be measured through a combination of licensing results and state-set standards. Such an initiative would be parallel to those that have taken place in K-12 education, with the various curriculum frameworks and curriculum projects. This effort would be time-consuming, but it would place California in the forefront of efforts to develop meaningful and comprehensive skill standards.

Core Measure 2

The second measure required by VTEA must reflect the rate of acquiring postsecondary degrees or credentials — a particular conception of completion. The Chancellor's Office has proposed using attainment of community college certificates or Associate degrees as the state measure of performance. VERATAC has proposed somewhat more detailed measures for this indicator, measuring completion as the number of students receiving an Associate degree, certificate, passing a licensing exam, or transferring to four-year colleges within a particular year divided by the number of occupational education majors (defined as those who have completed at least one SAM A through C course and who have accumulated 12 or more vocational units in one 2-digit TOP code within the previous 5 years) within the same year. Roughly, then, this measure represents the proportion of occupational majors who complete credentials or licenses, or who transfer.

One alternative for VERATAC to consider is changing the state measure from the proportion of students completing to the proportion of programs meeting some completion standard. The latter measure is more consistent with efforts to improve the quality of particular programs, while a measure based on individual students will be affected as students move into and out of programs with high and low completion rates — for example, into and out of nursing programs. If the aim of the state is to improve low-performing programs, then a state measure responsive to this goal is more appropriate.

VERATAC has also recommended that the state standard be set at 60 percent completion; while completion rates have been slightly higher in recent years, (63% for the 1993-94 cohort and 69% for the 1995-96 cohort), a standard set slightly low will still provide the necessary incentive for low-placement programs to improve.

In terms of devising local performance indicators, the obvious measure parallel to

the state measure would be the proportion of vocational education majors completing certificates, Associate degrees, and licensing exams, measured for both specific occupational areas and for colleges as a whole. Then this local measure would be used to trigger any of the state actions outlined above (in a - f).

Changes in State Performance: There are likely to be several effects on this measure of performance over time, some increasing and some decreasing this measure of completion rates even if the quality of programs does not change. If students are becoming less well-prepared, as many in California think, then either completion rates may fall (affecting Core Measure 1) or placement rates may fall, affecting this core measure. As with Core Measure 1, one solution to this problem is to measure completion and then placement separately for those passing diagnostic exams or all necessary remedial/developmental courses, and those who are not prepared in this sense; such an approach would require uniform information about preparedness measured in this way.

In addition, completion rates are surely affected by the overall economy. (This is a serious issue addressed in a subsequent section on long-run improvements in this measure of completion.) In boom times, students appear to leave programs when they find appropriate employment; indeed, many local administrators claim that their best students leave and that some employers "raid" their best students after the first year, reducing completion rates and skewing the composition of completers. In recessions, conversely, we might expect completion to increase. In theory it is possible to determine how sensitive completion has been to the overall state employment rate, though we doubt that data are available for a long enough period to clarify this pattern. As with cyclical effects on grades, one alternative would be to wait until this danger is more clearly present.

Regional Variation: We suspect it is likely that completion rates vary substantially among colleges, particularly as the preparedness of students vary, as the amount of remedial/developmental education they require varies, and as their goals and intentions vary. One possibility, therefore, would be to analyze the potential effects of regression adjustments for measures of completion. This would require exploratory studies of the effects of student characteristics on completion rates. It's also possible that the state of local labor markets influence completion — for example, with completion lower in strong and fast-growing labor markets — and this too should be considered in any exploratory analyses. The state might therefore include in its state plans the rationale and methodology for carrying out such regression adjustments (Option 6).

Long-run Developments: The measure of completion proposed by VERATAC is a conventional measure, given the data available. However, there are several problems with this measure, and improving the measure of completion might be a goal over the next several years.

The first and most serious issue — indeed, the single most frequent complaint among occupational deans and faculty about any issues in occupational education — is that completion of certificate and Associate degrees fails to incorporate completion as measured in other ways. Adding information about licenses received helps to some extent, but other kinds of completion not measured in conventional certificates and Associate degrees include:

- Students who earn state-required licenses, or "private" credentials, or industry-generated credentials, from industry associations and individual firms, and then leave for

employment.

- Students who intend to complete only a few courses necessary for advancement, and then leave the college with their goals completed.

Collecting information about these kinds of completion will require information from students themselves, for example collected in exit interviews, or (for those who intend to complete only a few courses) collected through questionnaires upon entrance. All colleges already collect data on student goals when they enter, and then collect information about "informed" goals at every registration — though there is surely some variation in what this information means. The Chancellor's Office could use the questionnaires already developed as a guide. (In addition, the Chancellor's Office used to carry out student follow-up studies, conducted first by Walter Brooks and then Nick Atma, and these earlier efforts may provide additional guidance.) For purposes of developing state measures of performance related to completion, it is necessary to have data that are consistent across the state — and so it would be necessary to develop a uniform system of student questionnaires upon entrance into community colleges and upon exit. While information on student goals is now collected and included in the state's data base, colleges update student goals at different times, leading to uncertainty about the point in a program that the data refer to. An alternative for the Chancellor's Office to consider, therefore, is the development of such supplemental information about other forms of completion. (See Option 8.) Under this procedure completers would include (1) those who earn certificate and Associate degrees; (2) those who earn licenses, even if they fail to complete certificates and Associates degrees; (3) those who report earning an industry-generated credential; (4) those who reported upon entry that their goals were simply to upgrade their skills, and who reported upon exit that they had successfully completed coursework they were seeking, or who had completed a certain number of courses. The results of such forms of completion should then be reported by age or experience of students, to test the common statement that older students are more likely to enroll for skills upgrading and leave after completing a few courses.

We note that information about completion rates is necessary not only for VTEA Core Measure 2, but also for the identification of eligible providers under WIA (see Section III below) — and so a more accurate and expanded conception of completion will be important for more than just VTEA.

Second, the measure proposed by VERATAC is not a truly longitudinal measure, and therefore does not accurately measure the proportion of students entering colleges who subsequently complete (in any sense of completion). It is instead a ratio of two numbers collected at the same time from different groups of students, and only approximates the correct longitudinal measure. The differences between true longitudinal measures, and those calculated from ratios two independent numbers are particularly large for colleges with increasing or declining enrollments, and the results in particular may understate completion rates for colleges with increasing enrollments (for which the denominator of the measure will be increasing faster than the numerator). The state's system of data is truly longitudinal, since it has data on students and their enrollment in every semester. It is therefore possible to calculate completion rates that are truly longitudinal, and the Chancellor's Office may want to start using truly longitudinal measures of completion. (See Option 9.) This would result in figures like the proportion of students entering in fall 1996 who were occupational majors and who then completed a

certificate, Associate degree, passed a licensing exam, or transferred by fall 1999 (or who completed in other ways). Typically, as the length of time over which completion is measure is lengthened, completion goes up but by diminishing amounts — and it may be especially important to use longer periods of time for various groups of disadvantaged students.

To be sure, the use of truly longitudinal data creates certain problems of timeliness. For example, a measure of completion three years after initial enrollment must examine a cohort that enrolled at least three years ago, or more likely four years ago given the time requirements for collecting data. However, what is important is not having results for the current cohort — since their completion cannot be known for several years to come — but rather understanding the patterns by colleges, by programs, and by local economic conditions, and for these purposes longitudinal measures are more appropriate and accurate than the current measures.

Finally, there are several detailed problems with the quality of data provided to the Chancellor's Office. For example, some colleges apparently award certificates and degrees to students automatically when they complete a certain number of units, while others require students to apply for the same credentials — and therefore have lower apparent completion rates. Where there are such differences in local practices, the Chancellor's Office might consider requiring uniform practice — especially now that there will be consequences to performance measures and consistency has become important in ways that were not true before. (See Option 10.) In addition, some colleges have begun to define early exit options, or exit points short of a certificate, particularly in programs for welfare recipients. If such options are to be considered as completion for purposes of Core Measure 2, then again some standardization of practice across colleges will be necessary.

Core Measure 3

The third core measure required by VTEA must reflect the rate of placement and retention in advanced training, employment, or military service. The Chancellor's Office has proposed measuring placement and retention by employment or transfer to other postsecondary education. More specifically, VERATAC has proposed using the state's system of Unemployment Insurance (UI) wage record data to develop two separate measures. Placement would be measured by the proportion of a cohort (excluding those who transferred to four-year colleges) who found employment as measured by UI records, and would be calculated separately for completers and for non-completers. Retention would be measured by the proportion of those found in UI wage records who had three or more consecutive quarters of employment within the first five quarters after college. VERATAC did not establish a standard for placement because such rates are seriously affected by state (and local) economic conditions (like the unemployment rate). In addition, some individuals are not included in the state's UI data — those who are self-employed, those in federal and state government positions, those in the military, and those who have moved out of state. Therefore some occupational areas — particularly in occupations with many self-employed individuals, such as real estate and cosmetology, and those who enter government as in police and fire programs — may show low rates of "employment" when the real problem is that UI data does not include many of their graduates.

For job retention, VERATAC recommended that the state measure be the statewide average for the 1997-98 cohort of 72 percent; establishing such a standard would therefore provide pressure and incentive for local programs to increase their retention over the next several years.

Given the difficulties of measuring placement with other sources of data — for example, with follow-up questionnaires to students, which typically have very low response rates — the UI data is by far the best source of information on placement available. However, the uneven coverage of the UI data is certainly a problem, and therefore the Chancellor's Office might examine statewide variation in placement as measured by UI data, to see whether the expected differences associated with self-employment, and other forms of non-coverage, are in fact problems. Second, we expect that placement should vary with local economic conditions, as well as with the personal characteristics of students, and the Chancellor's Office might analyze the placement figures for such variation. Only when there is better understanding of precisely how such placement measures behave should there be an effort to establish either a state standard, or a local standards for individual programs to meet. Indeed, we anticipate that any local standard for placement will have to have some adjustments, at least for local economic conditions and probably for student characteristics as well.

The state's measure of retention could take several different forms. The current proposal — the proportion of individuals employed (i.e., located in UI data) who are employed two consecutive quarters or more — describes a statewide average. However, an alternative — especially if the state wants primarily to improve the performance of the lowest-performing programs — is to shift to the proportion of local programs that meet the 72 percent state average. Then as local programs improve their performance, state performance will improve as well. The local indicator appropriate to retention is then the proportion of those employed (i.e., located in UI data) with two consecutive quarters of employment, and the state could use any of the approaches outlined in (a) - (f) above to encourage local programs to improve their retention.

Job retention as measured by two or more consecutive quarters of employment is also a performance measure that is almost surely influenced by local economic conditions. In regions with more small businesses with high turnover rates, in labor markets with higher unemployment rates or affected by seasonal patterns (like construction, agriculture, and timber, for example) we might expect lower levels of retention unrelated to the quality of programs. Therefore the Chancellor's Office might analyze the variation in retention rates across the state, and over the long run local indicators of retention might need to be adjusted for differences in local economic conditions.

Changes in State Performance: Both placement and retention are likely to be influenced by statewide economic conditions, including the overall employment rate, the mix of sectors and occupations, and trends in employment practices (like the practice of using large proportions of contingent or temporary workers). While this is not an issue for the moment, future calculations of placement and retention should bear these factors in mind.

Unintended Effects: The most serious side effect of using placement rates to measure the performance of local colleges is that it provides incentives for programs to "cream", or enroll only the most able students. Conversely, women in non-traditional

fields, minorities, recent immigrants and non-native speakers of English, and disabled students are all likely to experience lower placement rates, and therefore local programs have incentives not to accept them. Such consequences would be contrary to the community colleges' role as the "people's college" and an open-access institution, of course. The separate calculation of placement for completers and non-completers is one way to mitigate the incentive to cream somewhat; in addition, if placement proves to be highly sensitive to the composition of students, then the use of adjustments in calculating local placement standards would also help colleges maintain high placement rates while still accepting a broad range of students. Finally, the Chancellor's Office should monitor the composition of students and look for any substantial changes in the composition of occupational students. (This is how the creaming in JTPA programs was detected, for example.)

Long-run Developments: One improvement over the long run would be to supplement the UI data with other information on placements, in order to gather information about individuals not covered by UI data. This would result in a "hybrid" data system, based on both UI data and questionnaire data from students. (See Option 11.)

In addition, an overall placement measure is probably not as appropriate a measure as a measure of related placement, or placement in occupations related to a student's field of study in college. Related placement is particularly important in occupational programs, many of which provide relatively occupation-specific preparation; if individuals do not find related employment, then the employment benefits of education may be low or even zero — as has been confirmed by some research with national data (Grubb, 1997 or 1999) and preliminary results with data from California, Texas, and Washington. However, the existing methods of inferring occupations, and therefore the relatedness of placement, from UI data are quite imprecise (Anderberg and Pfeiffer, 1998, Special Issues Addendum #1). Collecting such data would either require an addition to the data collected by the UI system, or would require gathering additional information directly from students — leading once again to a hybrid data system. The current pilot project undertaken by the Employment Development Department (EDD) to collect information directly from employers on the occupation of their employees covered by UI has yielded promising results, and appears to confirm that much higher employment benefits of finding related employment. It's important to recognize that there is likely to be bias in these results, since results are likely to be available from large firms only. The state of California might therefore consider extending this pilot project, and continuing to investigate alternative ways of measure the relatedness of employment. (See Option 12.) We note that information about related placement may in the future be required by WIA of those education and training providers that want to be eligible to provide training services (see Section III below), and so information about related placement may be necessary for this purpose as well.

One way to distinguish community college occupational programs from the very short-term programs provided by JTPA, adult education, or the ROC/ROPs is that community college programs can aspire to provide the broader and deeper preparation that prepares individuals for more permanent jobs, or that provide access to upward mobility over time, and for jobs with higher earnings. Some of the effects of more intensive programs are captured in the two-quarter employment measure, but not all of

them. Therefore the Chancellor's Office might experiment with alternative measures of the quality of employment as measured by longer-run measures of employment (over a period of time longer than a half year), upward mobility over time as measured by real earnings growth, and earnings themselves. (See Option 13.) For example, measures of earnings growth between the year before completion and three years after completion have been calculated, and are certainly better measures of long-run employment than are employment rates just after completion. In addition, the use of such longer-run measures is consistent with the measure required by Performance-Based accountability, which includes the employment rate after three years. However, before such measures of "quality" are established, either as state measures under VTEA or as local indicators of performance, there should probably be extensive analysis of existing patterns as well as consultation with local colleges.

Core Measure 4

The final core measure required by VTEA is the rate of non-traditional employment, where "non-traditional programs" is defined according to the balance of men and women; thus men in programs with more than 75% women and women in programs with more than 75% men would be considered to be in non-traditional programs. The Chancellor's Office has proposed using participation and completion in non-traditional employment as the two relevant state measures. The state measures would then be the proportion of men within the state enrolled in non-traditional programs, and the proportion of women enrolled in non-traditional programs. The rates should be calculated separately for men and women because the incentives are not the same: typically there are wage incentives for women to enroll in non-traditional programs, but there are often disincentives for men to enroll in occupations typically held by women, which tend to pay less than "male" occupations. Sometimes social conditions influence gender patterns; for example, publicity about child abuse has caused men to avoid child caring occupations.

VERATAC decided not to establish a state standard for this core measure until there is better understanding of what appropriate standards might be, and more consultation with local colleges about this performance measure. This decision seems appropriate. Creating incentives or requirements for male participation in traditionally female programs, when such incentives might work directly against what students themselves want, seems unwise. Similarly, while it may be appropriate to eliminate artificial barriers to women participating in non-traditional programs, this is different from creating incentives or requirements for local programs to enroll specified proportions of women, at least in the absence of information that this is what women students themselves want (or say they want). For the moment the alternative of providing local programs with information about the gender composition of their non-traditional programs may be sufficient for another round of discussion about appropriate measures.

However, while the state may not want to establish a standard at this time, it could still ask local programs to develop a report stating existing measures being taken and improvements proposed, for all programs that fall below the state averages. This step, which is intermediate between voluntary and mandatory standards, would be consistent with pulling up the lowest-achieving colleges and programs. It might also provide valuable information to the Chancellor's Office about potential practices for other

colleges to consider. (See Option 14.)

The issue of completion of non-traditional programs is simpler, however: if either men or women are enrolled in non-traditional programs, VERATAC has decided that the state's goal should be equality in completion rates between men and women — since lower rates of completion for women in "male" fields — or of men in "female" occupations would be a cause for concern. Therefore the state performance measure should be the difference between male and female completion rates, measured separately for traditional male occupations and for traditional female occupations. The state's initial performance measures should be the average state figures for 1998-99; thus continuous improvement in retention within non-traditional programs implies that this difference over time will move towards zero. The local performance indicators should, in parallel, be the difference between the completion rates for men and for women in non-traditional occupational programs.

The current data indicate that there are numerous programs that present difficulties — and none of them are surprising. Typical areas with more than 75% men include aeronautical and aviation technology; agricultural power supply; automotive technology; construction crafts technology; diesel technology, drafting and design, and drafting technology; electronics and electrical technology; engineering; environmental hazardous material conservation; and fire control. Those with more than 75% women include consumer education and home economics; cosmetology; dental technicians, fashion, and French (not really of concern under VTEA). While a few of these occupational areas show changes in gender composition from year to year, in most of them the patterns are quite stable — consistent with the notion that gender segregation is stable and responds largely to long-run movements in wages, opportunities, and gender conceptions.

Changes in State Performance: At the community college level, occupational segregation by gender has been much more resistant to change than it has been for professional and managerial occupations requiring baccalaureate and professional degrees (Blau and Ferber, 1992). The changes that have taken place at the top of the occupational structure have been responses to long-run pressures to gain women access to traditionally male professions. Similarly, we would anticipate that any real changes in the state's measure — particularly on the proportion of women in non-traditional fields — would be relatively slow, in response to long-run changes in shortages, wage patterns, and publicity about alternative occupations. Dramatic responses to cyclical variation, and to long-run demographic changes in California, seem less relevant to this measure.

Regional Variation: It's quite possible that the gender composition of programs varies substantially among colleges, as local labor markets and their acceptance of non-traditional gender roles vary, though we have no information about such potential patterns. One option, therefore, is for the Chancellor's Office to carry out exploratory studies of the variation among colleges in both enrollment and completion rates in non-traditional occupations. (This is quite similar to Option 6.) This would also provide better information about whether certain specific occupational areas have particularly large variation in enrollment and completion rates, and whether problems are relatively uniform across programs and colleges rather than being confined to a few programs and colleges

Unintended Effects: Pressures to include individuals in non-traditional programs

could lead to several problems. At the community college level — in contrast with high school, for example, which has historically been the focus of federal legislation for vocational education — students are presumed to be adults, well-informed about their choices, and responsible for their own decisions. Therefore imposing requirements on colleges to meet certain enrollment targets contradicts the presumption that students ought to be making these choices. Of course, community colleges students may not be making fully-informed decisions, and complaints about the lack of guidance and counseling are legion at the community college level (as at the high school level). However, imposing requirements on local colleges is an awkward way to encourage them to improve their guidance and counseling — even though VTEA specifically allows local programs to use federal funds for "career guidance and academic counseling" and "nontraditional training and employment activities" (Section 135(c)(2) and (14)). Furthermore, in some studies, women in non-traditional occupations have found it difficult to find employment in non-traditional fields, and therefore persuading them to enroll in such areas may not be beneficial to them. An unintended consequence of forcing local programs to enroll higher proportions of women — or men in "female" occupations, which might be even harder — is that these individuals would be made worse off. This possibility again suggests that performance standards be avoided until more is known about current patterns and the possibilities for improving them. Another option is for the Chancellor's Office to continue working to improve the quality of career-oriented counseling and guidance, particularly incorporating information about non-traditional fields, but without imposing local performance standards.

Long-run Developments: The Perkins legislation is quite precise about the definition of "non-traditional training and employment", defining them as occupations in which less than 25 percent of individuals employed are of one gender (Section 3(17)). The underlying purpose, of course, is to eliminate gender-related barriers to employment. However, important this goal may be, there are barriers to certain occupations for other groups that should also be of concern. For example, black and Hispanic individuals are often under-represented in technical fields, in part because of lack of familiarity with such options and in part because of weaker K-12 preparation in math. One option, therefore, is for the Chancellor's Office to carry out analyses of other patterns of under-representation — by race and ethnicity, by economically disadvantaged status, and by linguistic status, for example (Option 165). (Because of sample sizes, some of these analyses may be possible only for large, urban colleges and districts.) Such analyses would provide information for subsequent deliberation about other conceptions of "non-traditional" occupations, and about the more general issue of barriers for groups of students to certain occupations that can in theory be overcome with improved guidance and counseling and with concerted efforts to identify and remove specific barriers (like math preparation).

III. Consistency with Other Performance Measures

In the current environment of accountability, different systems of measuring the performance of community colleges have proliferated. In this section we consider two others that have been developed specifically in California — Performance-Based Accountability (PBA), enacted in SB 645 and sometimes referred to as the state report

card on education and training; and the Partnership for Excellence (PFE), which provides additional funding for community colleges in exchange for improved performance on a number of measures. In addition, the federal Workforce Investment Act (WIA), enacted in 1998, contains a number of performance measures that will probably affect a large number of community colleges — though the effects of WIA are still quite uncertain. Therefore community college in California face no fewer than four independent sets of performance measures.

In examining the three systems (in addition to that required by VTEA) in this section, we are principally concerned with three issues:

(1) Is the process that creates performance measures generated one that allows community colleges an appropriate voice? This is important because of our concern, articulated in the introduction, that a process without adequate discussion of potential measures and their effects cannot lead to improvement in the accountability systems over time. Moreover, an accountability system that is seen as illegitimate may lead to local efforts to circumvent its intention, undermining its potential in improving local programs.

(2) Are the performance measures themselves consistent with one another, in the sense that local improvements in one measure also lead to improvements in other related measures? It's important not to place local colleges and programs in a bind, where improvement on one measure causes them to look worse on another measure. (This might happen, for example, where short-run and long-run employment measures coexist.) In addition, the different systems should minimize the reporting requirements on local colleges.

(3) Do the different accountability mechanisms all contribute to the ultimate goal of local program improvement? If they do not, or if they do not add to the incentives in other accountability systems, then they may be ineffective or redundant.

The Workforce Investment Act (WIA)

There is currently great uncertainty about performance measures in WIA for at least three unrelated reasons. First, the Department of Labor has signaled that it will develop precise ways to measure its required performance measures — rather, as in VTEA, allowing states to develop their own measures — in order to assure comparability across states. However, at this writing (June 1999), the federal Department of Labor (DOL) had not yet completed work defining these measures, or clarifying other ambiguities related to performance.

Second, the state of California has not yet passed legislation creating a state Workforce Investment Board (WIB), though various pieces of legislation are currently being considered. Therefore there is currently no state entity with which to discuss performance — or any other dimension of WIA. The state must, according to WIA, develop a state plan before April 1, 2000, in order to begin implementing WIA by July 1, 2000. This means that the first state plan for WIA is due at the same time as the revised (second year) state plan for VTEA. It would, of course, be desirable for the two state planning processes to overlap so as to assure consistency between VTEA and WIA (see Option 16). However, in the absence of a state WIB there is currently no way to assure that this will happen.

Third, there remains considerable ambiguity about the power of states and of local WIBs, an issue addressed in a subsequent section.

Given these uncertainties, it is difficult for community colleges to know precisely what will happen under WIA. However, in one sense, the performance measures required by WIA do not create many special problems for community colleges because they are consistent with measures already being required by other accountability systems. WIA contains two different kinds of performance measures. One set, defined in Section 122, is required for those providers of education and training who want to be eligible to provide training services. Information about local education and training programs cannot be provided by One-Stop Centers — which purport to be all-purpose centers for information about education and training options — unless they have been certified as eligible (see Sec. 122(d)(2)(f)); because we assume that all (or most) community colleges will want to have their programs listed with One-Stop Centers, we assume that they will want to provide the performance information required to be eligible. There are three measures required here:

(1) Program completion rates. These could be consistent with the completion rates required by VTEA Core Measure 2, including expanded completion rates taking into account completion of licenses, industry-generated credentials, and certificates at early exit points as well as certificates and Associate degrees. However, uncertainty about specifically how DOL defines these measures still exists.

(2) The rate of entering unsubsidized employment, or employment for which no public subsidy is available to employers. This is consistent with the employment rate required under Core Measure 3. However, WIA also specifies that this employment rate could be based on the "occupation related to the program conducted"; in this case community colleges will need supplemental information about related placement. (See Option 12.)

(3) The wages at placement. This is easily calculated from the current Unemployment Insurance wage record data. However, we note that this is a short-term measure, rather than a longer-term measure discussed above. Given potential differences between short-term and longer-term measures of performance, the full-information solution is to provide both measures, for the entire set of education and training providers.

These three measures must be calculated for all students enrolled in WIA-approved programs — as is appropriate for providing information about the quality of programs to prospective students. However, there are many details that remain to be worked out. For example, the employment rate calculated for VTEA Core Measure 3 will be reported separately for completers and non-completers. For instance, should the WIA measure of employment be based on completers only, on those students who have completed some minimal amount of education (like 12 units), or on all students who enter regardless of how little they complete? For purposes of programs providing information to One Stop Centers, an important consideration is comparability of information across very different kinds of programs; it would be absurd to compare calculation rates of two-year Associate programs with completion rates of 15-week job training programs. Indeed, the issue of comparability across very different types of programs is a knotty problem that emerges also in PBA, one that we think the state has not yet fully addressed. For the moment, therefore, we can only suggest that the WIA planning processes at both the state and the local levels include sufficient representation from education providers like community colleges. (See Option 17.)

Then, for local providers of education and training that do provide training, the following performance measures must be calculated only for "participants who receive assistance", not for all students:

(1) Entry into unsubsidized employment, among those completing the program. Here again "completion" needs to be defined, in order to determine the group for whom this measure will be calculated, and the period of time after completion also needs to be defined. However, this measure is roughly consistent with VTEA Core Measure 3.

(2) Retention rates in unsubsidized employment, 6 months after the first day of employment. This rate can be measured by employment rates in the third quarter after a completer first is registered as being employed. This is also consistent with the retention measure under VTEA Core Measure 3.

(3) The earnings of completers, six months after initial employment. This is readily measure by UI wage record data, as earnings during the third quarter after post-completion employment commences. While this is not one of the measures required by VTEA, it is one of the proposed measures for the Chancellor's Office to consider over the long run (see Option 13).

(4) The rates of licensure, attainment of degrees, or other measures of skills of the completers of such programs. This is consistent with the VTEA Core Measure 2 on completion rates. It raises, as does Measure 2, the issue of how broadly completion should be defined. By collecting additional data to broaden the conception of completion (see Option 8), the results would be useful for WIA measures as well as VTEA measures.

Overall, then, the WIA performance measures we have described so far are roughly consistent with those required by VTEA, though there remain many details to be worked out. However, there are several other complications. One is that, according to DOL, these performance measures must be calculated, only for participants receiving services under WIA, separately for adults (over 21), for youth 19-21, and for dislocated workers. Evidently, it will be necessary for local One-Stop Centers to generate information about which individuals are awarded ITA's (Individual Training Accounts), and whether or not they are dislocated workers. Then either the WIA system can calculate performance measures, using the UI wage record data, or they can provide this information to the Chancellor's Office to calculate WIA performance measures. The former method probably facilitates comparisons among different types of WIA providers (e.g., community colleges versus ROC/ROPs versus community-based organizations); the latter facilitates comparing WIA and non-WIA participants within community colleges, and would be more useful to colleges seeking improvement.

WIA also specifies that the state must establish local measures of performance; if a local area fails to meet these levels of performance, then it must receive technical assistance (including help in developing a performance improvement plan, consistent with VTEA), and if such failure continues for two years, the Governor may reorganize the local program.

In developing local indicators, the economic, demographic, and other characteristics of the populations to be served shall be taken into account. This suggests that the kinds of regression adjustments we proposed for future VTEA measures (see Option 6) will also be useful to WIA measures; therefore the methodologies of such adjustments could be jointly developed. (See Option 18.)

In addition, the performance accountability system requires collecting two

measures of "customer satisfaction", one developed for employers and another for participants. The participant measures could be collected in exit interviews, if these were standardized for the state as a whole (see Option 8). The employer measures will require a new survey of employers, with all the difficulties of high costs and low response rates typical of employer surveys. However, the current pilot project by EDD to collect information about the relatedness of employment might serve as a vehicle for broader measure of employer satisfaction.

There remains one additional source of great uncertainty. WIA specifies that "the Governor, or the local board, may require a provider to submit other program-specific performance information to obtain such subsequent eligibility" (Sec. 122(d)(2)(B)), This raises the specter of local WIBs specifying a variety of additional performance information in order for local providers to be eligible — even though the Act also specifies that local boards that require expensive information provide local programs with "access to cost-effective methods for the collection of the information involved". At the moment it remains unclear precisely how this provision will develop; while local boards may not exercise this prerogative, it is also a provision that would result in many different performance measures being required around the state, a situation that would make participation in WIA difficult for community colleges. Therefore the Chancellor's Office should participate actively in the development of WIA performance measures, in part to forestall local colleges being caught in a crazy-quilt of different measures around the state (see Option 19). While the recommendation to participate actively in WIA deliberations may seem obvious and banal, in the past community colleges in California and many other states have often ignored developments in job training as being "not their business" — though in the context of WIA such refusal to participate could cost them dearly in the future, when they want to participate in WIA.

There remains one further issue to consider: Will the WIA performance measures serve to improve the quality of local programs, or will they simply be measures to monitor and reward (or punish) states and local providers? The requirement of performance data for programs to be eligible to be WIA participants is surely useful: it's consistent with information that community colleges will need to provide for VTEA anyway, and it will force other providers who otherwise escape accountability — like community-based organizations and ROC/ROPs — to measure their performance. But performance measures for clients who are subsidized by WIA will potentially be much less useful. In the first place, there are likely to be relatively few such individuals; then the data will be reported for local providers and programs, separately for those 19- 21, for adults, and for dislocated workers. The problems of small sample sizes and cells with too few individuals to draw any meaningful conclusions — problems that now affect the community college's data system, despite having many more students than WIA is likely to have — are likely to be substantial. The results of these performance measures might be useful for the state, but they are much less likely to be useful for local program improvement.

In the second place, the kinds of individuals who will receive training services under WIA will be a highly selected group — those who cannot find a job even after they have received core services and then intensive services. They are likely to be the individuals with the greatest barriers to employment — the least formal schooling, the least work experience, the most health and mental health problems, the greatest family

problems, the greatest need for support services — and therefore will not be comparable to the "regular" community college population. In fact, no training program has ever been shown to be effective with this population,

and it's unlikely that WIA-sponsored programs will be particularly effective either — though no doubt some will be marginally more effective than others. It will therefore be difficult to compare the performance of this group with the performance of other community college students, and equally difficult to know what information about performance with this group means for programs generally.

Performance-Based Accountability (PBA)

SB 645, passed in January 1996, gave the State Job Training Coordinating Council the responsibility for developing measures of performance for a variety of employment-related programs, including community colleges, job training under JTPA, welfare-related programs under JOBS, the Employment Service, vocational rehabilitation, and the Employment Training Panel. Adult education, ROC/ROPs, and the Food Stamp Employment and Training Program were included in the Act but not in the first set of reports. PBA is being progressively implemented; that is, initial experiences will inform subsequent efforts to extend the PBA system. In general, the purpose of the Act was to develop performance measures, in order to create pressure for local improvement. This aspect of PBA seems to have been relatively successful: a number of participants report that the programs subject to PBA are now more conscious of their outcomes. (From this vantage, it will be valuable to include adult education and ROC/ROPs in subsequent efforts, since these are programs that badly need greater attention to effectiveness.)

In addition, the effort to develop comparable performance measures for a variety of different programs assumes that comparisons across programs are valuable. This proves to be much more problematic, as we will outline below.

PBA has so far defined six performance measures:

1. The rate of employment, defined as the proportion of completers found in UI wage record data. This is similar to the employment rate used as VTEA Core Measure 3. It suffers from the same problems in under-reported employment, since UI wage records are incomplete, and would therefore benefit from the efforts to develop a "hybrid" data set (see Option 11).

2. The length of employment retention, defined as the employment rate the third year after program participation. While this is not currently a measure required by VTEA, it is certainly consistent with the idea that postsecondary occupational education ought to focus on preparation for long-run employment, and with the option that the Chancellor's Office develop longer-run measures of employment (see Option 13) including longer-run employment rates and earnings. These measures are also simple to calculate from the existing UI wage record data.

3. Earnings before and after program participation: This measure examines earnings in the four quarters after education, and compares it with earnings in the last quarter before completion. Again, this is not a measure currently required by VTEA, but it is consistent with longer-run measures that the Chancellor's Office might consider (option 13). The measure of earnings before and after programs completion is also the form in which many community college results have already been reported (e.g.,

Friedlander, 1993; Sanchez and Lanaan, 1997 and 1998).

4. Rate of change in Unemployment Insurance status: This measure, while not related to any of the VTEA measures, is closely related to employment patterns (see measures 2 and 3 above) since only those who are unemployed are eligible for UI. This measure is therefore completely consistent with incentives embedded in various other performance measures to enhance the employment rates of community college students.

5. Rate of change from tax receiver (on TANF or SSI) to taxpayer: This measure, not calculated in the first year because of the lack of data, is most relevant for welfare recipients. This measure is therefore consistent with employment rates and earnings patterns, though measured for welfare recipients only. This measure raises the question of the special effects of different forms of education, training, and other employment-related services for welfare recipients, some of whom have proved to be extremely hard to move into employment (see the programs noted in footnote 5, for example). It is certainly useful for community colleges to report their effects separately for welfare recipients as well as other special groups of students; for one effort to do so see Wisely (1998), who found very low rates of completion for welfare recipients. However, a more focused alternative is for the state to conduct a special study of the effectiveness of different programs in moving welfare recipients into self-sufficiency, particularly in this period of time when the most-employable recipients are already off the welfare rolls. (See Option 20.)

6. Rate of advancement to higher education: This measure, also not calculated in the first-year report, is a conventional transfer rate from community colleges to public four-year colleges. It is applicable to both academic and occupational programs; indeed, there are some indications that transfer rates are as high from occupational programs as from academic programs, from such areas as business, information technology and computing, health occupations, and engineering technologies with obvious paths to four-year colleges (Palmer, 1988; Grubb, 1991). For occupational programs, transfer rates are also considered in VTEA Core Measure 3, which counts transfer as one of the components of "success", along with employment or military service. (The actual measure proposed excludes students transferring to four-year colleges so that they do not show up as "failures.") Transfer is a phenomenon, like "completion" that is quite controversial in community colleges, because so many students enroll without any intention of transferring or with unclear conceptions of what transfer and four-year college requires. Therefore, the PBA Committee should consider reporting transfer rates for different populations of students — for those in academic and occupational programs separately; for students with different amounts of coursework completed; for students who stated upon entry that they intended to transfer, compared to those without such intentions. (See Option 8 for the origin of such data.) And, like many other aspects of performance in occupational education, transfer rates vary with the economic status and family background of students, and probably with local economic conditions; therefore the PBA Committee might consider a study of such variation among colleges, parallel to the study of regional variation in occupational performance measures (see Option 6). These different analyses would provide more refined information about transfer patterns than is available from simple performance measures. (See Option 21.)

We note first of all that the PBA performance measures are — with the possible exception of #5, focused on the success of welfare recipients — also measured by or

included in measures required by VTEA. Therefore we see no real inconsistencies between VTEA accountability and PBA. By the same token, because the two sets of measures overlap so much, the PBA measures will be largely redundant once the required VTEA measures are adopted. PBA adds to the VTEA measures in two ways, however — in requiring a performance measure related to welfare recipients (#5), and in requiring an explicit transfer rate (#6). Both of these would benefit from more detailed analysis, as our Options 20 and 21 clarify.

Performance-Based Accountability is widely described, both by participants and by its own publications, as an "incremental" system that should learn from its initial efforts in continuing to develop performance measures. While it is not our role to carry out a full evaluation of the PBA system, we do offer several observations about the current process.

First, the process of establishing PBA measures and the process of establishing VTEA measures are independent of one another, although there are some individuals who serve in both. VTEA measures are considered by VERATAC, all of whose members understand the workings of community colleges; in contrast, the PBA Committee is dominated by individuals without experiences in community colleges, many of whom do not provide education at all. While this clash of different programs may be precisely what the PBA legislation intended, it is

Second, the processes by which PBA measures and VTEA measures are calculated are quite different. VTEA measures are calculated by the Management Information Services and Policy Analysis Division within the Chancellor's Office, while PBA measures are calculated under a contract with AMPG. The kinds of exploratory analyses and more detailed analyses that we have consistently proposed as options are therefore relatively easy to carry out for VTEA measures, while such variations for PBA measures require a modification of the contract with AMPG. Furthermore, the process of calculating performance measures for occupational education ends up generating expertise about statewide patterns within the Chancellor's Office, while the similar processes for the PBA Committee do not result in greater expertise among government officials. In this period of greater accountability, the state might want to use its resources in ways that maximize the opportunities for data analysis and exploration, and that maximize the development of expertise within the state; therefore the state might want to end the process of subcontracting for data analysis with outside consultants, and instead build up the state's own expertise with data analysis and performance measures of various kinds. (See Option 22.)

Third, unlike the VTEA and WIA performance measures, the PBA accountability system does not define local accountability measures, nor does it require some response from local programs if they fail to meet such measures. (This is also true of the Partnership for Excellence, examined below.) It therefore relies on voluntary responses to low performance. Within an accountability system that measures only state levels of performance, there is an incentive for every local program to be unconcerned about its own performance, since local improvements — especially local improvements that come at great local cost — cannot possibly improve the state's measures by very much. Therefore the PBA process might, in parallel with the VTEA and WIA accountability processes, incorporate a system of local performance measures well. Alternatively, at least for community colleges, PBA might rely on the local VTEA and WIA performance

measures — augmented perhaps by a transfer measure and measures of effectiveness of welfare recipients — since there is so much overlap.

Finally, one aspect of PBA has been quite difficult, and is conceptually almost impossible to carry out successfully: comparability among different types of programs. Many participants have pointed out that the programs now being compared through PBA are extremely different in their purposes: community colleges provide short and long-term education; job training programs often provide only job search assistance and short-term training, and often act as subcontractors to other agencies rather than direct providers; welfare-related programs serve a population with particular and often multiple barriers to education and employment; vocational rehabilitation also serves a distinctive population, often with services completely different from those found in community colleges or job training. The idea that any performance measure could compare across these program areas — rather than serving to compare specific programs within each program area — is extremely difficult, and needs to consider vast differences in the kinds of individuals served, services provided, and goals addressed (including both employment and non-employment goals). Where other states have created data systems that combine many different program areas, they have generally shied away from such direct comparisons. For example, Florida has assembled a data system (FETPIP, the Florida Education and Training Performance Indicator Program) that included virtually every social program in the state, including welfare and the correctional system. However, FETPIP is used to track individuals as they move among programs, not to compare success across different programs. Similarly, Washington State has amassed data on the effects of various education and training programs; however, the only measure that is used to compare programs is a benefit-cost calculation, which is arguably a measure applicable to any kind of public program.

(This measure has been used, for example, to indicate that community colleges have short-term benefits to participants at the expense of taxpayers, and long-run benefits to both participants and taxpayers as increased taxes and reduced welfare benefits more than offset the initial costs of providing education; however, adult education and JTPA youth programs, with costs outweighing benefits, need "substantial changes"; see Workforce Training and Education Coordinating Board, 1996.) Our conclusion is simply that the goal of comparability across very different programs areas can probably not be readily achieved, and that the PBA accountability process should probably concentrate on using its measures to enhance performance within each of the areas under its purview, rather than comparing across these areas.

Partnership for Excellence

Finally, the Partnership for Excellence is a program which has provided additional funding to community colleges — at least \$100 million a year — in return for enhanced performance. The program contains five state-level performance measures, to be met by 2005:

1. An increase in the number of transfer students from 69,574 to 92,500. The intent underlying this measure, to increase transfer, is consistent with VTEA Core Measure 3 and with PBA measure #6.
2. An increase in the number of degrees and certificates awarded, from 80,799 to 110,500. The intent to increase completion, is consistent with VTEA Core Measure 2 and

the completion rate required by all programs who want to be eligible for WIA.

3. An increase in the rate of successful course completion from 68.1 percent to 70.6 percent, with targets highest (from 77.2 percent to 80.0 percent) for occupational courses and lowest (from 60.3 percent to 62.5 percent) for basic skills courses. "Successful" completion is defined as passing with a grade of C or better, consistent with VTEA Core Measure 1.

4. An increase in the numbers of successfully completed apprenticeship courses, advanced vocational courses, introductory vocational courses, businesses benefiting from contract education, individuals benefiting from contract education, and individuals receiving fee-based job training (e.g., through subcontracts from job training programs). While successful completion of apprenticeship and advanced occupational courses is consistent with VTEA Core Measure 1, the measures related to contract education are new and different measures. The measure related to fee-based job training may be difficult to meet now that JTPA has been transformed into WIA, with very different priorities for spending money

5. An increase in the numbers of students completing coursework one level above their prior level of basic skills. The measure, related to the success of developmental education in enhancing the competencies of its students, is certainly consistent with VTEA Core Measure 1 if this measure were redefined in terms of competencies rather than grades (see Option 7). Otherwise, however, it is a different goal related more to the performance of developmental education than to occupational education.

In general, then, the spirit underlying PFE is consistent with performance measures in VTEA, WIA, and PBA. However, the performance goals are measured in extremely awkward ways, since they are (with the exception of #3) expressed as numbers rather than rates. There are, of course, two ways to increase numbers over several years: one is to increase rates of performance with a stable base of students, and the second is to increase the numbers of students while the rates of performance stay the same. For community colleges in growing regions of the state —largely, for all but rural community colleges — targets defined in terms of numbers are likely to be easy to meet in the coming years. And given the counter-cyclical nature of enrollment, any downturn in the economy in the coming years is likely to increase enrollments and therefore make it easier to reach these targets. We therefore suspect that the steady expansion of community colleges is likely to take care of the targets embedded in PFE. (Conversely, in areas of the state with declining enrollments, or in periods of time when enrollments decline for reasons outside the control of colleges, it might be impossible to meet targets stated in terms of numbers.) The solution, of course, is to express targets for PFE in terms of rates rather than numbers (Option 23).

A second awkward aspect of PFE is that, like PBA, it has no local performance indicators. As we mentioned above, the incentive for local colleges to improve is simply missing if state measures are not translated into local performance indicators — the free rider problem once again. (This issue shows up in complaints that PFE funds are not being spent for program improvement, but instead are being used for various pet projects and politically-motivated purposes — a charge that would take armies of accountants to verify, we should say.) Therefore PFE — and PBA as well — are only weakly related to local improvement. The state might therefore consider some consistency in the state-local

dimensions of its various accountability systems — that is, making sure that any system of state accountability measures also incorporates a set of local performance indicators as well (Option 24). Like the local indicators to be developed for VTEA (see Options 3 and 6) and WIA, these local indicators could also be adjusted for student characteristics and intentions and local economic conditions, rather than assuming it is appropriate for all colleges and programs throughout the state to meet the same local standards.

The Goal of Consistency

With the imminent development of WIA accountability measures, there are no fewer than four separate sets of accountability measures affecting California community colleges. Fortunately, there prove to be relatively few major conflicts among these four accountability systems and their performance measures, though a number of detailed issues remain to be resolved. However, the processes developing each of the four are now independent of one another — though the procedures for setting standards are linked by several individuals who serve on multiple committees, including several analysts — and the procedures for developing data sets and analyzing them are also independent, a fact that may limit data exploration and analysis. The goal of promoting greater understanding about students and local programs — their behavior and progress through education and training, their passage into employment and subsequent success — varies among these accountability systems. The process of establishing state goals is not, in our view, as valuable as certain more detailed analyses would be (see, for example, Options 5, 6, 8, 9, and 11), but PBA and PFE currently do not engage in such detailed analyses. The goal of promoting local effectiveness varies substantially among them: PBA and PFE lack local indicators matched with state measures; the performance measures in PFE are awkwardly stated; and PBA will be largely though not entirely redundant once VTEA accountability mechanisms are put into place. And of course there remains substantial uncertainty about what will happen under WIA.

Given the current variation in the quality of accountability systems, the state might consider consolidating some or all of them. (See Option 25.) For example, the two that pertain exclusively to community colleges — VTEA and PFE — could be combined within the Chancellor's Office. The accountability systems that include several different kinds of programs — PBA, and soon WIA — present greater difficulties, partly because there has been a history in California of weak cooperation between education agencies and those responsible for job training. However, creating a unit responsible for accountability in a variety of employment-related programs would provide another forum in which to debate the different approaches to accountability, the possibilities for consistency among them, the areas where consistency is impossible because of different program goals and populations, and the best ways of encouraging local improvement. Then the different accountability systems in the state might work to reinforce one another, rather than remaining essentially independent as they now are.

In the end, the decisions to consolidate accountability systems are part of larger, political decisions whether or not to coordinate different types of employment-related programs in the state. WIA strengthens the development of One Stop Centers, envisioned as comprehensive employment-related service centers; WIA invites states to submit joint state plans for a wide variety of programs including adult and youth job training, secondary and postsecondary vocational education, adult education, and the

Employment Service, vocational rehabilitation, and several others; and it provides incentive funds for states with strong performance in three programs — job training, vocational education, and adult education. The WIA governance process could incorporate a broader variety of programs, including those (like community colleges, ROC/ROPs, and ETP) that are essentially state funded, in addition to the federally-funded activities under WIA. If California begins moving to coordinate its various education and training-related programs through the WIA process, then it would make sense to consolidate accountability systems as well. But if California continues its history of weak coordination at the state level,

then it is unlikely that its accountability systems can be consolidated in any meaningful sense. Indeed, consolidating accountability mechanisms, without enhancing coordination of programs themselves, could only undermine the technical quality and the efficacy of existing accountability systems. Therefore the state of California must decide as a whole whether it will begin moving toward greater coordination in its recent system of employment-related education, training, and other services, or whether it will continue with a fragmented system, before such decisions can be made about accountability.

IV. Other Issues Generated by Performance Measures: Funding and Resources

The development of performance measures, coming from both state and federal legislation, cannot be isolated from other aspects of how community colleges function. In several ways, the performance measures already generated and their subsequent refinement create incentives and imperatives that contradict other patterns in California. In this section we note three potential problems — one related to the incentives in enrollment-based funding, the second related to the special funding problems of occupational education, and the final one concerned with the resources necessary for analytic work — that should be considered in addressing the overall problem of accountability and performance.

Quantity and Quality:

Performance Measures versus Enrollment-based Funding

The performance measures required by the legislation reviewed in this report create new incentives for local colleges to improve performance in particular ways — or, more precisely, to improve their indicators of performance, which can be improved without improving performance itself (as the history of JTPA reveals). At the same time, the core funding of community colleges in California (as in most states) is generated through formulas that reward colleges for increasing enrollments. Indeed, there is currently no funding that is directly tied to performance: Perkins and WIA funds are allocated to the local level by formulas that do not include performance; the Partnership for Excellence allocates its funding — its uncertain and varying funding, we might point out — according to an enrollment-based formula, and the state report card requires performance measures to be reported but without any funding attached to them. In many cases, improving performance requires resources — for example, for improved guidance and counseling, placement and transfer centers, improved instruction in occupational and developmental courses — that may be greater than programs can generate in increased

enrollment, and conversely increased enrollment and therefore funding can be achieved in ways that reduce performance. Local colleges are therefore under contradictory incentives, with the major fiscal incentive being to increase enrollment, while the incentives to enhance performance are still largely symbolic and hortatory, with no money attached to them. College presidents may therefore feel themselves justified in continuing business as usual, stressing enrollment over performance.

There are only a few ways to reconcile the incentives for enrollment and the incentives for performance. One option, of course, is to move toward performance-based funding, in order to modify or weaken enrollment incentives. Funding formulas can also combine enrollment-driven and performance-driven components; for example, Florida plans to start allocating 15 percent of its occupational funding through performance-driven formulas starting sometime in 1999. However, we note that Florida has been developing performance measures since the late 1980s, and has always had a more state-controlled system of community colleges and vocational schools than California; we suspect that what might work in Florida would take more time to develop in California. And prudence suggests that the effects of performance-based funding in Florida be evaluated before other states move in this direction.

A second option would be for the state to provide categorical funding for those institutional practices that, while they might enhance performance, do not increase enrollments in any direct or certain way. These might include, for example, enhanced career-oriented guidance and counseling, improved matriculation practices, placement and transfer centers, teaching-learning centers to improve the quality of teaching, and incentives for faculty to become prepared in teaching methods. (The funding for such activities and centers could also be based partly on enrollments, on the utilization of such services, on student demographics and local labor market conditions, and on specific indicators of performance, to address the difficulties individual colleges face and to reward high-quality efforts.) Such an option would provide the resources to enhance performance, whereas the proposed performance measures create pressures but without the resources to improve performance.

The default option, of course, is to let a system of contradictory incentives persist. Under this option, we fear that some colleges will make valiant efforts to attain performance standards — particularly those that are already conscious of quality, and those with relatively high levels of funding or that are growing anyway and therefore will have sufficient funds. But others will respond more to enrollment incentives, particularly those that are strapped for funds or that are low-performing and view the attainment of local performance standards as impossible. This is a scenario under which the good become better and the mediocre become worse, intensifying rather than narrowing differences in performance over time.

The Special Funding Problems of Occupational Education

Many of the local problems in occupational education are related to a fundamental issue rarely addressed in legislation. The basic funding for community colleges (and most other educational institutions) is driven by conceptions of academic education, where the basic cost of providing instruction varies with the number of students but not with the kinds of programs offered. Sometimes, in K-12 education, there are additional state resources for high-cost students like disabled students, or for high-cost programs like

vocational programs, or for students in high-cost districts like urban districts or rural areas. A few states also provide differential funding for high-cost programs in community colleges. But in California, there are no such supplements for high-cost programs, or high-cost students. The basic conception of education seems to be driven by academic classrooms with a teacher, students, a blackboard, some chalk, and little else.

But occupational programs are different from most academic programs in several ways, and their funding requirements are different. Some have higher costs for current expenditures, either because class sizes have to be kept low in workshops and labs, because workshops and labs have extended hours, because of the need (or desirability) to create internships and other work experiences, or because of the requirement to develop employer advisory committees. And some have higher costs for capital expenditures, particularly in areas — like machining, automotive occupations, certain health occupations, computer-related occupations, drafting — where space requirements for workshops are substantial, equipment costs are high, and equipment keeps changing. (Of course, there are substantial differences among occupational areas, and some academic areas including science also have substantial needs for space and capital equipment.) Capital equipment is not separately budgeted in California, so that occupational instructors need to badger their administrations and local employers for materials and equipment, and often use federal (VTEA or Perkins) funds for basic capital equipment or for updating their equipment and materials.

The effects of differential costs, in a funding system that fails to recognize such differences, are several. One is that occupational instructors seem to spend more of their time scrounging materials and equipment, and finding internships and other work-related placements, than their academic colleagues; they often resent the disparities between the roles they have to play and those of their colleagues who need not worry about space, equipment, advisory committees, or adapting to rapidly-changing technology. In addition, administrators in community colleges are often quite "rational" in their decisions: since the institution's fiscal incentive is to expand high-enrollment, low-cost programs — to generate "profits" to be used elsewhere — there are clear fiscal incentives to limit high-cost, low-enrollment programs, even if these are in occupational areas with high community demand and high placement rates. Finally, in the current situation there are no slack resources, and program innovation and improvement almost always require some slack. These are conditions, then, under which improvement in response to performance measures becomes difficult, particularly since there is no guarantee that the effects of increased performance will generate additional resources for the programs anytime in the future.

This problem is one that predates the pressures for accountability and performance measures. But the development of performance measures, concentrating (at least in VTEA and WIA) on occupational programs, exacerbates these problems. The only real solution would be for California to move toward a funding system in which cost differentials among programs are recognized, and in which capital funding is a regular component of community college allocations.

State and Local Resources for Accountability

It should by now be obvious that the development of performance measures and standards, and their further refinement over time, will require additional resources — at

both the state and local levels. Many of the options we have outlined require additional research with existing data; some of them require new data to be collected; and some of them require data to be made more widely available, in more understandable forms, than is currently the case.

At the state level, it seems relatively clear that there are inadequate resources in the Chancellor's Office. We have outlined a number of additional analyses of existing data sets that need to be carried out if existing performance measures are to be further developed and refined. The complaint that state data often comes to local colleges in formats that are difficult to understand may require resources to develop more user-friendly formats. The development of a hybrid data system, based in part on UI data and in part on data collected through student and employer follow-up questionnaires, will require additional work at the state level. And the consultation processes with local colleges, as state performance measures are put in place and as the state develops local indicators and some mechanism for enhancing local performance, will take additional resources as well. While local administrators are sometimes exasperated by the information available from the state level, they recognize that state resources are inadequate to the tasks being demanded through state and federal legislation, and that state resources for analysis and consultation need to be enhanced.

At the local level, accountability will also require additional resources if it is to be done right. Some colleges do not now have institutional researchers who can interpret performance indicators to administrators and instructors; others will need to increase resources in institutional research. The development of a unified mechanism of collecting information from entering students about their purposes, and the development of student and employer follow-up questionnaires if the state decides to supplement UI data with a hybrid system, will require still other local resources. Furthermore, the interpretation of state- and locally-generated data — for example, the effort to determine why particular programs have low placement or completion rates, or why women are not enrolling in non-traditional employment — and the development of alternative solutions is something that, while it can be promoted at the state level, must take place locally, with local researchers, administrators, faculty, and counselors. All of this will take resources, in some form. If there are no additional resources, then either the time required to implement performance measures will detract from other responsibilities, or performance measures will not be given the attention they need.

We fear that the consequences of under-funding various accountability measures is simply that the state will have to remain content with relatively simple performance measures, rather than developing any of the alternatives we have outlined and making progress toward more refined measures of performance. At the local level, the consequences of under-funding are likely to be a lack of responsiveness to local indicators, either because local administrators and instructors don't know what they entail, because they don't have the resources to respond, because they come to view local indicators as illegitimate, or simply because the performance-related incentives are so much weaker than those attached to enrollment. In short, if the state of California wants its complex system of accountability and performance measures to work as it should — to improve the quality of local programs, in the interests of students, employers, and taxpayers alike — then it will need to provide sufficient resources as well as initiative in the creation of performance measures.

V. A Summary of Options

Throughout this report we have presented a series of options. Some of these are actions that the Chancellor's Office could pursue; others would be the responsibility of other agencies and committees, including the state Workforce Investment Board created under WIA, or the Performance-Based Accountability Committee, or they would require the legislature and the governor to enact legislation or appropriate funds. In this section we simply summarize these options.

Option 1: The Chancellor's Office could formalize the roles of various members of VERATAC so that they represent various constituencies — occupational deans, chief instructional officers, and institutional researchers among them — and serve as mechanisms of communication between these groups, local colleges, and the Chancellor's Office.

Option 2: The state could develop a pilot project with a small number of colleges — both those with strong offices of institutional research, and those without institutional researchers — in order to translate state data into formats that occupational deans and faculty can more readily understand, and to develop appropriate comparisons so that local colleges can see how they are doing relative to others.

Option 3: If local programs fall below local performance standards, then the option of requiring local improvement plans is consistent both with requirements in VTEA and with what seems most feasible in the short run. This option also provides the Chancellor's Office with information about what local colleges are doing to improve. This information, plus information about what other successful colleges are doing, could be the basis for an annual report on promising practices so that all colleges could learn from their peers. This would involve the Chancellor's Office in providing a kind of technical assistance to improve local programs that is complementary to the role of performance measures in creating incentives to improve.

Option 4: In setting state standards, the Chancellor's Office could create incentives for the worst-performing colleges to improve by setting standards that measure the proportion of colleges falling below some standard. This method of setting state standards recognizes that there are ceiling effects, and recognizes that the greatest problem and room for improvement is in low-performing programs. For example, for VTEA Core Measure 1, the state standards could be that 77 percent of programs have successful completion rates of 70 percent or better. Then improvement on this standard would require that those programs with less than 70 percent C's or higher improve their performance, while better-performing colleges continue performing above this level.

Option 5: The Chancellor's Office could consider including data on students in need of remedial/developmental education into its system, both for this purpose and to monitor the performance of low-achieving students in general. If the state decides to take this approach, then it will need to confront the variation in the diagnostic tests and standards used from college to college. While it is possible to use unstandardized

estimates of the proportion of students needing remediation in state data systems, it would be better from the perspective of uniformity to shift to a statewide diagnostic process — though this would require agreement among 107 colleges on what process to use.

Option 6: Because local performance surely varies with the preparedness and composition of students in different colleges, the Chancellor's Office could carry out a pilot study of the effect of such variables as the preparedness of students (as measured for example, by diagnostic tests at matriculation or high school grades), language or immigrant background, and income measures like eligibility for student grants and loans. For employment-related outcome measures, local economic and employment conditions might also influence outcomes. Such studies could then become the basis for regression adjustments to local performance measures, similar to those that have been used for JTPA.

Option 7: Over the long run, the Chancellor's Office could consider measuring academic as well as occupational competencies in VTEA Core Measure 1, by including grades in related academic courses. Other possibilities include the inclusion of information from licensing exams and industry-generated credentials. A more difficult alternative would be to develop state standards for all occupational areas, a process that would require extensive discussion to reach consensus on such standards.

Option 8: The Chancellor's Office could use the existing system of student questionnaires upon entrance into community colleges, standardized across all colleges, and another questionnaire for those leaving community colleges. These could then be used, in conjunction with other information about the receipt of credentials and licenses, to develop more refined measures of completers than are currently available.

Option 9: For all measures, the Chancellor's Office may want to start using truly longitudinal measures for all those measures — like completion rates — that examine the behavior of students over time. Longitudinal measures are more accurate and interpretable than the current measures, which are ratios of two independent numbers, and problems in timeliness can be overcome by concentrating on stable patterns.

Option 10: The Chancellor's Office might consider requiring uniform practice in awarding certificates and Associate degrees — now that there will be consequences to performance measures and consistency has become important in ways that were not true before.

Option 11: The state could develop a "hybrid" data system, adding to UI wage record data with surveys of students unlikely to be found in the UI wage records — for example, those in programs where self-employment is common, those moving out of state, and those working for federal and state government.

Option 12: The state of California could consider extending the current pilot project being undertaken by EDD to collect information about the occupations of

employed individuals, in order to calculate measures of the relatedness of employment to prior education and training. It could also continue to investigate alternative ways of measure the relatedness of employment.

Option 13: The Chancellor's Office could experiment with alternative measures of the quality of employment as measured by longer-run measures of employment, over a period of time longer than a half year, upward mobility over time as measured by real earnings growth, and earnings themselves. For example, measures of earnings growth between the year before completion and three years after completion have been calculated, and are certainly better measures of long-run employment than are employment rates just after completion.

Option 14: While the state may not want to establish a standard for enrollment in non-traditional occupations, it could still ask local programs that fall below the state averages on these measures to report existing measures being taken and improvements proposed. This step, intermediate between voluntary and mandatory standards, would be consistent with pulling up the lowest-achieving colleges and programs. It might also provide valuable information to the Chancellor's Office about potential practices for other colleges to consider.

Option 15: The Chancellor's Office could carry out analyses of patterns of under-representation other than gender — for example, patterns by race and ethnicity, by economically disadvantaged status, and by linguistic status — in different occupational areas, particularly those with high earnings and growth rates. Such studies could form the basis for other conceptions of "non-traditional employment", and would provide information about the more general issue of barriers to certain occupations.

Option 16: State plans for the initial implementation of WIA and for the second year of VTEA (the revised state plan) are both due April 1, 2000. The two planning processes could be coordinated so that the resulting accountability measures are consistent with one another.

Option 17: The development of performance measures under WIA and under VTEA could be coordinated with one another, at both the state and local levels, in order to work out many small details of such measures and to assure consistency.

Option 18: The local measures of performance under VTEA and under WIA could be jointly developed, since the purposes are similar and the methodologies for adjusting these measures for social, demographic, and economic conditions ought to be similar as well. Completely separate systems of local indicators — as might happen, for example, if WIA contracted its performance measures with an entity completely independent of the group developing VTEA measures — would only increase confusion at the local level.

Option 19: The Chancellor's Office, and community colleges generally, should be well-represented in state and local councils (like WIBs) that make decisions about state

and local accountability measures under WIA. One reason is that many decisions remain to be made about WIA performance measures, and consistency with other community colleges measures is important. A second reason is that the potential power of local WIBs to specify other performance measures must be balanced against the potential harm to local colleges of facing a bewildering variety of performance measures.

Option 20: As a complement to the measure #5 required by PBA, the state could invest in a specific study of the success of various education and training programs in moving welfare recipients into self-sufficiency.

Option 21: The PBA Committee could consider a more extensive study of transfer, to provide greater understanding of the transfer rates required among its performance measures. This would involve calculating transfer rates for different populations of students — for those in academic and occupational programs separately; for students with different amounts of coursework completed; for students who stated upon entry that they intended to transfer, compared to those without such intentions. (See Option 6 for the origin of such data.) And since transfer rates vary with the economic status and family background of students, and probably with local economic conditions, the PBA Committee might consider a study of such variation among colleges, parallel to the study of regional variation in occupational performance measures (see Option 6).

Option 22: In this period of greater accountability, the state might want to use its resources in ways that maximize the opportunities for data analysis and exploration, and that maximize the development of expertise within the state. Therefore the state might end the process of subcontracting data analysis with outside consultants (as happens with PBA measures), and instead build up the state's own expertise with data analysis and performance measures of various kinds.

Option 23: In order to achieve consistency among accountability systems and to enhance the incentives for improvement, rather than merely expansion, the PFE performance measures could be expressed in terms of rates rather than numbers.

Option 24: The state could enhance the consistency of the state-local dimensions of its various accountability systems — that is, making sure that any system of state accountability measures also incorporates a set of local performance indicators as well, particularly for PBA and PFE that now lack local measures. Like the local indicators to be developed for VTEA and WIA, these local indicators could also be adjusted for student characteristics and intentions and for local economic conditions (see Option 6), rather than assuming it is appropriate for all colleges and programs throughout the state to meet the same local standards.

Option 25: The state might consider consolidating some or all of its accountability systems, including the technical analysis that is required. The two that pertain exclusively to community colleges — VTEA and PFE — could be combined within the Chancellor's Office. The accountability systems that include several different kinds of programs — PBA, and soon WIA, and the incorporating VTEA and perhaps

PFE as well — could be developed within a unit responsible for accountability in a variety of employment-related programs. Then the different accountability systems in the state might work to reinforce one another, rather than remaining essentially independent as they now are. However, the consolidation of accountability systems probably requires greater coordination of different programs than now exists, something that might (but need not) be accomplished through the WIA governance mechanism.

References

- Anderberg, M., & Pfeiffer, J. (1998, May). A Field Guide to Automated Follow-Up. Commissioned by the Employment and Training Administration, U. S. Department of Labor.
- Betts, J., & McFarland, L. (1992, July). Safe port in a storm: The impact of labor market conditions on community college enrollments. *Journal of Human Resources* 30(4):741-765.
- Blau, F., & Ferber, M. (1992). *The Economics of Men, Women, and Work*. Englewood Cliffs, NJ: Prentice-Hall.
- Friedlander, J. (1993, Fall). Using wage record data to track the post-college employment and earnings of community college students. Santa Barbara, CA: Office of Academic Affairs, Santa Barbara City College.
- Grubb, W. N. (1991, March/April). The decline of community college transfer rates: Evidence from national longitudinal surveys. *Journal of Higher Education*, 62(2), 194-222.
- Grubb, W. N. (1996). *Learning to Work: The Case for Re-Integrating Education and Job Training*. New York: Russell Sage Foundation.
- Grubb, W. N. (1997). The returns to education in the sub-baccalaureate labor market, 1984-1990. *Economics of Education Review*, 16(3), 231-246.
- Grubb, W.N. (1999). *Learning and Earning in the Middle: The Economic Benefits of Sub-Baccalaureate Education*. Occasional Paper. Community College Research Center, Teachers College, Columbia University.
- Grubb, W.N., and Associates (1999). *Honored but Invisible: An Inside Look at Teaching in Community Colleges*. New York and London: Routledge.
- Grubb, W. N., Badway, N., Bell, D., King, C., Herr, J., Prince, H., Kazis, R., Hicks, L., & Taylor, J. (1999, January). *Toward Order From Chaos: State Efforts to Reform Workforce Development "Systems"*. MDS-1249. Berkeley, CA: National Center for Research in Vocational Education.
- Grubb, W.N., Brown, C., Kaufman, P, and Lederer, J. (1989, April). *Innovation versus Turf: Coordination Between Vocational Education and Job Training Partnership Programs*. MDS-001. Berkeley: National Center for Research in Vocational Education.
- Joint Management Team, *Regional Workforce Preparation and Economic Development Act* (1998, December). *Regional Workforce Preparation and Economic Development*. <http://www.regcolab.cahwnet.gov/rclciwdp.htm>.
- Palmer, J. (1988). *Bolstering the community college transfer function: An ERIC review*.

Community College Review, 14, 53-63.

Sanchez, J., and Lanaan, F. (1997, Winter). The Economic Returns of a Community College Education. *Community College Review* 25(3):73-76.

Sanchez, J., and Lanaan, F. (1998, May). Assessing the Post-College Earnings of Students: Benefits of Attending California Community Colleges — Methods, Analysis, and Implications. Costa Mesa, CA: Coast Community College District, Vocational Education and Institutional Research.

Wisely, C. (1998). AFDC Students in the California Community Colleges, 1992-93. Unpublished M.A. Dissertation, California State University, Sacramento.

Workforce Training and Education Coordinating Board (1996). Workforce Training Results: An Evaluation of Washington State's Workforce Training System. Olympia: Workforce Training and Education Coordinating Board.

This procedure would simply ask each college to report, based on their local diagnostic tests, the proportion of students passing and failing according to locally-specific criteria. We ignore in this report the much larger problem of creating valid diagnostic procedures; for problems with current tests, see Grubb and Associates (1999), Ch. 5.

The same four performance measures are mentioned in both Sec. 122 (d)(A)(ii) and in the section on core indicators of performance in Sec. 136(b)(2)(A).

We note that such reorganizations have been quite rare under JTPA, and there is no reason to think they will be any more common under WIA.

On the importance of being at the table, see also Grubb et al. (1999).

See, for example, the evaluations of the Minority Female Single Parent Demonstration project and of New Chance, both reviewed in Grubb (1996).

We note that self-sufficiency, which is conceptually the same as the PBA conception of being a "taxpayer" rather than "tax receiver", is not the same as moving off welfare, since individuals can leave TANF and still be receiving publicly-subsidized child care, medical care, housing, social services, food stamps, and other goods and services. In the past several years TANF has been successful in reducing the numbers on welfare without necessarily moving them into self-sufficiency.

Arguably — though many analysts (even economists) would acknowledge that benefit-cost analysis is inadequate because of its inability to encompass benefits and costs that cannot be monetized. For example, the goal of independence, valuable in both welfare-related programs and vocational rehabilitation, cannot readily be incorporated into benefit-cost measures.

The two need not be distinguished from one another: One effective way of teaching developmental education is to create learning communities of developmental courses and occupational courses, integrated with one another so that students see the point of mastering basic skills. An effective learning community would result in higher completion rates of all courses and a greater rate of advancement in basic skills.

In the context of examining coordination in a variety of states, we earlier concluded that coordination was relatively weak in California; see Grubb, Brown, Kaufman, and Lederer (1989). The current effort, the Regional Workforce Preparation and Economic Development Act, does not provide any greater hopes for coordination since initial report provides no specifics (Joint Management Team, 1998).

This procedure would simply ask each college to report, based on their local diagnostic

tests, the proportion of students passing and failing according to locally-specific criteria. We ignore in this report the much larger problem of creating valid diagnostic procedures; for problems with current tests, see Grubb and Associates (1999), Ch. 5.