

# **CALIFORNIA STATE UNIVERSITY**



## **A Pilot Evaluation of the Outcomes of Elementary Subject Matter Programs In the California State University System**

**Study Sponsored in Part by the  
California Commission on Teacher Credentialing**

**Spring 2003**















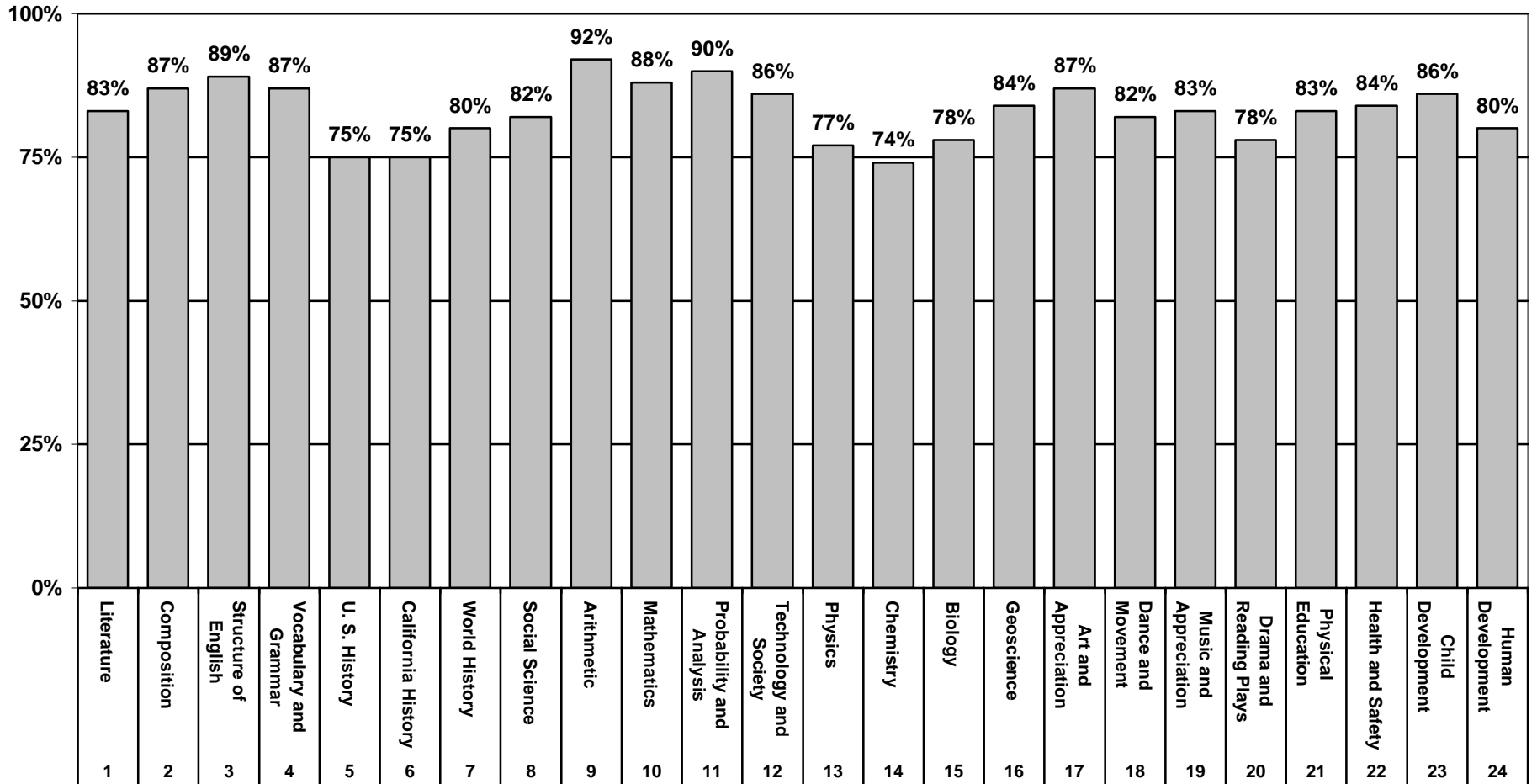
## An Interpretation of Supervisors' Evaluations in the Pilot Study

Figure 1 on page 8 displays the *percentages* of CSU teaching graduates who were *well-prepared* or *adequately-prepared* to know and understand each subject, according to the teachers' supervisors. These percentages were shown in Column (2) of Table 1. The graph highlights the fact that, with some exceptions, the effectiveness of CSU preparation varies according to the broad areas in which the 24 subjects are taught. The first four subjects comprise the broad area of *reading-language arts*, followed by the broad areas of *history-social science* (subjects 5-8), *mathematics* (9-12), *science* (13-16) and the *visual and performing arts* (17-20). In the aggregate, the responding supervisors reported the CSU graduates' knowledge and understanding to be very strong in the broad areas of *mathematics* (particularly arithmetic) and *reading-language arts* (particularly the structure of English). They also judged the graduates' knowledge and understanding to be somewhat less strong in the broad area of *visual and performing arts* (with graduates having the weakest understanding of drama), and considerably less strong in the areas of *history-social science* (with the weakest understanding being in U. S. history and California history) and *science* (with the weakest knowledge being in chemistry). In the aggregate, the supervisors also reported the CSU graduates' knowledge and understanding of *physical education* (21), *health* (22) and *human development* (23 and 24) to be at intermediate levels relative to subjects in the other areas. These aggregate findings apply only to *the overall system of campus-based programs* as these programs were experienced by a single cohort of graduates who began teaching in 2001-02. Neither the findings nor the interpretations describe individual programs or campuses within the CSU system.

In examining the percentages in Table 1 and Figure 1, it is important for the pilot study's margin of error to remain in mind. The *reported percentages* are the most reliable estimates of preparation effectiveness in 24 subjects. The percents in Table 1 and Figure 1 have a *greater probability* of being "true" than any other percentages. But their accuracy is *not certain* because they are based on a finite sample of participants drawn from a larger population, and because errors may have occurred in measuring the effectiveness of preparation. In estimating the "true" percentages, the next-most-probable values (after the reported percentages) are *one point greater and less than the reported values*. After those, the next most likely percentages are *two percent higher and lower* than the values in Table 1 and Figure 1, and so on. Given the size of the response group and the magnitude of the reported percentages, we can have 95 percent confidence that the true percentages are no more than six points greater or lower than the percents in Figure 1 and Table 1.

As an alternative to testing the CSU graduates to directly measure their subject-matter expertise, the pilot evaluation solicited assessments from their school-site supervisors. As a measure of the CSU system's overall effectiveness in subject-matter preparation, the *validity* and *reliability* of the supervisors' reports (in the aggregate) derive from the following facts: The graduates were selected randomly from the entire cohort of CSU graduates; 63 percent of the selected supervisors elected to participate in the pilot evaluation; 89 percent of the responding supervisors had supervised six or more new teachers in their administrative careers; 94 percent had supervised the designated CSU graduates for more than six months; 97 percent had observed the graduates' classrooms during active instruction on three or more occasions; 87 percent had conferenced with the graduates about their teaching on three or more occasions; the evaluation data suggest that the participating supervisors evaluated only those subject areas in which they had observed or conferenced with the graduates; and a teacher's knowledge and understanding of subject-matter is commonly a focus of observation and conversation when a supervisor visits a classroom or conferences with a teacher. Alternative sources of evaluation data with equivalent validity and reliability may not be available to the university.

**Figure 1: Percentages of K-8 School Supervisors Reporting that Randomly-Selected CSU Teaching Graduates Were Well-Prepared or Adequately Prepared to Know and Understand Twenty-Four School Subjects that the CSU Graduates Taught**



## Evaluation of CSU Program Effectiveness by CSU Teaching Graduates

First-year classroom teachers who had completed CSU programs of subject-matter preparation were asked to report on *their knowledge and understanding* of the same school subjects that were listed in Table 1. The pilot evaluation assumed that teaching a subject provides an opportunity for teachers to “come to grips” with their knowledge and understanding of that subject. Accordingly, when CSU graduates responded to each subject in the list of 24 school subjects, they were asked this question: *When you taught material related to this subject in your class, how well did you know and understand the material?* The evaluation did not ask the CSU graduates *in the abstract* how well they knew and understood the subjects. An abstract self-evaluation would be difficult for respondents to complete, and would likely yield responses that were somewhat inflated. Instead, the pilot study focused on knowledge and understanding of *material that the graduates had taught*. The term *material* denoted an open-ended category of ideas, information, skills, abilities and attitudes that form the content of a subject and that the CSU graduates had actually taught in their classes during the first school year after they received their CSU-based credentials. This open-ended denotation was *by design* because a wide array of content elements are taught in different schools and districts, even within specific subjects such as literature, history, music and health.

CSU graduates were directed to make no response “if you did not teach a subject in your classroom this year.” Like their supervisors, then, the graduates assessed the knowledge and understanding *that they actually used*, not the knowledge and understanding that they thought they possessed. The CSU graduates also had four response options that were equivalent to those of their supervisors.

- 3 = When I taught material related to this subject, I knew and understood it *well*.
- 2 = When I taught material related to this subject, I knew and understood it *adequately*.
- 1 = When I taught material related to this subject, I knew and understood it *somewhat*.
- 0 = When I taught material related to this subject, I knew and understood it *not at all*.

On the following page, Table 2 summarizes the responses of CSU graduates to the pilot evaluation question about their knowledge and understanding of the 24 subjects. (For an explanation of the four types of statistics in Table 2, see the preamble to Table 1 found on page 5.) Given the fact that some subjects are taught at *limited grade levels* (e.g. California history, world history, physics, chemistry and biology), and that state standards and school district policies do not emphasize the importance of some other subjects for *any grade level* (e.g. geoscience and dance), and that subject specialists teach some subjects in place of classroom teachers in some schools (e.g. physical education and music), we should expect the numbers of respondents in Column (1) to vary considerably. In fact, the *Ns* in Table 2 mirror these state and district policies very accurately. Again, participants in the pilot evaluation appear to have been reliable in adhering to the admonition that they assess their knowledge and understanding of *the subjects they had taught* in their first-year classrooms.

**Table 2:**  
**CSU Graduates' Evaluations of Their Knowledge and Understanding of the**  
**K-8 School Subjects in Their Teaching Assignments**

| <b>K-8 School Subjects in Which<br/>CSU Graduates Evaluated<br/>Their Knowledge and Understanding</b> | <b>Descriptive Statistics for<br/>Each School Subject</b> |                 |                    |                  |
|---|---|-----------------|--------------------|------------------|
|   | (1)<br><b>N</b>   | (2)<br><b>%</b> | (3)<br><b>Mean</b> | (4)<br><b>SD</b> |
| 1. Fiction and Non-Fiction of Varied Genres, Periods and Cultures                                     | 185   | <b>81 %</b>     | <b>2.15</b>        | 0.78             |
| 2. Writing, Composition and Oral Language (Including Listening)                                       | 207   | <b>88 %</b>     | <b>2.31</b>        | 0.72             |
| 3. Structure of English Including Sounds, Words and Sentences   | 209   | <b>86 %</b>     | <b>2.26</b>        | 0.73             |
| 4. Vocabulary, Word Derivations, Spelling and Grammar   | 209   | <b>82 %</b>     | <b>2.20</b>        | 0.76             |
| 5. Major Events & Developments in U. S. History (Not California)                                      | 169   | <b>64 %</b>     | <b>1.76</b>        | 0.82             |
| 6. Major Events and Developments in California History  | 151   | <b>60 %</b>     | <b>1.74</b>        | 0.86             |
| 7. Major Developments in the History of Peoples Outside the U. S.                                     | 151   | <b>56 %</b>     | <b>1.66</b>        | 0.82             |
| 8. Concepts of Economics, Geography, Government, Psychology   | 155   | <b>67 %</b>     | <b>1.88</b>        | 0.85             |
| 9. Solving Problems in Arithmetic Including Shapes & Measures   | 194   | <b>89 %</b>     | <b>2.43</b>        | 0.70             |
| 10. Solving Problems in Mathematics Including Pre-Algebra   | 178   | <b>90 %</b>     | <b>2.45</b>        | 0.72             |
| 11. Understanding Estimation, Probability and Basic Data Analysis                                     | 178   | <b>85 %</b>     | <b>2.29</b>        | 0.78             |
| 12. Technology (Including Computers) and Its Effects on Society                                       | 142   | <b>76 %</b>     | <b>2.01</b>        | 0.88             |
| 13. Physics Including Topics Such as Energy, Matter and Motion  | 133   | <b>62 %</b>     | <b>1.76</b>        | 1.01             |
| 14. Chemistry Including the Structure and Properties of Matter  | 117   | <b>57 %</b>     | <b>1.63</b>        | 0.99             |
| 15. Biology Including Topics Such as Life Forms and Evolution   | 143   | <b>77 %</b>     | <b>2.05</b>        | 0.84             |
| 16. Topics in Astronomy, Geology, Oceans, Climate and Weather   | 150   | <b>68 %</b>     | <b>1.87</b>        | 0.92             |
| 17. Lessons in Art Including Appreciation, Drawing and Crafts   | 173   | <b>76 %</b>     | <b>2.05</b>        | 0.83             |
| 18. Lessons in Dance Including Social Dance & Creative Movement                                       | 114   | <b>60 %</b>     | <b>1.66</b>        | 1.10             |
| 19. Lessons in Music Including Appreciation, Singing & Instruments                                    | 141   | <b>63 %</b>     | <b>1.78</b>        | 1.05             |
| 20. Drama Lessons Including Making Plays & Reading Them Aloud   | 146   | <b>66 %</b>     | <b>1.86</b>        | 0.98             |
| 21. Physical Education Instruction in Fitness, Skills and Sports                                      | 175   | <b>78 %</b>     | <b>2.11</b>        | 0.79             |
| 22. Instruction in Health Including Nutrition, Diseases and Safety                                    | 167   | <b>77 %</b>     | <b>2.05</b>        | 0.79             |
| 23. The Teacher's Knowledge of Child Growth, Learning & Thinking <sup>3</sup>                         | 182   | <b>90 %</b>     | <b>2.41</b>        | 0.74             |
| 24. The Teacher's Knowledge of Human Development After Childhood                                      | 156   | <b>82 %</b>     | <b>2.27</b>        | 0.81             |

<sup>3</sup> Subjects 23 and 24 were included in the pilot evaluation because state law requires that they be included in an accredited subject-matter program for Multiple Subject Teaching Credentials. The fact that teachers do not teach these subjects did not appear to interfere with the graduates' capacity to assess their own knowledge and understanding.

## **An Interpretation of CSU Graduates' Evaluations in the Pilot Study**

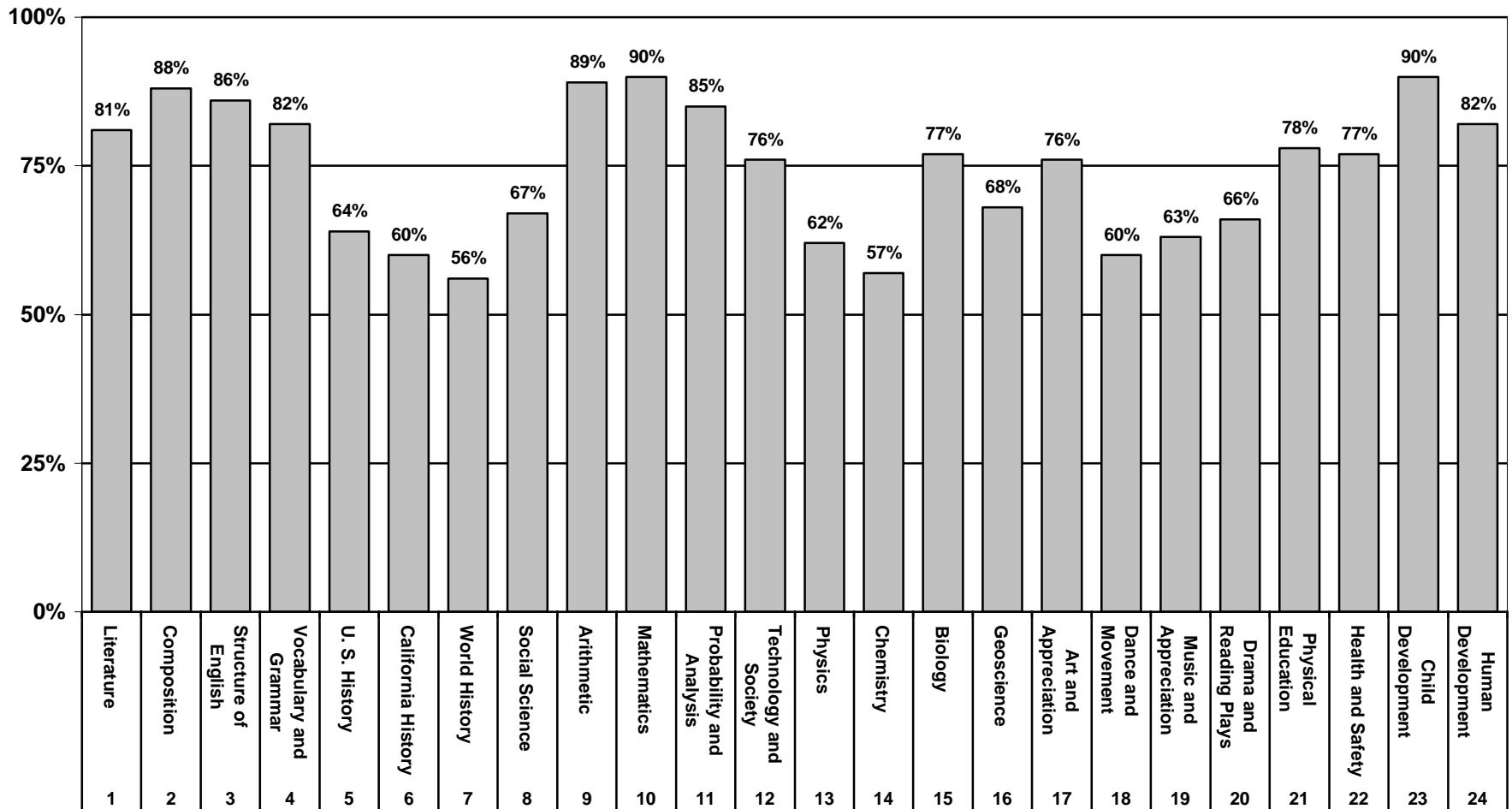
Figure 2 on the following page highlights the percentages of graduates who were *well prepared* or *adequately prepared* to know and understand the material they taught in each subject, which were also shown in Table 2 above. In the broad area of *reading-language arts*, the graduates' assessments of their knowledge and understanding were roughly equivalent to those of their supervisors in Table 1 and Figure 1. Similarly, the graduates' assessments within the broad area of *mathematics* (Items 9-11) were almost equivalent to their supervisors' evaluations. In other subjects, however, the graduates reported their own knowledge and understanding to be weaker than was reported by their supervisors. CSU evaluations of the effectiveness of *professional* teacher preparation programs consistently yield similar findings in which *supervisors report the preparation of first-year teachers to be stronger than is reported by the teachers themselves*. The reduced judgments of the graduates are attributable to their inexperience compared with the many years in which their supervisors have gained perspective and acquired realistic expectations about how much first-year teachers should know and be able to do. Despite the fact that many beginning teachers assessed their subject-matter knowledge and understanding more critically than did their supervisors, the two participant groups were similar to each other in rank-ordering the broad areas: both groups reported the graduates to be strongest in the broad areas of *mathematics* and *reading-language arts*, followed by *physical education* and *health*; the *visual and performing arts*; *science*; and *history-social science*.

## **Overall Interpretations of the Pilot Evaluation Results**

*An Interpretation Pertaining to Validity and Reliability of the Pilot Evidence.* In all subjects of the school curriculum, the pilot evaluation questions were directly aligned with the *academic content standards* and the *curriculum frameworks* of the State of California. These are the same standards and frameworks that govern California's school curriculum and accountability system. Unless the CSU graduates and/or their supervisors misunderstood the pilot evaluation questions, their replies have considerable validity for CSU programs of subject-matter preparation. There is no evidence that the participants misunderstood the questions; in fact, the pilot evaluation yielded considerable evidence to the contrary. Response rates were high; response patterns reflected the directions given to the participants; the numbers of respondents to specific questions varied in predictable ways; and very few individuals in the sample indicated they were confused or unclear about the questions or the response options. As was the case with the pilot evaluation of CSU *professional* programs in 2001, *the present study suggests that CSU graduates and their school-site supervisors are valid and reliable in answering questions about the subject-matter preparation of teachers in the CSU system.*

*An Interpretation Pertaining to the Utility of Evaluating the Outcomes of Subject-Matter Programs.* By focusing on as many as 24 distinct subjects of the school curriculum, the pilot evaluation yielded a profile of evidence regarding the CSU's *systemwide strengths and weaknesses* in subject-matter preparation (K-8). Even with relatively small numbers of respondents, the pilot study produced very different results for distinct subjects of university preparation and school teaching. If the pilot's findings lack utility or value, it is because they only describe the outcomes of the *entire CSU system* while shedding no reliable light on the outcomes of specific programs at particular CSU campuses. Responsible academic officers are not in a position to use the pilot findings in academic decisions because the pilot evaluation was small in scale. With the participation of larger numbers of graduates

**Figure 2: Percentages of Randomly-Selected CSU Teaching Graduates Reporting that They Were Well-Prepared or Adequately-Prepared to Know and Understand Content Material that They Taught in Their K-8 Classes and That Related to Twenty-Four Subjects of the School Curriculum**



and supervisors, a *full-scale evaluation* would provide outcomes-oriented evidence that (1) would be *aligned* with the teaching responsibilities of CSU's future graduates, (2) would be *specific* to each CSU campus, and (3) would answer *uniform evaluation questions* among all CSU campuses and programs. CSU campuses do not have access to such information from other sources. Given the increasingly important role of subject-matter preparation in K-12 schools, campuses have a growing need for campus-specific data regarding program strengths and weaknesses.

*An Interpretation of the Potential Value and Benefits of Expanding the Scope of the Evaluation.* If a full-scale evaluation included the same questions and response options as the pilot study (with few major changes in their language), the CSU system would receive *longitudinal evidence* of changes that may be occurring in the *effectiveness* of subject-matter programs over time. Additionally, CSU campuses and the entire system would learn *still more* about subject-matter preparation if some or all of the following additional questions were also included in future evaluations.

- Whereas the pilot evaluation focused on the preparation of K-8 teachers who earn Multiple Subject Credentials, expansion of the evaluation scope would enable CSU to ascertain its effectiveness in the subject-matter preparation of *high school teachers* who earn Single Subject Credentials. State academic standards apply to high school students as well as children in the lower grades. Moreover, state accreditation standards for Single Subject Credentials are changing quickly. Beginning in 2003, it would be feasible to assess the outcomes of subject-matter programs in English and mathematics, followed by studies of other subjects in subsequent years.
- It would be important for the CSU to ascertain how its teaching graduates assess the actual value and helpfulness of *key elements* of their subject-matter preparation, based on their teaching experiences. For example, graduates could be asked to evaluate advisement, the availability of program information, their sense of belonging to a “cohort,” and the assessments used to ascertain their end-of-program expertise. Answers to such questions would be particularly valuable if the program elements and features to be evaluated by the graduates were aligned with current and forthcoming accreditation standards for subject-matter programs (SB 2042).
- In order to “unpack” the *sources* of subject-matter knowledge and understanding on the part of CSU's first-year teaching graduates, they could be asked how many courses they took in each subject, how many of these courses were completed at CSU campuses and at community colleges, and how many were completed with a grade of B or better. In addition to revealing whether or not subject-matter expertise comes primarily from lower- and upper-division courses in the subjects, responses to these questions could reveal how many courses are needed for typical graduates to be well prepared in each area.
- At least one large-scale evaluation of subject-matter preparation should include a random sample of CSU teaching graduates who qualified for credentials by passing standardized state examinations in their subjects. By including such a sample in an evaluation, the CSU could, for the first time, compare the relative subject-matter expertise of exam passers and program completers – two groups of new teachers who have met the state's subject-matter requirement in alternative ways.

These questions (and others like them) illustrate the potential value and benefits of sponsoring larger-scale evaluations of CSU subject-matter preparation.

*An Interpretation of the Importance of Alignment Between K-12 Curriculum and Teacher Subject-Matter Preparation.* In two areas of the K-8 school curriculum – mathematics and language arts – all students in all grades are instructed by their classroom teachers every day. Student advancement from grade to grade, and student eligibility to graduate from high school, depend on learning to read and compute. All CSU graduates in self-contained classrooms teach reading and math; this is not true of any other subject. Significantly, the pilot evaluation indicates that, in recent years, more than 85 percent of CSU’s teaching graduates have been *well prepared* or *adequately prepared* to know and understand the content of California’s language arts and mathematics curricula.

Compared with the evidence of strong effectiveness in the core elementary subjects of language arts and mathematics, the pilot-study findings suggest that preparation in other subjects is less effective. Here, the most likely explanation stems from the sequence of recent changes in California education policies. California’s new academic standards for grades K-8 were adopted and implemented *while and after the pilot-study participants enrolled in CSU subject matter programs*, not before they did so. Due to the standards, K-8 schools dramatically shifted the focus of curriculum and instruction in several subject areas. In the broad area of history-social science, for example, the State Board of Education dropped the prior focus on world history and world cultures, which was in effect *while the pilot-study participants were preparing to teach*. In place of this focus, the Board adopted a set of standards that emphasize United States and California history. *In the schools, this new focus went into effect while the pilot-study participants were securing their credentials and their initial teaching positions.* Additionally, the new academic standards for K-8 students were not the only policy changes to take effect *while the pilot participants were earning credentials*. California also adopted new accreditation standards for the subject-matter programs that are the focus of this study. For the first time in California history, subject-matter accreditation standards require CSU programs to be *fully aligned and congruent* with the K-8 academic standards for pupils. The pilot-study participants were CSU students during a prior period when the state’s accreditation standards encouraged each CSU campus to emphasize the subjects that the campus deemed to be important. These “pre-2042 standards” did not call for any level of alignment between subject-matter teacher preparation and the state’s learning expectations for K-8 students. *Some pilot-study participants experienced a “disconnect” between subjects they studied in college and content they were expected to teach in K-8 classrooms a few years later. Such disconnected experiences probably occurred most frequently in science and history-social science, and would prompt many reasonable teachers to question the value of their subject-matter preparation in these subject areas.*

The pilot-study findings suggest that the most important priority for the next few years will be for CSU programs of subject-matter preparation to complete their alignment with the new academic standards for K-8 students. Curriculum alignment is most likely to strengthen the effectiveness of subject-matter programs – particularly in science and history-social science – throughout the CSU. The most productive course for policymakers would be to foster the earliest possible alignment of teacher programs with K-8 standards.

For the foreseeable future, it would be most responsible for the CSU system to sponsor a series of carefully-designed evaluations to track subject-matter programs and ensure they become fully aligned with the state’s high-stakes standards for K-8 children. Additional policy changes during this period are only likely to undermine the alignment process and the CSU’s opportunity to track and monitor its continued progress toward excellence in the preparation of teachers for California’s K-8 students.