

1 JOHN F. DAUM (SB #52313)
2 FRAMROZE M. VIRJEE (SB #120401)
3 DAVID L. HERRON (SB #158881)
4 PAUL B. SALVATY (SB #171507)
5 O'MELVENY & MYERS LLP
6 Embarcadero Center West
7 275 Battery Street
8 San Francisco, California 94111-3305
9 Telephone: 415.984.8700

10 Attorneys for Defendant State of California

11 **SUPERIOR COURT OF THE STATE OF CALIFORNIA**

12 **CITY AND COUNTY OF SAN FRANCISCO**

13 ELIEZER WILLIAMS, et al.,) Case No. 312 236
14)
15 Plaintiffs,) Date Action Filed: May 17, 2000
16)

17 vs.)
18)

19 STATE OF CALIFORNIA, DELAINE)
20 EASTIN, State Superintendent)
21 Of Public Instruction, STATE)
22 DEPARTMENT OF EDUCATION, STATE)
23 BOARD OF EDUCATION,)
24)
25 Defendants.)

26 STATE OF CALIFORNIA)
27)
28 Cross-Complainant,)

29 vs.)
30)

31 SAN FRANCISCO UNIFIED SCHOOL)
32 DISTRICT, et al.,)
33)
34 Cross-Defendants.)
35)

36 **EXPERT WITNESS DECLARATION RE MARGARET RAYMOND, Ph.D.**

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

I, Paul B. Salvaty, declare as follows:

1. I am an attorney with the law firm of O'Melveny & Myers LLP, counsel of record herein for defendant State of California ("the State").

2. The State has provided a list of persons whose expert opinion testimony the State intends to offer at trial of this action, either orally or by deposition testimony. The list includes Dr. Margaret Raymond, to whom this declaration refers.

3. Dr. Raymond has agreed to testify at trial.

4. Dr. Raymond will be sufficiently familiar with the pending action to submit to a meaningful oral deposition concerning the specific testimony, including any opinions and their bases, that Dr. Raymond is expected to give at trial.

5. Dr. Raymond's fee for providing deposition testimony, consulting with the State, conducting research and other activities undertaken in preparation of the attached report is \$185 per hour.

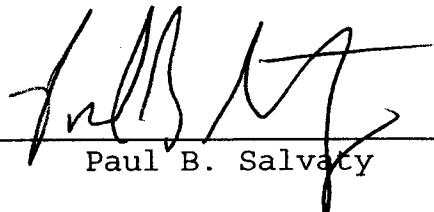
6. Pursuant to Section 2034(f)(2)(A) of the California Code of Civil Procedure, attached hereto as Exhibit A and incorporated herein by reference is a *curriculum vitae* providing Dr. Raymond's professional qualifications.

1 7. Attached hereto as Exhibit B and incorporated herein by
2 reference is Dr. Raymond's expert report. Pursuant to Section
3 2034(f)(2)(B) of the California Code of Civil Procedure, the
4 following is a brief narrative statement of the general substance
5 of the testimony that Dr. Raymond is expected to give at trial.
6 Dr. Raymond rebuts several of plaintiffs' expert reports, arguing
7 that State and plaintiffs have the same goals but that the
8 State's approach (focusing on outcomes rather than ill-defined
9 input standards) is far more likely to achieve positive results;
10 that plaintiffs fail to demonstrate that any of their sweeping
11 policy changes would improve student achievement at all, let
12 alone more than the current policies being pursued by the State;
13 that many of plaintiffs' experts' opinions are internally
14 inconsistent and contradictory of one another; and that
15 plaintiffs' various proposals are not feasible.

16
17 The foregoing statements are only a general summary of
18 the issues and conclusions discussed and documented more fully in
19 Dr. Ballinger's expert report.

20
21 I declare under penalty of perjury that the foregoing
22 is true and correct.

23
24 Executed this 18th day of April, 2003, at Los Angeles,
25 California.

26
27 
28 Paul B. Salvaty

MARGARET E. RAYMOND

205 Lou Henry Hoover Building
Hoover Institution
Stanford University
Stanford, CA 94305
(650) 725-3431 tel.
(650) 723-1687 fax.
raymond@hoover.stanford.edu

Professional Experience

- July 2000 – present** **Hoover Institution, Stanford University**
CREDO (formerly the Center for Research on Education
Outcomes) Director
- Feb. 1999 – June 2000.** **University of Rochester,**
Rochester, New York.

Direct the activities of the Center, a research organization designed to improve the development and dissemination of empirical evidence about the effect of education reform efforts. Activities include tracking of reform efforts in the United States, assistance in the creation of sound evaluation research and dissemination of results. The primary aim of the Center is to better inform education policy decisions.

- March 1986 – June 2000.** **Raymond Associates, Inc.**
Rochester, New York. President

Privately held woman-owned research and consulting business. Areas of expertise include public policy research and evaluation, organizational development, and telecommunications policy. Projects have included strategic planning, market analysis, program evaluation, forecasting, training, information systems and management consulting. Clients include federal, state and local government agencies, community non-profits, industry associations, regulatory agencies and private corporations.

Entrepreneurial activities include business start-ups for four firms, resource development for privately held companies, creation of coalitions for market expansion and joint ventures.

- October 1988 - April 1992.** **Rochester Telephone Corporation**
Rochester, New York.
Manager, Regulatory Policy and Research.

Analysis of technical, regulatory and market implications of legislative, regulatory and management proposals to determine directions for company. Recommendations based on economic, policy, and statistical analysis. Leadership in national forums addressing open network architecture, telecommunications infrastructure, and personal communications services. Analysis and recommendations regularly shared with corporate officers.

May 1981 - September 1988. Center for Governmental Research
Rochester, New York. Principal Research Analyst.

Responsibility for principal research design and execution of contract research. Projects included program evaluation, multivariate analysis, strategic planning design, data system design, economic analysis, projections and forecasting, marketing, management analysis, and training, in areas of criminal justice, education, welfare, public finance, economic development, youth services, child abuse/neglect, desegregation, job training and labor force analysis. Administrative responsibility for long range staff planning, personnel hiring and evaluation.

University Teaching Experience

- Adjunct Associate Professor**, Public Policy Analysis Program,
University of Rochester, 1996 - 2000
- Adjunct Assistant Professor**, Public Policy Analysis Program,
University of Rochester, 1992 - 1996
- Visiting Assistant Professor**, College of Applied Science Technologies,
Rochester Institute of Technology, 1992.
- Visiting Assistant Professor**, College of Arts and Science,
University of Rochester, 1985-1986.
- Instructor**, College of Arts and Science,
University of Rochester, Spring semester, 1982-1983.
- Graduate Teaching Fellow**, College of Arts and Science and
School of Medicine and Dentistry, University of
Rochester, 1979-1982.

Education

University of Rochester

- | | | |
|-------|------|---|
| Ph.D. | 1985 | Political Science
<u>Fields: Public Policy, Analytic Methods,
Federalism, Organizations</u>
Thesis written on Executive Decision Making |
| M.A. | 1983 | Political Science |
| M.S. | 1982 | Community Medicine |
| M.S. | 1980 | Public Policy Analysis |

Boston University

- | | | |
|------|------|--|
| B.A. | 1976 | Psychology, <u>with distinction.</u>
Degree conferred <u>Summa Cum Laude.</u> |
|------|------|--|

Professional Affiliations

- American Evaluation Association
- Policy Studies Organization
- Association of Public Policy and Management

ELIEZER WILLIAMS ET AL. vs. STATE OF CALIFORNIA
Superior Court of the State of California, San Francisco

DOJ Docket Number: 06121 140 SA2000CV0694

EXPERT WITNESS REPORT

Prepared by:
Margaret E. Raymond

April 2003

Williams v. State of California
Report by Margaret E. Raymond
April 2003

Introduction

I was asked to participate in this case as an expert for the State of California on the basis of recent work that I have completed on accountability systems and the determinants of student achievement. My work draws on training I received at Boston University (B.A., Psychology, *summa cum laude*) and the University of Rochester (M.S. Public Policy Analysis; M.S. Public Health; Ph.D. Political Science) and twenty years doing public policy research and program evaluation throughout the United States.

In 2000, I was asked to assist in the preparation of analysis for the defense in *Campaign for Fiscal Equity v State of New York, et al.* (New York). I have not provided testimony in education-related litigation.

I have experience as an expert witness in over 20 regulatory proceedings before the Federal Communications Commission, the Vermont Public Service Board, and the Washington State Public Utilities Board.

Summary of Testimony

Many of the State's education goals are in line with the desires of plaintiffs. The State has a strong interest in the education of its youth, and is interested in setting high standards for student academic achievement. There is little dispute that current levels of academic achievement for all the State's students do not universally meet the expectations that the State has established. Moreover, the State agrees with plaintiffs' central argument that every student deserves qualified teachers, adequate instructional materials and clean and decent facilities that are conducive to learning. On these general matters, plaintiffs and the State are in complete agreement.

The difference lies not in the ends, but primarily in the means to achieve them. Accordingly, it rests with plaintiffs to prove that their choice of means is superior to those chosen by the legislature, the Board of Education for the State of California and the California Department of Education. Plaintiffs have the burden to show that the current decisions and policies about 1) the recruitment, training and retention of teachers, 2) the mechanisms for ensuring adequate instructional materials for students, and 3) the policies related to construction and maintenance of facilities are ineffective and completely unreasonable. Further, plaintiffs must demonstrate that their proposals are a superior solution to providing a Constitutionally-adequate education to California's student population. They have not done so.

Plaintiffs have taken the position that "qualified teachers", "adequate instructional materials" and "decent facilities" are important to every child's education. The State agrees generally with these principals and currently is pursuing numerous programs and devoting billions of dollars toward ensuring that these conditions are present in every California school. But plaintiffs dismiss -- without discussion -- the State's efforts to achieve these goals. Plaintiffs argue that the State has a constitutional obligation to prescribe specific "input" standards for each of the three areas that plaintiffs deem most important. Plaintiffs further assert that the State has a "non-delegable duty" to monitor schools to ensure that plaintiffs' proposed input standards are being met. Plaintiffs also argue that the State has abandoned its duty to exercise its authority by its choice of accountability systems and its delegation of control of many operational decisions to local authorities. Finally, plaintiffs make specific recommendations to correct the alleged deficiencies in classrooms and in the state accountability system. Their proposals include:

- Eighty percent of teachers in every school should be fully credentialed in the near term, moving to 100 percent of all teachers in the future
- One current textbook for each student per subject
- Adequate physical facilities

On its face, the case advanced by plaintiffs has almost unassailable appeal. Who can be against fixing broken windows or ensuring that students be taught by high quality teachers?

To contemplate opposition to plaintiffs' proposals is to invite ridicule and scorn as "anti-kids" or "Scrooge" or, even worse. None are enviable monikers. Yet, this case is worthy of opposition – strong opposition – arising not from an aversion to the cause of public education, but because the merits of the case are flawed from a legal vantage and dangerous from a public policy vantage.

Plaintiffs' position is precarious from both theoretical and applied perspectives. The claim that textbooks, certificated teachers, and facilities are critical factors pre-supposes that these are the "right" things to focus on and that the current programs being pursued by the State are "wrong". Yet, the proof to support that position is entirely lacking. Instead, the experts that plaintiffs rely upon present a heterogeneous and rather haphazard *mélange* of ideas that lack adequate evidence of their importance.

The single unifying theme of their testimony is that the State has erred in its choice to focus on outcomes rather than continuing the historical attention to inputs. But no substantiation is provided that establishes the minimum standards for the inputs of their choice. How should a "qualified teacher" be determined? How would a district be able to determine what instructional materials are "adequate" and which are inferior? Which features of facilities should we pay attention to, and what constitutes adequacy? As at least one of plaintiffs' own expert witnesses has acknowledged, there are no universally accepted standards for these factors, and vigorous academic and political debate has yet to bring them forth. [Mintrop, p. 5; Russell deposition, p. 103, l. 23] This is one explanation for the current direction of California education policy; by focusing on outcomes, the State Board of Education and the California legislature implicitly acknowledge the diversity of views and leave the specific definition and measurement to local school authorities. Thus, one district may choose a three-year rotation of textbooks while another chooses a different cycle.

Moreover, even if there were specific definitions and standards proposed by plaintiffs, the case they have advanced would remain shaky. There is a shocking dearth of evidence presented by plaintiffs about California schools. Research used by plaintiff witnesses is highly selective and does not present any definition of what creates good education. Further, the case presented by plaintiffs ignores a considerable amount of evidence –

developed under stricter scientific research conditions than the research cited by plaintiffs – that the elements they propose to regulate are not the most influential in creating good student academic achievement. Their proof of the role of these factors is extremely weak, and even contradicted by their own witnesses. This report will present analytic results from California schools to demonstrate the weakness of plaintiffs' assertions.

This report also shows that the plaintiffs have mistakenly alleged the abrogation of State authority in its choice of accountability. Plaintiffs' argument is founded on a view of centralized control that runs deeply counter to the current organization of education in the United States and ignores many of the advantages local control can offer. Plaintiffs recognize the need for local flexibility, but their proposals actually would require enormous centralized control and give little room for local adjustment. In a state the size and diversity of California, such an approach is risky at best. It is difficult to conceive of a set of regulations that could adequately serve all the circumstances in the State's schools, yet that is what the plaintiffs seek to do. In contrast, the point of the current State focus on accountability for outcomes is to allow local districts to focus more directly on their own particular priorities and problems.

Moreover, plaintiffs are incorrect to argue that the results of current policies are a breach of responsibility by the State. The State has acted within its authority to develop an approach to accountability that is reasonable and appropriate. The legislature enacted the Public Schools Accountability Act in 1999 after input from the Board of Education, legislators and their constituents and other interested parties, including the California Teacher's Association. Opportunities were available to any interested person or party to express opinion. Opportunities were available to influence the opinions of legislatures through lobbying or grass-roots efforts. But none of these avenues were utilized. Instead, plaintiffs are attempting to circumvent the normal policy making process. Not satisfied with attempting to redefine policy through the courts, their focus on forcing upward accountability through the California Department of Education is in complete disregard for the constitutional process in California for managing policy disputes, namely the election of the legislature and the governor.

Finally, the remedies supported by the plaintiffs are unworkable and counter productive to the very results that they claim to be pursuing. Dictating the allocation of funds to specific but undefined uses such as “adequate facilities” will not yield the advances that are being promised. The promoted changes to the state accountability system would completely corrupt the system without producing any compensating gains.

Plaintiffs haven't shown that the field of education policy has developed a reliable production function for education that highlights the factors at issue in this case.

There exists today considerable debate about the critical elements needed to produce a sound education. Despite assertions from the plaintiffs' experts, there is still a great deal we do not know about the mix of factors and the extent to which they vary by types of students or school contexts. Certainly, the fact that plaintiffs must rely on the specific professional judgments of their experts to gauge the effects of system inputs and efforts underscores their continuing need to substitute subjective and/or variable assessments for objective and empirical evidence. Plaintiff expert Mintrop explicitly acknowledges this in his testimony. [Mintrop, p. 6.] Another plaintiff expert, Michael Russell, says the State's accountability program must take into account all “essentials for learning”, but he cannot say what those are. He further admits that there is disagreement over the very definitions of those elements and the extent to which they influence student achievement. [Russell deposition, p. 94, l. 25; p. 101, l.19-25] Thus, plaintiffs cannot offer empirical proof that their specific proposals would be effective in improving student learning -- because such evidence does not exist.

To be confident that plaintiffs' claims have merit, it would be necessary to study the effects of each of their proposals under conditions of controlled circumstances: that is, to study the effects on student achievement in schools where the factor is abundant compared to schools where the factor is scarce, controlling for other possible effects. Instead, plaintiffs offer a series of subjective inferences harvested from studies that have little variation and that do not control for independent influences. Thus the claim is: “Factor X must be important because where it exists, students do well” without considering “Are there cases where students do well where the factor is scarce?” and “Could students do well in schools where

the factor is present, not because of the factor itself, but because the factor is associated with other (potentially unmeasured) variables that provide the real drivers of student achievement?”

An example of this approach can be found in Mintrop’s analysis of 26 high-distress districts. [Mintrop, p. 15.] He finds that the proportion of so-called impact factors such as Free or Reduced Price Lunch participation, year-round school or the proportion of emergency credentialed teachers is higher in these districts than for the state as a whole. But he does not use the data available to him to test whether these factors are systematically related to outcomes or compare their effects to districts with low rates of these factors. Thus, we are asked blindly to accept the conclusion that adjusting the facilities or teacher quality in the manner preferred by plaintiff will improve student achievement.

Likewise, Mintrop offers specific suggestions for change in the State oversight of schools without any empirical support that the changes, if adopted, would make any difference. As a single example, he advocates for the development of “standards of adequacy for learning conditions and teaching practice.” Implicit in the suggestion is the admission that these standards do not already exist, and therefore are untested. The assumption that consensus on these matters is even possible is itself a stretch.

Plaintiffs’ case fails in two regards: (1) the purported proof that is offered is over-reaching; and (2) plaintiff witnesses and others have acknowledged counter examples that disprove the claim that these factors are essential.

Failure # 1: The Proof offered by Plaintiffs is Over-Reading.

Plaintiffs attempt to “prove” that student achievement is directly and significantly impacted when a student has at least 80% “fully” credentialed teachers, when a student has a “current” textbook to “use in class and take home,” and when a student has “decent” facilities (the definition of which is uncertain). But plaintiffs’ experts fail to provide any hard data or analysis to support their specific proposals. Even Mintrop himself accedes that “educational inputs... and processes... are connected to educational outcomes, but not in a

straightforward way. Rather they all interact with each other in ways that vary considerably according to local contexts.” [Mintrop, p. 5.]

Consider the evidence offered for the value of facilities. Sobol’s testimony affirms the consequent without any test whatsoever of its validity under control conditions. He claims that facilities are important. [Sobol, p. 8] It simply cannot be the case that facilities are important if it is possible to identify cases where they have not influenced the outcome of students. The statement is not tested against a real alternative, so we do not know which if either is correct based on his testimony. While he claims to have seen “students unlikely to succeed achieve to high standards when given the right kinds of teaching and support,” he does not provide any confidence in his position because he has not followed the basic precepts of even the simplest social science investigation. [Sobol, p. 6]

Even if the input standards proposed by plaintiffs do impact student achievement, the burden would still rest with them to prove that these factors were the most significant drivers of student outcomes and, therefore, worthy targets for limited education resources. They have not done so. As a thought exercise, consider the hypothetical: the evidence is solid that facilities are a statistically significant factor in producing student achievement gains. Even if this were the case, it would remain necessary to show that the magnitude of the effect was larger than other potential factors, such as strong leadership, libraries or teacher quality. How much more effective would increasing the number of fully credentialed teachers be than, say, banning cell phone usage by students during classes or revising the cafeteria menu to provide more wholesome selections? The case advanced by the plaintiffs does not provide the pertinent analysis that permits a reasoned comparison of their proposals.

As a general proposition, plaintiffs’ experts seem to want to shift the focus away from student outcomes and onto a variety of inputs that may or may not matter to a particular student in a particular school. This is clear in Russell’s report from its opening pages. Russell assumes accountability systems ought to focus on the environment that schools create, not on the results that schools produce in their students. [Russell, p. x.] This distinction is extremely important to the foundations of this case, as it highlights the

difference between outputs, that is, what the school does operationally, and outcomes, which concern the difference schools make in their students' education achievement.

Other plaintiff witnesses follow the same path. Most of Mintrop's testimony does not consider the effects of factors on student educational outcomes. His contention is that accountability systems should be able to diagnose the root causes of performance but the performance in question is framed as motivation of teachers, classroom environment or a positive work climate. [Mintrop, p. 7-9] The lack of attention to the key focus of the case – namely the educational attainment of students in California schools – undermines confidence in the case put forward by plaintiffs.

Darling-Hammond pretends to focus on outputs but her analysis is selective and flawed, and is another strong example of plaintiffs' quasi-serious basis of evidence. Darling-Hammond's testimony is a good example of this error. She asserts that a study by the Public Policy Institute of California (PPIC) found teacher qualifications were the strongest predictors of student achievement after controlling for the effects of socioeconomic status. [Darling-Hammond, p. 16] But, in fact, these factors are not statistically significant in many of the models in the full PPIC study. [Julian R. Betts, Kim S. Rueben, and Anne Dannenberg, "Equal Resources, Equal Outcomes? The Distribution of School Resources and Student Achievement in California, Public Policy Institute of California," February 2000 <http://www.ppic.org/main/publication.asp?i=64>] Mention of an amorphous group of "other studies" that show similar findings in Darling-Hammond's testimony receives no substantiation, and therefore, can receive no weight. [Darling-Hammond, p. 16]

The point is reinforced by my own research showing that teacher experience is important, but that certification *per se* is not an absolute requirement for successful teaching. By examining the performance of Teach for America teachers (who enter the classroom without traditional or alternative certification) compared to other new teachers in the Houston Unified School District, the study showed TFA teachers on average did as well or better than their peers. [Raymond, Margaret, Stephen Fletcher and Javier Luque. Teach for America: An Evaluation of Teacher Differences and Student Outcomes in Houston, TX 2001. <http://credo.stanford.edu>] The study created value-added measures of each teacher's

performance and then compared new TFA teachers to new teachers that were hired in Houston in the same years. By controlling for student socioeconomic status, language, special education status and other demographics, the contribution of individual teachers to their student's test score performance could be isolated.

Darling-Hammond is quick to dismiss the study in its entirety because the profile of new non-TFA teachers reports that significant numbers do not have at least a bachelor's degree. Her contention is that the comparison is unfair if the control group isn't minimally educated. This claim, however, is in error. (To be clear, the TFA report presented data suggesting that large numbers of Houston teachers are without college degrees, but we later found that to be incorrect. The error stems from a coding change within the HISD information systems that made it appear that teachers were without degrees when in fact they had them. Nonetheless, the erroneous figures are mentioned by Dr. Darling Hammond with zeal each time she discusses the report.) So Darling-Hammond's strongest critique of the lack of certification is groundless.

Darling Hammond ignores another study released by PPIC showing that the only teacher variable that influences student achievement is teacher experience. [Christopher Jepsen and Steven Rivkin. "Class Size Reduction, Teacher Quality, and Academic Achievement in California Public Elementary Schools," Public Policy Institute of California, June 2002. <http://www.ppic.org/main/publication.asp?i=155>] This oversight is surprising because she cites a companion report produced by the same firm, suggesting at least a minimum regard for the quality of research being produced at PPIC.

Failure# 2: Plaintiff's own witnesses and others provide counter examples that disprove the claim that the input factors at issue are essential.

Thomas Sobol, a leading expert for the plaintiff, actually refutes their claims in an article published in 1997. He writes

What do we see, when we stand back and survey the nation's schools? To begin with, we see many fine schools with many successful students and many hard-working, dedicated teachers, often accomplishing great things

against great odds. *Such schools flourish in our urban centers and in our rural areas as well as in our favored suburbs.* [Emphasis added.]

[Sobol, Thomas. Beyond Standards: The Rest of the Agenda. *Teachers College Record*, Volume 98, Number 4, Summer 1997. p. 630.

<http://www.tcrecord.org/Content.asp?ContentID=9600>]

He goes on to say that many schools have poor facilities and struggling teachers and outmoded textbooks. It bears mention, however, that Sobol also cites as causes of failure the lack of school safety, the lack of community, and the absence of a sense of purpose among students. In his testimony, Sobol lists 18 factors that contribute to a sound education, but does not show that the three at issue in this case are paramount. [Sobol, p.5] Even in his own litany of factors, it is impossible to know which are of primary importance. Nor have plaintiffs' experts provided convincing evidence that their preferred proposals are better than the programs and policies currently in force in California schools. We are asked to accept their professional judgment.

Regardless, if the three inputs at issue in this case were essential, then it would not be the case that students and schools could overcome the odds of not having them, and yet Sobol himself admits that it happens. By logical reasoning, therefore, one cannot accept the validity of the claim that these factors are essential.

There is no quibble that the three proposed solutions – sufficient textbooks, quality teachers and adequate facilities -- play a role in the production of good education. But the definitions of “sufficient”, “quality” and “adequate” are elusive and highly subjective. Moreover, it is a large leap to accept that these elements are only effective in the precise formulations advanced by the experts. Given the State and local resources that are already devoted to these matters, there needs to be a high standard of proof that these elements are essential to all schools and to all students in the same way. It is also the burden of the plaintiffs to prove that their proposals will in fact produce the results they claim and in a manner superior to the State's current approach. Russell admits this himself. [Russell deposition, p. 106, l. 25]

More importantly, plaintiffs' argument is proven specious on empirical grounds. In Table 1, the 2002 API scores for the 39 plaintiff schools are listed. The API scores range from 425

to 805 on a scale of 200 to 1000. Median scores for the State were 689 for elementary schools, 668 for middle schools, and 635 for high schools. [API Growth Results 2001 – 2002. <http://www.cde.ca.gov/news/releases2002/re132attach1.pdf>]. Based on plaintiffs' complaint, each of these schools has deficiencies in at least one of the three elements. Yet eight of the plaintiff schools – more than 20 percent of the total -- are above the state medians for their school type (elementary, middle or high school). This simply could not be possible if the plaintiffs' argument were valid. Instead, it supports the position advanced here – that schools use resources in different ways to produce different results. Even when confronted with operating challenges, there are schools who can realign their resources to compensate and to overcome the deficiencies. They should have the latitude and support to do so.

To further examine the validity of the positions advocated by the plaintiffs, data from the 2002 API datasets constructed by the California Department of Education were used to construct econometric models of educationally challenged schools in California. The results show that the benefit that would be realized by plaintiffs' proposal to increase the proportion of fully credentialed teachers is not as dramatic as their experts assert.

The analysis was conducted as follows. A sample of 584 schools in California was selected from the population of California public elementary, middle and high schools. The population of schools includes operating charter schools, but excludes alternative schools, continuation schools, adult schools and schools that did not operate throughout the 2001-2002 school year.

The criteria for selecting schools into the sample were constructed to select equivalently challenged schools. Across the plaintiff schools, the group average was constructed on three measures: 1) the percentage of the student body that was reported as minority; 2) the percentage of the student body that qualified for Free or Reduced Price meals; and 3) the percentage of the teachers at the school who held full teaching credentials. Any California school in the population who met or exceeded all the criteria (that is, *higher* proportions of minorities, *higher* proportions of Free / Reduced Price meals students and *lower* proportions of fully credentialed teachers) was selected into the analysis. Thus the resulting sample of

584 schools consists of all the plaintiff schools and comparison schools that look like them on key attributes. All the schools could be considered educationally challenged.

These schools were analyzed to determine the significance and magnitude of influence for fully credentialed teachers. Table 2 presents the results. This model explains the variation in 2002 API scores across schools as a function of student background and the availability of fully credentialed teachers. The model uses 565 schools for which all the required data was available. (If a school has a missing value for any of the variables in the model, the school observation is dropped.) The model shows three very important results.

First, the model is statistically significant overall, but not particularly powerful. The F statistic measures capability of the independent variables (student background and the availability of fully credentialed teachers) as a set to explain the variation in API scores. The value of the F statistic gives us confidence that the model is useful. However, another statistic, R^2 , measures the extent to which the model completely explains the variation in API scores. A perfect model would explain 100 percent of the variation in API scores and have an R^2 of 1. The model in Table 2 explains about one quarter of the observed variation in API scores. This finding illustrates one of the primary points of this report – there is a considerable amount of what transpires in schools that cannot be explained. It is difficult to capture the precise relationship between inputs and outcomes. This finding strikes a blow to the notion that specific formulae could be effective in all cases.

The second point concerns the discernable impact of teacher credentials. Table 2 also presents the values for the individual coefficients for each of the factors used to explain API score variation. The model coefficients measure the marginal effect of a one-unit change in the independent variables on the dependent variable. Using District Mobility as an example, if the mobility of students in and out of the district increases by 1 percent -- the unit of measure for mobility is a percent – then the API would be expected to fall by 4.66 points, all other things being constant.

The coefficient for the proportion of fully credentialed teachers is positive and significant. But the magnitude of the coefficient is quite revealing. For each percentage gain in the

proportion of credentialed teachers, we would expect to see only .58 points of gain in the API. This means that if a school was to increase its share of fully credentialed teachers by 10 percent – an ambitious achievement – it could expect only a 6 point gain in its API. That change is roughly equal to the measurement error associated with the API according to some researchers. [Kane, Thomas J. and Douglas O. Staiger, “Volatility in School Test Scores: Implications for Test-Based Accountability Systems” Brookings Papers on Education Policy 2002 (Washington DC: Brookings Institution, 2002), p. 247.

http://www.dartmouth.edu/~dstaiger/Papers/KaneStaiger_brookings2002.pdf]

The third result is even more striking. The model controls for differences in student background. The model also controls for differences in teacher force composition. After those effects are restricted, the plaintiff schools still appear to be performing at substantially lower levels of API compared to their equivalent peer schools. For reasons other than challenging student populations or the size of the credentialed workforce, plaintiff schools performed on average 32 points worse on the API in 2002 than schools that look just like them.

The model is not able to discern exactly what about the plaintiff schools makes them underperform relative to their peers. We can, however, say that the variance in the API scores themselves and the control of the teacher element as a possible factor point to operational differences across schools. These schools have similar available resources as a result of their profiles, so the way those resources are utilized has strong effect on their performance.

To test the plaintiffs’ argument further, the models were re-run substituting the proportion of emergency credentialed teachers for the proportion of fully credentialed teachers. These results appear in Table 3. The coefficient in question is -.19 and it is not statistically significant. This means that it is not possible to assert that changes in the share of emergency credentialed teachers will have an impact on the average test scores for a school, all other things being equal. It is impossible to square this empirically derived result with the position of the plaintiffs regarding the value of fully credentialed teachers and the potential detriment of using non-fully credentialed teachers.

There is an inherent contradiction in the position taken by plaintiff: they demand local flexibility but are advocating greater control by State education authorities.

Plaintiffs' experts advocate a proposal that is inherently contradictory. They seek to dictate input standards and hold districts accountable for them. They would remove local latitude for decision-making relating to their three inputs, but then insist that localities have differing needs.

The approach seeks to turn the current operating practice – common performance standards across districts and schools coupled with local discretion to adapt practice to meet specific local needs – on its head. The new regime would prescribe the way schools must operate and then allow flexibility in the accountability standards. Yet, the approach is intended to produce a high academic result for all students. This construct is inherently contradictory.

Consider the case that the plaintiffs desire: all schools are required to allocate resources the same way for the same things. By dictating inputs, we are assuming that local practice is equivalent; that is, that the contribution of the factor is constant across all schools. We would have to be confident that the resource is utilized the same in all settings. Using social studies textbooks as an example, schools with high academic achievement would be required to have the same textbook policies as those schools with low academic achievement, even if the schools used textbooks in fundamentally different ways. This would be true even if the high achieving school's students had mastered the State's social studies content standards better than anyone else in the state. One example of differing use might be that high achieving schools might use a textbook as a jumping off spot and augment it by bringing other resources into the classroom, or by integrating social studies, reading and writing into a blended lesson plan. Since we do not know the full contribution of inputs on average (i.e., the production function for education discussed above), we also do not know if the contributions differ by various school characteristics. The case by plaintiffs assumes we do.

This assumption cannot be squared with the strong case that plaintiffs advance that the needs of students differ across the state. Even if we do not agree on the causality for those

differences, we can agree that differences do exist. It cannot be the case that the same tools work the same ways on students with different educational assets. One could consider the feasibility of using genealogical databases to explore the settlers' expansion into California; some schools and teachers would readily incorporate these resources while others would not know where to begin.

Even plaintiffs' witnesses acknowledge the weakness of their proposal. In Sobol's 1997 article, he states, "If we have learned anything from the reform efforts of recent decades, it is that top-down, formulaic, 'teacher-proof' remedies do not work." [ibid, p. 631.] He further reinforced the point in a presentation at Columbia University Teachers College in late 2000. There, he stated

Decentralization has had some good effects, historically, I think. And it has [had] some unfortunate consequences. Without making that whole long speech – historically, it has helped us to – I think you could argue that it has kept schools closer to the people, in places outside the big cities at least, annual vote of the budget, election of local school board members and so on. Kept the public schools closer to the people that let's say the bureau of transportation, motor vehicles, about which most people don't have warm and fuzzy feelings. It's also that decentralized nature of the system that has made it very difficult, if not impossible, for partisan groups or individual demagogues to seize control of the public education system. [Sobol, Thomas. School Finance (Part2) *Teachers College Record*, 10/15/00, <http://www.tcrecord.org> ID number 10596.]

His position cannot be squared with the position plaintiffs are taking in this case. His report in this case even says so. [Sobol, p. 12, para. 24]

Plaintiffs' proposals are unworkable.

The plaintiffs have not considered the reality of a budget constraint – especially curious in the current fiscal environment. The costs of fully implementing their proposals would be exorbitant. Even before the K-12 education system of California was faced with the state budget deficit, plaintiffs' case did not consider whether the costs of implementing the proposals could be borne at all or could be better allocated to other uses. Perhaps one of the reasons plaintiffs took their case

to court in the first place is because the costs involved would doom their proposals in all other venues simply because they are outside the feasible budget for education. Other such examples exist: it might be ideal to have one teacher for each student, but the costs are more than the public can bear regardless of the potential benefits. There are myriad things the State could do to improve education if it had unlimited resources, but that scenario is a fantasy, the real question concerns what the State should do with its very limited resources.

Even if the costs of plaintiffs' various proposals were manageable, the question remains whether taking the funds that are required for this plan and putting them to other purposes would be preferable. Perhaps equivalent investments in early childhood education or stay-in-school incentives would produce greater results. Before the proposals from plaintiffs can be supported, the relative benefits from these alternatives should be clearly articulated. It should be the burden of plaintiff to do so.

However, plaintiffs' testimony offers no comparison with other uses. We are left to assume, just based on their assertions, that these proposals represent the highest and best use of education dollars. Returning to the models presented in this testimony, we have solid empirical evidence that in the area of teacher qualification, where data are available, the marginal impact of their proposal is much smaller than other potential investments.

The model in question appears in Table 4. This model is based on only 129 of the 584 schools in the comparison group because fewer schools provided information on the number of core academic courses. However, this number of schools is sufficient to support a comparison of different uses of education dollars. Recall that the coefficients in these models shows the marginal change that would be produced in the API for schools if there was a one unit change in the variable in question, all other things being equal. The variable "Percent of Fully Credentialed Teachers" is measured in percentage units. The variable "Number of Core Academic Courses" is measured in single course offerings. Table 4 shows a coefficient for "Percent of

Fully Credentialed Teachers” of .3622, meaning that a one percent increase in the proportion of fully credentialed teachers produces roughly one-third of a point of API gain. By contrast, the coefficient for “Number of Core Academic Courses” is 3.461, nearly 10 times greater. By increasing the number of core academic courses by one, we would expect the effect on API scores to be about 3.5 points. To create the same impact, a school could increase the proportion of fully credentialed teachers by 10 percentage points or add a core course.

There are no available data to test the impact of the remaining two proposals. However, from this single analysis the implications are straightforward: we cannot rely on the professional wisdom of the plaintiffs’ experts to choose the most effective ways to improve the educational attainment of California’s students.

In the present environment of cutbacks and program suspensions, many other programs that are underway would have to be terminated to make room for this unproven and open-ended enterprise. The proposal ignores the budget realities we all face. The fact is that the plaintiffs’ experts do not address issues of cost, nor identify sources of new funds required to carry out their proposals. They ignore the fiscal side of their arguments completely.

Further, plaintiffs’ proposals would disenfranchise parents. By claiming to know what is best for all students, plaintiffs are removing the option for parents to be co-creators of the education programs that best meet the needs of their children. Dictating rigid practices and requirements signals to parents that their role is at best secondary and that the education of their children is best left to the experts. It is difficult to imagine parents willingly giving up their involvement in so important an area of their children’s lives. So, as a policy position, the proposal is politically untenable.

The State of California has acted within its authority to create an accountability framework that is reasonable and appropriate. The result is a program that presents positive incentives for students and schools to improve their performance.

In 1999, the PSAA was enacted by the California State Legislature and signed by Governor Gray Davis. The Act introduced a new statewide performance measurement system – the California Academic Performance Index – that utilizes a consistent measure of academic performance across all schools in the state. While other components of the program such as remedial intervention and peer referent groups were adopted then and at a later date, the API is the keystone of the state school accountability system.

The API uses a computational formula to create a composite measure for schools based on the scores of their eligible students on mandatory annual examinations conducted as part of the state's STAR testing program. Students in grades 3-11 take STAR exams. Individual student scores are shared with schools, districts and parents, and aggregates of the scores are available by school and by populations of interest, which are determined according to federal and state guidelines.

The original legislation calls for incorporation of attendance rates and graduation rates into the API, should those measures be found to be consistent, reliable and universally reported. These measures have not yet been incorporated into the API, and the question of whether and when to incorporate these additional measures is the subject of ongoing discussion, analysis and debate. A report to the Secretary for Education in 2002 found these data elements currently to be unreliable and sketchy, but more importantly, showed that they diluted the focus of the API in unproductive ways. [Fletcher, Stephen and Margaret Raymond, "The Future of California's Academic Performance Index, Report Prepared by CREDO for the Secretary for Education of the Governor's Office, 2002.

<http://www.ose.ca.gov/whatsnew/CREDO.html>]. Specifically, the measures currently lack standard definitions, and are collected differently in different schools. The resulting figures are too ambiguous to rely upon. While changes in the definitions have been developed this year by the staff at CDE, local implementation of the new definitions has not occurred universally.

By using only test scores as the basis of school API scores, the state legislature has an accountability system that is focused exclusively on outcomes. This focus is consistent with the legislation, which clearly specified that the primary intent of the law was to measure the academic progress of children in California schools. [California Education Code Section 52050.5(j)] The State Legislature was clearly within its statutory authority to decide what kind of accountability policy was best for the State, even if plaintiffs don't agree with it.

This is not to suggest that the State ignores inputs, however. The Education Code is rife with statutes that specify such things as earthquake safety [Section 17282]; increasing the supply of highly qualified teachers of mathematics [Section 44400] or establishing nutritional guidelines for school meals [Section 49531]. However, in setting the parameters for accountability, the State has chosen a course focused on outcomes. Outcomes are the final measure of the success of schools in preparing their students, and they are the basis on which entry into employment, colleges, and other post-secondary avenues are judged.

Mintrop's definition of accountability includes the assertion that "accountability" is bi-directional; that is, that the State is held accountable by communities and citizens for adequate and equitable provision of education. This is incorrect. "Good" accountability is focused where responsibility is located, and since the majority of influence in decisions concerning education occurs at the local level, it is completely appropriate that the accountability system be focused there. Further, it is incorrect to think of executive agencies as entities to hold responsible. The constitutional provisions of the State place the executive agencies under the direct supervision of the elected governor. The governor and the elected members of the State legislature craft the policies that the California Department of Education is charged to carry out. Our electoral process is designed to provide the ultimate form of accountability – job loss for poor performers.

Since the system is but three years old, with the additional elements of the II/USP and alternative accountability programs even newer, it is too soon to judge the full impact of the State policy. In fact, in April of 2002, the CDE delivered to the Board of Education a six year plan which placed full implementation of the accountability program at the end of 2003.

[Source: The Academic Performance Index (API): A Six Year Plan for Development (2001-2006) http://www.cde.ca.gov/psaa/api/Six_Yr_Plan.pdf] Until the full spectrum of the program has had a chance to operate, it would be premature to judge its impact.

Russell provides a contradictory criticism of California's accountability program. On the one hand, he is critical of the phase out of the school and teacher rewards, but immediately turns around and criticizes the system again because it is unstable and constantly changing. The long implementation period appears lost on plaintiffs' experts.

There is reason to expect that the accountability program will be effective in motivating schools to improve school performance. Research completed in 2002 by Hanushek and Raymond shows that states with accountability systems did produce greater gains on the National Assessment of Educational Progress (NAEP) than states that lacked them. [Hanushek, Eric and Margaret Raymond, 2003b. "Lessons about the Design of State Accountability Systems." In *No Child Left Behind? The Politics and Practice of Accountability*, edited by Paul E. Peterson and Martin R. West. Washington, DC: Brookings Institution Press. <http://www.ksg.harvard.edu/pepg/>]. That study involved a statistical analysis of the changes in NAEP scores, comparing those rates of change for states that adopted accountability systems to those that did not. The results showed that accountability counts. This rigorous empirical analysis trumps the contradictory study cited by Russell in his deposition and that was quoted widely in the *New York Times* and elsewhere. In earlier work done by researchers at Arizona State University, the claim was made that accountability systems had no effect or even impaired educational progress. However, as fully explained in a recent article, the methodology followed in the ASU study lacks any semblance of acceptable research design and is disturbingly overreaching in its conclusions. [Amrein, Audrey L. and David C. Berliner. 2002. "The impact of high-stakes tests on student academic performance: An analysis of NAEP results in states with high-stakes tests and ACT, SAT, and AP test results in states with high school graduation exams. Tempe, Arizona: Educational Policy Research Unit, College of Education, Arizona State University. http://www.asu.edu/educ/epsl/EPRU/epru_2002_Research_Writing.htm]

An outcomes based accountability system achieves the conditions that plaintiff asserts should be met— a) it offers localities more latitude than any alternative, b) it correctly focuses on the true objective of education policy, that of educational achievement of students, and c) it adopts a common yardstick by which to gauge the effectiveness of schools.

Expansion of the API would damage the current system with no realizable benefits. Adding inputs and process measures would dilute the incentives to focus on outcomes, allow schools to game their results without true gains. Imagine including an input factor in the API; say, the number of computers per 100 students. Further imagine equal weights in the final API computation for the test score results and the number of computers per 100 students. Of course, other weights are possible, but having them equal illustrates the problem most clearly. The final API score could be affected by schools doing two things: raising the average test scores for the school and/or increasing the number of computers per 100 students. Given the comparative low cost and ease of computer purchases over the alternative, the path of least resistance is clear.

Nor would the API be enhanced by addition of course grades, a popular fall back recommendation. The problems of grade inflation and manipulation are well known in the post-secondary levels of education. [Rosovsky, Henry and Matthew Hartley. *Evaluation and the Academy: Are We Doing the Right Thing?* American Academy of Arts and Sciences, 2002.

http://www.amacad.org/publications/monographs/Evaluation_and_the_Academy.pdf.]

Its prevalence is only recently been discussed in elementary and secondary schools. Use of grades would completely defeat the use of standardized test scores that are intended to eliminate such subjective and variable practices.

Williams v. California
 Tables for Report by Margaret E. Raymond

Table 1 2002 API Scores for Plaintiff Schools

555	Garfield Year-Round Elementary
472 ^a	Stonehurst Elementary
566	Burbank Elementary
567	Berry Creek Elementary
425	Kennedy High
485	Helms Middle
570	Dailey Elementary
805	Brightwood Elementary
726	Mark Keppel High
732	Freeman (Daniel) Elementary
660	Parent (Frank D.) Elementary
714	Robinson (Jackie) Elementary
489	Belmont Senior High
474	Crenshaw Senior High
442 ^b	Dorsey (Susan Miller) Senior H
452	Fremont (John C.) Senior High
519	Huntington Park Senior High
439	Jefferson (Thomas) Senior High
728	Cahuenga Elementary
600	Gulf Avenue Elementary
713	Rosewood Avenue Elementary
530	Belvedere Middle
432	Carver (George Washington) Middle
480	Harte (Bret) Preparatory Intermediate
590	Marina del Rey Middle
530	Webster (Daniel) Middle
497	Lynwood Middle
582	Gascon (Joseph A.) Elementary
554	Tenaya Middle
573	Bryant Elementary
558	Luther Burbank Middle
516	Edison-McNair Academy
499	Chavez (Cesar) Academy (Elem)
800	Castlemont Elementary
488	Watsonville High

Williams v. California
Tables for Report by Margaret E. Raymond

633	Cloverdale High
643	Mt. Whitney High
621	Redwood High
463	Balboa High School

Notes a. School did not have API scores for 2002, 2001 or 2000. The 1999 API score was substituted.
b. School did not have API score for 2002. The 2001 figure was substituted.

Williams v. California
 Tables for Report by Margaret E. Raymond

Table 2 Econometric Model Results

Source	SS	df	MS	Number of obs = 565
Model	553579.949	7	79082.8499	F (7, 557) = 24.35
Residual	1808819.21	557	3247.43126	Prob > F = 0.0000
				R-squared = 0.2343
				Adj R-squared = 0.2247
				Root MSE = 56.986
Total	2362399.16	564	4188.65099	

API 2002	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
% Minority	-1.342271	.445087	-3.02	0.003	-2.216525 - .4680164
% Fully Cred. Teachers	.5840205	.1801964	3.24	0.001	.230073 .937968
% Free/Reduced Lunch	-.1761775	.2442243	-0.72	0.471	-.6558908 .3035357
% English Learners	.3597235	.167129	2.15	0.032	.0314434 .6880036
% Student Mobility	-4.656176	.7570641	-6.15	0.000	-6.143226 -3.169126
% Parent Not HS Grad	-1.423336	.1745109	-8.16	0.000	-1.766116 -1.080556
Plaintiff School	-32.6192	11.44292	-2.85	0.005	-55.09574 -10.14265
Constant	760.3324	47.12043	16.14	0.000	667.7769 852.8878

Williams v. California
 Tables for Report by Margaret E. Raymond

Table 3 Econometric Model Results

Source	SS	df	MS	Number of obs =
Model	523127.28	7	74732.4685	565
Residual	1839271.88	557	3302.10392	F(7, 557) = 22.63
Total	2362399.16	564	4188.65099	Prob > F = 0.0000
				R-squared = 0.2214
				Adj R-squared = 0.2117
				Root MSE = 57.464

API 2002	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
% Minority Students	-1.59194	.442296	-3.60	0.000	-2.460712 - .7231679
% Emer.Cred. Teachers	-.19407	.18436	-1.05	0.293	-.5561955 .1680561
% Free/Reduced Lunch	-.23460	.2456393	-0.96	0.340	-.717091 .2478943
% English Learners	.34867	.1686328	2.07	0.039	.0174338 .6799015
% Student Mobility	-5.15493	.7465237	-6.91	0.000	-6.621278 -3.688587
% Parents Not HS Grad	-1.37349	.1753883	-7.83	0.000	-1.717989 -1.028983
Plaintiff School	-32.50658	11.54817	-2.81	0.005	-55.18987 -9.823291
Constant	833.94990	41.25944	20.21	0.000	752.9068 914.993

Williams v. California
 Tables for Report by Margaret E. Raymond

Table 4 Comparison of Different Investments in Schools

Source	Sum of Squares	df	MS	Number of obs =	129
Model	124714.563	8	15589.3204	F(8, 120) =	4.42
Residual	422835.918	120	3523.63265	Prob > F =	0.0001
				R-squared =	0.2278
				Adj R-squared =	0.1763
				Root MSE =	59.36
Total	547550.481	128	4277.73813		

API 2002	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
% Minority Students	-.7829225	.6114771	-1.28	0.203	-1.993605 .4277595
% Fully Credentialed	.3621691	.4052734	0.89	0.373	-.4402439 1.164582
% Students F/RP Meal	-1.005723	.364449	-2.76	0.007	-1.727307 -.2841397
% English Learners	-.363981	.3443322	-1.06	0.293	-1.045735 .3177727
% Student Mobility	-3.905788	1.795814	-2.17	0.032	-7.461376 -.3502009
Year Round School	-25.8066	12.43904	-2.07	0.040	-50.43504 -1.17816
# of Core Courses	3.460734	1.577083	2.19	0.030	.3382196 6.583248
Plaintiff School	-20.82156	15.92712	-1.31	0.194	-52.35615 10.71302
Constant	630.7556	88.43651	7.13	0.000	455.6575 805.8538

received
 APR 1 1993
 Bits Received