EDUCATION POLICY ANALYSIS ARCHIVES

A peer-reviewed scholarly journal
Editor: Gene V Glass
College of Education
Arizona State University

Copyright is retained by the first or sole author, who grants right of first publication to the **EDUCATION POLICY ANALYSIS ARCHIVES**. **EPAA** is a project of the Education Policy Studies Laboratory.

Articles appearing in **EPAA** are abstracted in the *Current Index to Journals in Education* by the ERIC Clearinghouse on Assessment and Evaluation and are permanently archived in *Resources in Education*.

PDF file of this article available:

This article has been retrieved 12 times since March 5, 2004

Volume 12 Number 10 **March 5, 2004** **ISSN 1068-2341**

Increasing Equity and Increasing School Performance—Conflicting or Compatible Goals?: Addressing the Issues in *Williams v. State of California*

Jeanne M. Powers
Arizona State University

Citation: Powers, J. M. (2004, March 5). Increasing Equity and Increasing School Performance — Conflicting or Compatible Goals?: Addressing the Issues in *Williams v. State of California. Education Policy Analysis Archives, 12*(10). Retrieved [Date] from http://epaa.asu.edu/epaa/v12n10/.

Abstract
This work addresses some of the arguments regarding equity in public education versus school performance at issue in the case of Williams v. State of California. The plaintiff's expert witnesses have argued that the state is responsible to reduce the inequities in California's public educational system. In contrast, the state's witnesses argue that some of the plaintiff's proposals have limited educational effects at the cost of reducing local autonomy. In this paper, I use four years of data from California's Public Schools Accountability Act (PSAA) to evaluate these claims.

Introduction

On May 17, 2000, the 46th anniversary of the decision in *Brown v. Board of Education* outlawing racial segregation in public schools, a class action lawsuit was filed on behalf of California's public school students in an effort to make the state address some of the inequities in California's public educational system (Purdum, 2000). Represented by the ACLU and civil rights organizations, the plaintiffs allege that the state is responsible for ensuring that all public school children across the state have the right to experience the same quality of textbooks, teachers, and classrooms. The plaintiff's experts have documented a range of inequities in California's public educational system and have further argued that these inequities are fundamentally unfair given the high stakes accountability program California initiated in 1999 with the Public Schools Accountability Act (PSAA). (Note 1) This legislation mandated the ranking of all California public schools based on their Academic Performance Index (API), which has been calculated based on the results from the state-mandated tests administered to students in grades 2 through 11 in the spring of each school year. By the winter of 2003, the fifth year of state school rankings will be released. Over three years later, the case of *Williams v. State of California* is still working its way through the court system; the trial is set to begin August 30, 2004. (Note 2)

In this article, I evaluate the claims made in one expert report written by Margaret Raymond of the Hoover Institution's Center for Research on Educational Outcomes (CREDO) on behalf of the state. (Note 3) In her report, Raymond utilizes the data generated in the wake of the PSAA to rebut the plaintiff's claims. There are four main sections to the article. First, I briefly outline some of the major claims made by Raymond, focusing specifically on her analysis of API data. Second, I describe the API and the other sources of data in the analysis. In this section I also discuss some problems with the data and the strategies I used to address these problems in my analysis. I also present the results of my efforts to recreate Raymond's analysis. In the third and fourth sections, I provide analyses of the API data addressing the issues of teacher qualifications and facilities, respectively.

Key Claims

In her report rebutting the plaintiff's arguments in the case of Williams v. State of California, Margaret Raymond (2003a) argues that the plaintiffs "haven't developed a reliable production function for education that highlights the factors at issue in this case" (6). In her discussion she focuses on three in particular, the quality of the teaching staff, facilities, and instructional materials. In her analysis she specifically focuses on the effect of teacher credentials on school performance as measured by the Academic Performance Index (API). In part, this analytical strategy is based on the availability of data. To my knowledge no state level datasets exist which provide information about facilities and instructional materials.

In elaborating her point about appropriate research strategies that would provide evidence supporting the plaintiffs' claims, Raymond further argues that "[t]o be confident that the plaintiff's claims have merit, it would be necessary to study the effects of each of their proposals under 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" (6). Yet interestingly, she doesn't appear to follow this strategy in her own analysis of school API scores, which she uses to draw conclusions about the influence of credentialed teachers on school performance. Moreover, the evidence for her assertions is the results from an inappropriate analysis of API data.

More specifically, Raymond argues that in "educationally challenged" schools, teacher credentials do not influence school performance as measured by the Academic Performance Index (API). However, one of her criteria for defining educationally challenged is the percentage of fully credentialed teachers at a school. If the percentage of fully credentialed teachers in a school is below the mean percentage of fully credentialed teachers of the approximately 39 Williams schools in her analysis *and* it also has greater percentages of free and reduced lunch students and minority students than the mean of the Williams' schools on these variables, then the school is defined as educationally challenged (Raymond, 2003a, 12). Using this sample of schools, Raymond constructs an econometric model which she argues shows that teachers credentials – measured by the percentage of teachers at a school that are fully credentialed – have relatively little effect on school API (Raymond, 2003a, 13). She further argues that the variables in her regression models explain relatively little of the variation in school API scores.

Before accepting these conclusions, it is important to examine the research strategy employed in this analysis more closely. One of the criteria for inclusion in the group "educationally deprived" is that there must be a relatively low percentage of fully credentialed teachers at a given school. Thus, Raymond uses a sample of schools in which there is little variation in the availability of fully credentialed teachers to construct what is basically a tautological argument. If the schools in her sample are relatively similarly situated in terms of the percentage of fully credentialed teachers on staff, then it is not surprising that her regression analysis suggests that teachers' credentials don't matter. This method of sample selection also explains why she obtains what she describes as a very low R-square for her regression models (13). In a sample of schools with little variation on the key explanatory variables, it is not surprising to find they have a relatively small effect on the model when you test for the influence of these variables on a dependent variable. It is well known that restriction in the variability of a variable attenuates that variable's correlation with any other variable. (Glass & Hopkins, 1996, 121-3) As I will also detail below, it is also questionable whether or not the comparison sample Raymond utilizes in her analysis is an appropriate comparison for the Williams schools.

A Brief Overview of the API

The Academic Performance Index (API) is a state-constructed measure of school performance mandated by the 1999 Public Schools Accountability Act. From 1999 to 2001, a summary score for each school was constructed by weighting student scores in each content area of the SAT-9 tests administered to students in grades 2 through 11 by their national percentile ranking (NPR) and then weighting each content area to create an overall score (California Department of Education Office of Policy and Evaluation, 2000, p. 9). For the calculation of the API for elementary and middle schools, the content areas were weighted in the following manner: mathematics 40%, reading 30%, language 15% and spelling 15%. For high schools, the following content areas were each weighted 20%: Reading, Mathematics, Language, Science, and Social Science. In 2001, the results from the California Standards Tests (CST) in Language Arts were incorporated into a 2001 Base API. In

addition, a 2001 Growth API using only the SAT-9 test results was also calculated. The 2002 Growth API was calculated using the formula for the 2001 Base API. The 2002 Base API utilized in Raymond's analysis incorporates the CST in Math for all grades and the History and Social Science tests for grades 10 and 11 as well as the results of the High School Exit Examination (HSEE). In the main part of the analysis I utilized the API scores that are the most comparable across the four years – 1999-2001 API scores calculated using SAT 9 results only and the 2002 "Growth" API, which incorporates the results from the California Standards Test (CST) in Language Arts only but otherwise is calculated from SAT 9 test scores. In the section of the analysis reproducing the Raymond analysis, I used the 2002 Base API to ensure relative consistency of results.

In addition to the API rankings, the API datasets made available by the California Department of Education contain additional variables measuring various types of school characteristics. Three main categories of variables are utilized in the analyses here: 1) variables related to students' background characteristics; 2) variables related to teacher characteristics; and 3) variables indicating the type of calendar the school follows. I discuss each type of variable in turn in the sections that follow.

Student Background Characteristics

% Reduced/Free Lunch is measured by the percentage of students eligible for reduced or free lunch. *Mobility* is a state-constructed variable that provides a measure of the transiency of the student population by indicating the percentage of test-taking students who first attended the school within the current school year. Students who first attended the district a given academic year were excluded from that year's API calculations. *% English Learners* denotes the percentage of students school-wide reported as English learners. Additional variables denote the percentages of students belonging to one of 7 racial groups at each school: African American, American Indian, Hispanic, Asian American, Filipino, Pacific Islander, and White. One additional student background variable available in the API data that Raymond utilized in her also requires more in-depth discussion. The variable for percentage of parents not high school graduates is one of a series of variables measuring the percentage of parents at the school that have reached a given educational transition (e.g. high school graduates, some college, etc.). More importantly for the discussion here, however is the variable also available in the 2002 Base API data indicating the percentage of parents responding to this question at the school, *% Response for Parent Education*.

Teacher Characteristics

The API data also contains variables denoting the percentage of teachers at a school holding full and emergency credentials. According to the CDE website, in the API datasets, it is possible for one teacher to be in both the fully credentialed and emergency credentialed categories. As a result, for some schools, the total of the percentages for "Fully Credentialed" and "Emergency Credentialed" may exceed 100. Another issue not addressed by the CDE in their discussion of the API data is the problem of missing information; for some schools the percentages of fully credentialed and emergency credentials add to *less* than 100. In order to more precisely assess the credentialing at schools, I used variables drawn from the California Basic Educational Data System (CBEDS) Professional Assignment

Information Form. These files, which contained records for approximately 325,000 teachers across the state, were aggregated by school and matched to the API data using the unique code for each school.

Fully Credentialed indicates the percentage of teachers who have completed a teacher preparation program and hold a preliminary, clear professional, or life credential. *Emergency Credentialed* indicates the percentage of teachers that hold an emergency credential. Emergency credentials are granted to individuals who are not qualified for a credential or internship but meet minimum certification requirements. These minimum requirements include: a passing score on the state's basic skills exam (CBEST); a bachelor's degree; and 10 semesters of college coursework in any four of the following areas -- language studies, literature, history, social science, mathematics, science, humanities, art, physical education, and human development (Darling Hammond, 2002.). In addition, teachers working on emergency permits must submit a statement indicating their intent to complete the credentialing requirements. Some teachers are designated as having a full credential AND an emergency credential. This group of teachers could include teachers who are credentialed in one field but teaching out of field, or teachers that are credentialed in another state and working on California state certification (Darling Hammond, 2002.). In the case of the latter, teachers are counted in a third variable, *Both Full and Emergency*. An additional variable was used to indicate whether the teacher's credential information was missing.

I also added a variable to the analysis that is also available in the Professional Assignment Information Form. *Years Teaching* counts the average total years of educational service among the teaching staff as teachers in any district, state or country. This figure includes teaching in private school settings but does not include any years teaching as a substitute teacher or in classified staff positions. Like the credential variable described above, an additional variable indicated whether or not the teacher's information was missing.

School Calendar

Indicator variables for the type of year round school were created from the CBEDS School Information Form Sections G through K. *Traditional* indicates that the school follows a traditional educational calendar with an extended summer vacation. *Year Round Single-Track* indicates that the school operates on a single-track year-round calendar with more frequent and shorter vacation periods (usually three a year ranging from three to five weeks in duration). The major change from the traditional calendar for the year round single-track calendar is the timing and duration of instructional and vacation periods; all of the staff and students are in school or in session at the same time (California Department of Education, 2000). *Year Round Multiple-Track* indicates that the school follows a year-round calendar where the students and faculty are divided into three to five groups that rotate throughout the year. This schedule is used to maximize enrollment at the facility; as one group of students and staff go on vacation, another returns for instruction. *Year Round Multiple- Track "Concept 6"* is a specific type of year round multiple-track calendar in which students have fewer instructional days than the other types of school calendars; instead the school day is lengthened so students receive the same number of instructional minutes as the other calendars. Whereas the other types of school calendars have 180 instructional days in the school year, "Concept 6" schools have 163 (Oakes, 2002). As a result, I distinguish these schools from the

other four types of year round multi-track calendars. The advantage of the "Concept 6" instructional calendar is that it allows schools to enroll 50% more students that it would be able to handle at the facility if it were to follow the traditional calendar (California Department of Education, 2001). In contrast, the other types of year round multiple-track calendars allow schools to increase their enrollment by 33% compared to the traditional calendar. An additional variable indicated if calendar information was missing.

Control Variables

Indicator variables denoting school type (elementary, middle, high) were created using the school type variable provided in the API dataset and included in the revised models. As Raymond notes, the median score for the state as a whole varied considerably by school type with elementary schools having the highest median scores followed by middle and then high schools. In a critique of another similar analysis of API scores, Rogosa (2002) argues that given these differences it is important to control for school type in these and similar analyses as school type might serve as a proxy for other unmeasured factors (23). Surprisingly, Raymond (2003b) makes a similar point in her analysis of the API scores of California charter schools and uses analytical strategies that take school type into account. However, she does not appear to control for school type in the analyses provided in the report under discussion here.

Reproducing the Raymond Analysis

In this section, I reproduce the Raymond analysis based on the information provided in her report. First, I detail the method of selection. Next I provide the descriptive statistics for the comparison sample and a discussion of how this group compares with the Williams schools in her analysis and the statewide sample. Finally, I recreate her regression analysis and also provide an alternative model with the corrected variables described above.

Selection Method

First, I selected out the 39 schools Raymond listed in Table 1 of her report from the 2002 API Base data. Of these, 36 had 2002 API scores and complete information on all variables. (Note 4) I calculated descriptive statistics on the three measures she used to select her sample of 584 schools: 1) the percentage minority students; 2) the percentage of students that qualified for reduced or free lunch; and 3) the percentage of teachers at the school who held full teaching credentials. Table 1 below provides the descriptive statistics for these three variables for the 36 schools:

Table 1
Descriptive Statistics for Raymond's Sample Selection Variables

	Minimum	Maximum	Mean	Std. Dev.
2002 Base API	425	805	575.56	104.01

% Reduced/Free Lunch	24	100	68.47	22.31
% Fully Cred.	27	100	73.61	17.05
% Minority	22	100	74.78	25.40

Raymond did not indicate which of the 6 possible non-white racial groups she included in the variables she calculated for the percentage of minority students. This variable is not included in the original 2002 Base API data and thus must be calculated from the variables indicating the percentages of students in each of 7 racial categories: African American, American Indian, Hispanic, Asian American, Filipino, Pacific Islander, and White. In the analysis described here, I defined "Percentage Minority" as the percentage of African American, Hispanic, and American Indian students at a school and the results roughly parallel her analysis.

According to Raymond, she used these sample means to select her cases, by choosing all of the cases *above* the sample mean for percentage of students on reduced/free lunch and percent minority students, and *below* the sample mean for percentage of fully credentialed teachers (Raymond, 2003, 12). In Raymond's analysis, this yielded a sample of 584 schools, 565 of which had information for all variables in the analysis. Using the means listed in Table 1 above to recreate Raymond's sample, I initially selected 593 cases. Of these, 574 had complete information on all variables. In addition, I also corrected and augmented the data using other datasets readily available from the CDE website per the discussion above.

Finally, before turning to a discussion of specific variables, it should be noted that there are 7444 schools in the 2002 API Base data with API scores. Of these, 7225 have complete information on all of the variables in Raymond's analysis. Using either of these figures, Raymond's sample of 565 schools is less than 10% of the schools assigned 2002 Base API scores in the state.

Table 2 provides descriptive information on all variables in the analysis for three groups: 1) a statewide sample of schools with 2002 API scores and information on all variables in the analysis (Column 1); the 35 "Williams" schools with 2002 API scores and information on all variables I used to recreate the Raymond analysis (Column 2); and the "comparison" schools chosen by following Raymond's inclusion criteria (Column 3).

Table 2
Descriptive Statistics for All Variables

	State Sample N=7225 Mean (S.D.)	**Raymond's Williams Sample N=35 (Note 5) Mean (S.D.)**	**Comparison Sample Using Raymond's Criteria N=547 Mean (S.D.)**	**Full Comparison Sample N=582 Mean (S.D.)**
2002 "Base" API	688.69 (105.37)	575.83 (105.52)	568.27 (61.36)	568.72 (64.75)

Elementary School	.71 (.45)	.43 (.50)	.78 (.41)***	.76 (.43)
Middle School	.16 (.37)	.26 (.44)	.18 (.38)	.19 (.39)
High School	.13 (.33)	.31 (.47)	.03 (.18)***	.05 (.22)
% Minority	48.73 (29.76)	74.23 (25.56)	92.48 (6.62)***	91.38 (9.91)
% Reduced/Free Lunch	48.58 (30.96)	68.14 (22.55)	90.20 (9.53)***	88.88 (11.94)
% Fully Cred.	86.83 (13.28)	73.16 (15.53)	60.83 (11.64)***	61.57 (12.25)
% Emergency Cred.	7.26 (9.52)	14.82 (12.63)	23.53 (13.21)***	23.05 (13.38)
% Both Full and Emer.	1.32 (3.64)	1.69 (2.64)	1.27 (2.02)	1.29 (2.06)
% Credential Missing	4.59 (6.64)	10.33 (6.65)	14.32 (9.08)*	14.08 (9.00)
% English Learner	23.64 (21.85)	36.29 (21.79)	52.28 (19.94)***	51.32 (20.40)
School Mobility	18.46 (14.10)	18.57 (9.29)	19.04 (14.85)	19.01 (14.57)
% Parents Not H.S. Graduates	20.16 (18.91)	32.69 (17.94)	43.17 (17.21)***	42.54 (17.42)
% Response for Parent Education	76.62 (24.48)	71.51 (24.18)	63.59 (29.14)	64.07 (28.91)

***t-test comparing sample means for Columns 2 and 3 statistically significant at p= .001
**t-test comparing sample means for Columns 2 and 3 statistically significant at p= .01
*t-test comparing sample means for Columns 2 and 3 statistically significant at p= .05

What should be immediately evident if you compare across columns are the differences in means for the three groups of schools shown in the first three columns. These three groups of schools are very different. Compared to the state as a whole, the 35 Williams schools included in the Raymond analysis are disadvantaged across all variables. However, Raymond's comparison group (Column 3) is much more disadvantaged than the Williams schools in her analysis. Comparison with the results presented in Table 1 suggest why this is the case. Raymond used the group mean for the Williams schools on the three selection variables. In addition, for a school to be included in the analysis it had to fit *all three* of the selection criteria rather than any one of the three criteria. This has the effect of selecting a more disadvantaged group overall for comparison by definition because it excludes schools which are comparable to the schools that are lower than the Williams schools' group means on the selection variables (in other words, the

relatively advantaged among the Williams schools). If we look at the standard deviations for the three groups, we can also see that there is relatively less variation in the comparison group for most of the explanatory variables than the other two groups. Because Raymond's Williams Schools group is so small (N=35), when it is combined with the comparison group as shown in column 4, it has a minimal effect on the means for the full comparison sample. To confirm this, I conducted t-tests on the sample means shown in Columns 2 and 3. The asterisks in Column 3 indicate that most of the differences in means between Raymond's Williams schools and her comparison group are statistically significant, which suggests that the schools are not appropriate comparison groups as Raymond contends.

Another striking difference between the Williams schools and the other two groups is that while the comparison group resembles the statewide sample in the distribution of school types (elementary, middle, high), almost half of Raymond's Williams schools are middle and high schools. Finally, what should also be noticeable from Table 2 is that in Raymond's group of Williams schools and the comparison group, there is a good deal of missing information in the two series of variables for teachers' credentials and parental education variables. When the teacher credential variables are corrected using the Professional Assignment Information Form as detailed above, higher poverty schools are more likely to have missing information compared to the state as a whole. We see this in both the Williams and the Raymond comparison groups, which on average are missing information on about 10 percent and 14 percent of their teachers' credentialing data, respectively. Similarly, in the average school in the state sample just under 25% of the parent education information is missing. However for the two comparison groups this figure increases to 28.5% for the Williams schools and 36% for Raymond's comparison group.

Regression Analyses

In Table 3, I present the results of initial regression analyses. In the first column I provide the reanalysis using Raymond's model in Table 2. In the columns that follow I provide models for the statewide sample (Column 2) and the comparison sample (Column 3). In both models I add the variables described above that correct for missing information in the credential and parent education variables. One of the problems with the statewide model shown in Column 2 is collinearity. Collinearity occurs when the two or more of the independent or predictor variables are highly correlated with one another. I address the issue of collinearity in more detail in the following section where I discuss the issue of the impact of credentials and other teacher characteristics on the model. However, I include this model here to illustrate the dramatic increase in the R-square for the statistical model once the sample size is increased. In this case, the R-square of .78 in the corrected model using the statewide sample (middle column) indicates that these variables explain about 78% of the variation in school API scores.

Table 3
Regression Models Following Raymond's analysis with Statewide Sample and Reconstructed Sample

	Raymond's	**Corrected Model**	**Corrected Model**

	Model in Table 2	Using Statewide Sample	Using Comparison Sample
Constant	694.98 (33.86)***	853.41 (7.35)***	790.06 (31.63)***
Middle School		-45.11 (1.66)***	-63.58 (5.98)***
High School		-107.58 (1.92)***	-92.87 (11.11)***
% Minority	-1.65 (.31)***	-1.01 (.04)***	-1.51 (.27)***
% Reduced/Free Lunch	.64 (.27)*	-1.53 (.04)***	.04 (.25)
% Fully Cred.	.35 (.21)	.12 (.07)	-.09 (.19)
% Both Full and Emer.		.26 (.17)	-.18 (1.06)
% Credential Missing		-.37 (.12)**	-1.28 (.25)***
% English Learners	.31 (.16)	.17 (.05)***	-.22 (.15)
School Mobility	-.35 (.17)*	-1.01 (.04)***	-.48 (.15)**
% Parents Not H.S. Graduates	-1.47 (.17)***	-1.02 (.06)***	-1.11 (.16)***
% Response -- Parent Education		.08 (.03)**	.34 (.08)***
Williams School	-23.26 (11.94)*	-18.07 (8.41)*	-10.33 (10.62)
R-squared	.18***	.78***	.39***

What we see from the comparison of the three models above is that the low R-square reported by Raymond and the even lower R-square yielded in my replicate analysis is due to two factors: 1) the criteria for selecting the sample as discussed above which reduces the variation within Raymond's sample on most of the explanatory variables; and 2) the omission of important control variables for school type and missing information in the credential and parent information variables, which are among the most statistically significant variables in the analysis. It is also worth noting that the coefficient for the Williams schools, while still negative, has decreased considerably with the corrected model (Column 3) and is no longer statistically significant.

Is Teacher Certification and Other Types of Teacher Training Important?

On the basis of these findings, Raymond argues that teacher certification has a negligible effect on school performance. In addition, she also cites an earlier study she conducted in the Houston Unified School District in which she argued that on average Teach for America teachers did as well or better than their peers (10). However, a more recent quasi-experimental study of Teach for America teachers in

5 districts in Arizona (Laczko-Kerr and Berliner, 2002) which matched Teach for America and other under-certified teachers with fully certified peers found 1) little differences in student performance between Teach for America teachers and other under-certified teachers (i.e. emergency credentialed teachers); and 2) students taught by fully credentialed teachers outperformed students taught by under-certified teachers. More specifically, their results indicated that students taught by fully-certified teachers gained approximately two additional months per academic year across subjects than students taught by under-certified teachers.

Similarly, in an analysis of 1999 and 2001 API scores for elementary schools in Los Angeles and San Diego, I found that the impact of teacher credentials versus teacher experience varies based on district context (Powers, 2003). For example, in Los Angeles, if we look at variables related to teachers' credentials and training, the main disparity between high poverty schools and low poverty schools is the percentage of emergency credentialed teachers. In contrast, in San Diego, where there is a relatively even distribution of emergency credentialed teachers across the district, the main disparity between high poverty and low poverty schools is the average years of teaching experience among the teaching staff. Not surprisingly, these differences are reflected in the results of regression models run separately for each district predicting the influence of student and teacher characteristics on school achievement as measured by the API. However, a more overarching conclusion that we can draw from this type of analysis is that these types of disparities across schools do matter, and that while public policies meant to address inequities might have to be tailored to the specific features of the district context, they are not inconsequential.

In the section below, I provide a reanalysis of the API data, adding the variable for teacher experience described above to the analysis. However, instead of using Raymond's list to select schools, I use a list of current Williams schools provided by the plaintiff's lawyers. I also utilize an alternative method of analysis. Rather than examine API at one point in time, as in the Raymond analysis, I examined the relationship between the API and the factors of interest – school demographic variables and teacher qualifications – for the group of Williams schools and the statewide sample with complete information on all variables from 1999-2002. (Note 6) This yielded 29 of the 34 Williams schools, and 6452 cases for the statewide sample. According to my calculations, this group of 6452 schools is approximately 83% of all the schools eligible for a 1999 API. (Note 7) The advantage of this strategy is that by using the same sample of schools over all four years of the analysis, we can examine changes in the explanatory variables over time. In addition, using a statewide sample also addresses Hoxby's (2003) concern that arguments for increased equity rely on data that is "representative of California public schools in general" (2). (Note 8)

As noted above, this analysis also utilizes the API scores that are the most comparable across the four years – 1999-2001 API scores calculated using SAT 9 results only and the 2002 Growth API, which incorporates the results from the California Standards Test (CST) in Language Arts only but otherwise is calculated from SAT 9 test scores. Each of these files was matched to the files created from the Professional Assignment Information Files described above for the appropriate year, and then all four years of API data were merged together. Finally, the cases with all information for all of the variables of interest were selected for analysis. In the case of the Williams schools and the statewide sample the 2002 sample means for the 4-year analysis were roughly similar to the sample means using the 2002 data,

(Note 9) which suggests that the selection criteria requiring 4 years of data did not result in a consequential loss of information for either group.

For ease of presentation, Table 4 provides descriptive statistics for the Williams schools and the statewide sample for 1999 and 2002. Like the Raymond sample above, middle schools and high schools are over-represented in the Williams sample compared to the state sample. 28% of the Williams schools are middle and 41% are high schools compared to 16% and 12% respectively. It is also worth noting that if we compare the means for the 35 schools used in Raymond's analysis shown in Table 2 to the means for this group, we see that the current Williams schools are relatively more disadvantaged. (Note 10)

Table 4
Descriptive Statistics for Statewide Sample of Schools with API Data from 1999-2002 and Current Williams Schools

	Williams Schools N=29		Statewide Sample N=6452	
	1999 Mean (S.D.)	2002 Mean (S.D.)	1999 Mean (S.D.)	2002 Mean (S.D.)
API	464.52 (65.51)	520.59 (87.68)	632.37 (131.76)	693.79 (112.35)
Elementary	.31 (.47)	.34 (.48)	.71 (.45)	.71 (.45)
Middle	.28 (.45)	.24 (.44)	.17 (.37)	.17 (.37)
High	.41 (.50)	.41 (.50)	.12 (.33)	.12 (.33)
% Minority	84.74 (20.03)	85.59 (18.26)	47.43 (29.24)	49.82 (29.48)
% Reduced/Free Lunch	75.05 (18.30)	75.62 (16.44)	49.03 (29.96)	48.91 (30.73)
% English Learners	46.59 (20.76)	44.83 (20.59)	24.00 (21.91)	24.44 (21.46)
Mobility	13.30 (18.50)	13.28 (6.72)	13.16 (12.74)	16.56 (9.49)
% Fully Cred.	71.25 (14.95)	69.75 (14.89)	86.71 (12.25)	87.17 (12.01)
% Emer. Cred	19.52 (14.37)	16.62 (13.92)	8.58 (9.90)	7.04 (8.38)
% Both Full and Emer.	2.90 (3.34)	2.24 (2.53)	1.92 (3.23)	1.27 (3.27)
% Missing	6.34 (3.24)	11.39 (6.97)	2.79 (4.56)	4.52 (5.90)
Yrs. Teach	12.31 (3.22)	11.63 (2.87)	13.24 (3.23)	13.25 (3.23)
% Yrs Tch. Missing	1.53 (2.01)	1.29 (1.45)	.59 (1.90)	.48 (1.83)

What should be immediately obvious from comparing across the columns is that on average, the Williams schools have very different profiles than the state average. Minority students and students receiving reduced and free lunch comprise the majority of the student populations in Williams schools. It is difficult to ascertain the changes in teachers' credentials over time because in both the Williams schools and the statewide sample, the percentage of teachers with missing information has increased by approximately 40 percent for the two groups from 1999 to 2002. (Note 11) Even in the unlikely scenario that all of the teachers with missing information were fully credentialed and that the percentage of teachers with emergency credentials has decreased by 2.9% in the Williams schools over the four-year period – twice the rate of the statewide sample – a large gap between the two groups remains. To put this figure in perspective, we might do well to remember that if the percentage of emergency credentialed teachers in the Williams schools was actually decreasing at that rate, it still will take more than 12 years for the Williams schools to reach the average for the state sample in 2002. In addition, while in the Williams schools sample the average years of teaching experience among the teaching staff has decreased slightly, in the statewide sample there is a very slight increase.

As I noted in the discussion of the replication of Raymond's analysis above, if we use the state sample of schools and the same group of independent variables in the model, the problem of collinearity among the independent variables results. Three variables among the group utilized by Raymond in her analysis are particularly highly correlated: *% Minority, % Reduced/Free Lunch*, and *% English Learners*. It is not surprising that *% Minority* and *% English Learners* are highly related since Latinos are the largest racial group in California's public schools and Spanish-speakers comprise the majority of the English learners in California. (Note 12) Similarly, given the strong relationships between race and poverty in the United States and the degree to which public schools continue to be segregated by both race and class, it is not surprising to see a strong relationship between these two variables. Tables 5 and 6 illustrate these relationships in two different ways. First, I divided the 6452 schools in Table 4 above into quartiles by the variable *% Minority* using the 1999 data. In Table 5 I present descriptive statistics for the first and fourth quartiles, the schools with the least and the most minority students, respectively, for 1999 and 2002. Of the 29 Williams schools in the analysis, 23 or 79% fall into the fourth quartile.

Table 5
Descriptive Statistics for Schools with Least and Most Minority Students 1999 and 2002

	Schools with the Least Minority Students in 1999 N=1616		**Schools with the Most Minority Students in 1999 N=1611**	
	1999 Mean (S.D.)	**2002 Mean (S.D.)**	**1999 Mean (S.D.)**	**2002 Mean (S.D.)**
API	775.52 (81.53)	810.47 (73.17)	485.02 (68.56)	579.91 (69.55)
% Minority	12.13 (5.26)	13.57 (6.37)	87.68 (8.53)	89.00 (7.52)

% Reduced/Free Lunch	17.97 (17.41)	16.91 (16.92)	81.41 (15.41)	83.00 (15.59)
% EL	5.75 (7.74)	6.13 (7.91)	48.21 (21.18)	47.60 (20.39)
School Mobility	10.42 (12.18)	13.53 (8.57)	14.39 (12.89)	17.57 (9.66)
% Fully Cred.	94.71 (5.78)	94.32 (6.48)	75.38 (13.57)	76.79 (13.92)
% Emer. Cred.	2.83 (4.45)	3.03 (4.62)	17.63 (11.85)	13.02 (11.07)
% Both Full and Emer.	1.10 (2.05)	.87 (2.33)	2.34 (3.33)	1.62 (3.68)
% Credential Missing	1.35 (2.65)	1.77 (2.97)	4.65 (5.87)	8.57 (7.48)
Avg. Yrs. Teaching	14.17 (3.15)	14.50 (3.17)	12.00 (2.99)	11.70 (2.83)
% Yrs. Tch. Missing	.26 (1.25)	.44 (1.93)	.92 (2.09)	.60 (1.69)

Table 5 allows us to see the strong relationship between many of the independent variables in the analysis. In general, schools that have relatively low percentages of minority students also have relatively low percentages of students eligible for reduced or free lunch, and relatively low percentages of English language learners compared to schools with high percentages of minority students. Mobility is also much lower in schools with less minority students. Likewise, on average, the schools with the least minority students also have much higher percentages of fully credentialed teachers (and conversely fewer emergency credentialed teachers or teachers with both full and emergency credentials indicating teaching out of field) and more experienced teachers. Tables 6 shows the bivariate correlations between these variables of interest over the four years of data, which allows to see trends over time in these variables. The most striking feature of Table 6 is that it illustrates the strong relationship overall between these variables. The lowest correlation between pairs of variables is just under .75. What we see in Table 6 is while the relationship between race and poverty at the school level is not only strong, but also consistently increasing from 1999-2002. It is also worth noting that with the exception of 1999, *% Minority* is also the most highly correlated with the variables for teacher credentials and experience (see Appendix).

Table 6
Correlations between Student Background Variables over Time

Bivariate Correlations	**% Minority & % Reduced/Free Lunch**	**%Minority & % ELL**	**% Reduced/Free Lunch & % ELL**
1999	.818	.754	.762
2000	.834	.750	.768

2001	.834	.746	.771
2002	.842	.745	.763

In Table 7 I present regression models for 1999 and 2002 using the full state sample. Because the yearly models use the same sample of schools, this strategy allows us to look at changes over time in the regression coefficients. In these models, I omitted the variable for *% Minority* after inspecting the regression diagnostics for the full model (Note 13). There were two main reasons for this choice: 1) this variable was marginally more highly correlated with the other independent variables in the model for three out of the four years in the analysis (Lewis-Beck, 1980); and 2) the percentage of students eligible for reduced or free lunch is the more theoretically interesting variable. (Note 14) I also included the controls for school type and the missing information on the teacher qualifications variables utilized in the regression models above (not shown). Finally, the last variable is an indicator variable denoting whether or not the school is one of the 29 Williams schools. Because the teacher credential variables add to 100%, they function as a set of indicator variables; *% Fully Credentialed* is the omitted comparison category.

Table 7
Regression Models

	1999		**2002**	
Constant	845.02 (1.76)***	819.40 (4.26)***	887.70 (1.73)***	863.01 (3.90)***
% Red./Free Lunch	-3.21 (.04)***	-3.21 (.04)***	-2.67 (.04)***	-2.68 (.04)***
% EL	-.80 (.06)***	-.79 (.06)***	-.57 (.05)***	-.53 (.05)***
School Mobility	-.18 (.06)**	-.14 (.06)*	-.93 (.08)***	-.89 (.08)***
% Emer. Cred.	-1.41 (.08)***	-1.24 (.09)***	-.34 (.09)***	-.17 (.09)***
% Both Full & Emer.	-1.59 (.24)***	-1.39 (.24)***	.10 (.21)***	.27 (.21)
Avg. Yrs. Teaching		1.77 (.27)***		1.66 (.23)***
Williams School	-16.36 (11.32)	-16.26 (11.29)	-38.24 (9.98)***	-37.88 (9.94)***
R-squared	.790***	.792***	.776***	.778***

In the first column for each year I present the models using just the teacher credential variables. In the second column I add the teacher experience variable to the model. The decrease we see in the coefficient for teacher credentials from the first model to the second is not surprising as the two variables are related. Schools with high percentages of teachers on emergency credentials are also more likely to have less experienced teaching staffs. However, both variables have a meaningful independent effect on the model, and the regression diagnostics indicate that while the teacher experience variable is not unrelated to the other variables in the model, it

is not so highly correlated that it might adversely affect the model.

The coefficient indicating whether or not the school is a Williams school is negative across all of the models, and in 2002 the coefficient is statistically significant. However, this coefficient should be interpreted with care because the 29 Williams schools are approximately one-half of one percent of the entire sample of 6452 schools. While the coefficient is negative, which can be interpreted that when all other factors in the model are statistically held constant, Williams schools do worse than other schools, two factors might be considered that temper this conclusion. First, this model controls only for the demographic characteristics of the student body and teachers' qualifications and not other unmeasured factors such as facilities and textbooks that might also influence student achievement. Second, the gap between Williams schools and all other school using the "raw" API in Table 5 is just over 173 points in 2002, which is considerably higher than the 38 point gap indicated by the coefficient for the Williams schools which we might read as the gap in school API once student and teacher characteristics are accounted for.

This strategy provides a good way to examine trends over time in the coefficients and the overall strength of the model. For example the decreases in almost all of the coefficients from 1999 to 2002 suggests regression to the mean. (Note 15) To confirm this interpretation, I examined the change in API from 1999 to 2002 by calculating the raw change score by subtracting the 1999 API from the 2002 API. Next I correlated the 1999 to 2002 percentage change score with *% Reduced/Free Lunch* and *% English Learners* for 2002. The bivariate correlations were .523 and .482 respectively (p= .001), which indicates that schools with higher percentages of poor students and English learners made greater gains in API over the four-year period. (Note 16)

The above strategy of analysis, which controls for school type, does not allow us to discern whether or not the variables of interest might work differently across school types. This is particularly important to consider given the predominance of elementary schools in the statewide sample (71%). In Table 9 I present the same models run separately by school type for 2002; descriptive statistics by school type are shown in Table 8. However, because there are relatively few Williams schools in each of the categories, (Note 17) and the purpose of this analysis is understanding the effects of teacher characteristics on each type of school, I omitted this variable from the analyses.

Table 8
Descriptive Statistics by School Type, 2002 API Growth Data

	Elementary N=4588	**Middle N=1084**	**High N=780**
API	706.67 (111.43)	674.60 (112.94)	644.66 (98.85)
% Reduced/Free Lunch	52.79 (31.36)	44.98 (27.56)	31.54 (23.69)
% EL	26.91 (22.78)	20.25 (17.35)	15.75 (14.28)
School Mobility	17.18 (8.18)	17.01 (12.54)	12.28 (10.57)
% Emer. Cred.	6.29 (8.25)	9.13 (9.10)	8.53 (7.33)

% Both Full and Emer.	1.09 (3.04)	1.70 (3.69)	1.75 (3.76)
Avg. Yrs. Teaching	12.99 (3.31)	13.38 (3.05)	14.55 (2.64)

What we see from Table 8 is that there are important differences across the school types. Students receiving reduced/free lunch and English learners are the most highly concentrated in elementary schools and least concentrated in high schools, with middle schools falling in between. This is in part because elementary schools tend to serve the smallest geographical areas, and are thus more likely to be economically segregated. The lower mobility in high schools could also be attributable to the larger geographical areas served by high schools as this variable measures within-district mobility; students whose families move frequently within the district would probably be less likely to have to change high schools than elementary or middle schools. While elementary schools have lower percentages of emergency credentialed teachers on average than middle and high schools, they also tend to have less experienced teaching staffs (although the average difference between middle and elementary schools is less than a percentage point).

Table 9
Regression Models by School Type, 2002 API Growth Data

	Elementary Coefficient (S.E.)	**Middle Coefficient (S.E.)**	**High Coefficient (S.E.)**
Constant	857.64 (4.49)***	827.76 (9.31)***	746.29 (15.22)***
% Reduced/Free Lunch	-2.65 (.04)***	-3.01 (.09)***	-2.20 (.12)***
% EL	-.51 (.06)***	-.60 (.13)***	-1.66 (.20)***
School Mobility	-.98 (.10)***	-.47 (.12)***	-.99 (.19)***
% Emer. Cred.	.18 (.11)	-.71 (.19)***	-1.42 (.30)***
% Both Full and Emer.	.93 (.26)***	-.51 (.41)	-1.86 (.55)***
Avg. Yrs. Teaching	1.71 (.26)***	1.55 (.04)**	2.17 (.87)*
R-squared	.777***	.820***	.689***

In Table 9 the regression models for each school type yield interesting findings. Of particular note are the results for the teacher credential and experience variables, all of which have the strongest effect in the high school model. This finding is masked in the model shown in Table 7 because of the predominance of elementary schools in the state sample. However, it is not surprising if we consider that teaching at the high school level requires the most specialized subject area training, which could also explain the strong negative effect of the variable for the percentage of teachers with both full and emergency credentials to the extent that it provides an indicator of the percentage of teachers teaching outside their subject area training. This finding is also consistent with those of Fetler (1999) who found that once student background characteristics are controlled, teacher training and experience were the strongest predictors of high school math achievement. (See also Darling-Hammond,

2000, more generally) In sum, these findings make it much more difficult to dismiss the effect of teachers' credentials on school performance as measured by the API, particularly at the high school level. To put these findings in more policy relevant terms, on average, a five percent decrease in the percentage of emergency credentialed teachers will increase school API by just over 7 points, which is close to the average target increase in API of 7.86 for this sample of high schools in 2002. (Note 18) The results obtained here also suggest that given the current budget constraints facing the state of California, a policy aimed at equalizing the quality of teachers across schools might be most effectively targeted at high schools, much like California's class-size reduction initiative targeted the lower grades.

To bolster this interpretation, I present a final analysis in this section in which I compare the 29 Williams schools with a group of comparison schools I created by matching each of the Williams schools with a similar school by selecting out all of the schools of the same type (elementary, middle, high) with the same value on the variable *% Reduced/Free Lunch* in the 2002 API Growth data. From the 29 lists that resulted, I chose a matching school for each Williams school by choosing the school with the most fully credentialed teachers that was also the closest match on the variable *% English Learners*. On average, this group had close to 19 percent more fully credentialed teachers than the Williams schools group. Thus, this comparison group most closely matches Raymond's call for comparing the achievement – or in this case what is more accurately described as school performance -- of the Williams schools with schools that are "abundant" in fully credentialed teachers, controlling for other factors. Table 10 provides descriptive statistics on the Williams schools and the "Abundant" comparison group on the variables of interest for 1999 and 2002. In addition, I also used t-tests to assess whether or not the differences in means across the two groups are statistically significant. What we see from Table 10 is that the three groups are roughly comparable in terms of student demographics. With the exception of *% Minority* between the Williams schools and the "Abundant" schools, none of the differences in means for the background variables are statistically significant. However, we also see that the major differences between the two groups of schools are in the variables measuring teachers' qualifications.

Table 10: Sample Means for Williams Schools and Comparison Schools

	Williams Schools N=29 Mean		**"Abundant" Schools N=29 Mean**	
	1999	**2002**	**1999**	**2002**
% Min.	84.74	85.59	70.89[a]	73.14[a]
% Reduced/Free Lunch	75.05	75.62	75.77	75.45
% English Learners	46.59	44.83	40.79	42.07
Mobility	13.30	13.28	8.81	12.59
% Emer.	19.52	16.62	7.35[aaa]	5.74[aaa]

% Fully Cred.	71.25	69.75	87.21[aaa]	88.42[aaa]
Yrs. Teach	12.31	11.63	14.17[a]	14.13[aa]

[a] Comparison with Williams schools mean statistically significant at p=.05
[aa] Comparison with Williams schools mean statistically significant at p=.01
[aaa] Comparison with Williams schools mean statistically significant at p=.001

Figure 1 shows the mean API for the two groups of schools from 1999 to 2002. What we see is that there is a consistent gap between the two groups of schools, a good portion of which we can reasonably attribute to the differences in teachers credentials across the two groups of schools as most of the background characteristics of students have essentially been held constant. (Note 19)

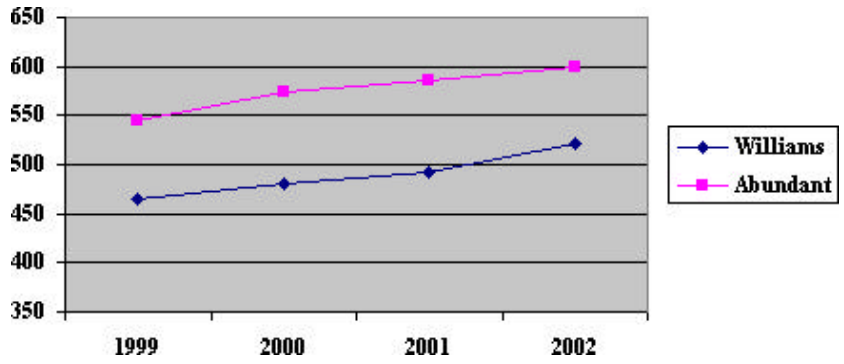

Figure 1. Mean API for Williams Schools and "Abundant" Schools: 1999-2002

Beyond Teacher Certification: Indirect Evidence of the Importance of Facilities

While most of Raymond's discussion focuses on teacher qualifications, her findings can also be read as providing indirect evidence of the importance of decent facilities on school performance. In one of her regression models with a smaller sub-sample of her comparison group, she includes a variable for whether or not the school is a year round school (Raymond, 2003a: Table 4). The regression coefficient is large (-25.81) and statistically significant at p=.05, which indicates that within this sample of schools (Note 20) with all other factors held constant, schools on a traditional calendar have an API score that is about 26 points higher than those on a year round calendar.

As noted above, there is no data available that would allow us to assess the effect of facilities on school performance as measured by the API. However, as I also noted, the year round multi-track calendars are utilized in California's public schools as a way to address overcrowding. According to the California Department of Education, if 20 percent of the students attending year round multi-track schools are "housed in excess of capacity at their school sites," the state and local school districts save approximately two billion in construction costs (California Department of Education, 2003: 92). Of the four types of multi-track calendars, "Concept 6" calendars are notable because they have fewer instructional days. Thus, we might consider the

use of year round calendars, and in particular the "Concept 6" calendar a proxy for inadequate facilities. In this section, I provide a more robust analysis of Raymond's findings which takes into account the different types of year round calendars by utilizing additional demographic data available from the California Department of Education that can be downloaded and matched to the API datasets.

The "Concept 6" calendar is used in only four districts across the state: Lodi Unified, Los Angeles Unified, Palmdale Elementary, Vista Unified. (Note 21) Lodi, Palmdale, and Vista are all small school districts with less than 30,000 students and well under 50 schools each. In Palmdale, all but one school in the district follows the "Concept 6" calendar. In Lodi Unified and Vista approximately half of the schools are "Concept 6" schools; however most other schools in Lodi follow a traditional calendar while in Vista, the majority of the remaining schools follow the other types of year round calendars. Because these three districts are so small and have such divergent patterns in their utilization of four types of calendars, I restrict the analysis to the Los Angeles Unified School District. The second largest school district in the country, the Los Angeles Unified School District also has a distribution of schools across the four types of school calendars that best allow us to test the effect of school calendar on school performance, controlling for other factors. This strategy also has the advantage of controlling for possible district effects that might distort the results if the four "Concept 6" districts were pooled for the regression analysis. Descriptive statistics for the 571 schools in the Los Angeles Unified School District with 2002 Base API (Note 22) scores and complete information on all variables are provided in Table 11.

Table 11
Descriptive Statistics for the Los Angeles Unified School District
2002 API Base Data

	Los Angles Unified School District N=571
API	635.81 (97.73)
Elementary	
Middle	**.14 (.34)**
High	**.09 (.28)**
% Reduced/Free Lunch	78.33 **(25.52)**
% English Learners	42.86 (23.30)
Mobility	15.54 (8.87)
% Fully Cred.	
% Emer.	14.63 (8.47)
% Both Full and Emer.	2.47 (2.35)
Yrs. Teach	11.52 (2.41)

Traditional	
YR Single	.02 (.15)
YR Mult.	.06 (.24)
"Concept 6"	.31 (.46)

Table 12 shows the regression model for the Los Angeles Unified School District. The same variables used in prior analysis were included, including the control variables for school type and missing information in the teacher variables (not shown). *Traditional* is the omitted comparison category. As a result, a positive coefficient on one of the remaining types of calendar variables indicates that schools operating on that type of calendar have higher API scores than schools running on traditional calendars. Conversely, a negative coefficient indicates that a school operating on the designated calendar has a lower API score than schools operating on a traditional calendar.

Table 12
Regression Model Testing for School Calendar

	Los Angles Unified School District API Base 2002 Coefficient (S.E.)
Constant	911.99 (19.67)***
% Reduced/Free Lunch	-1.91 (.13)***
% English Learners	-.72 (.15)***
Mobility	-1.17 (.25)***
% Emer.	-1.97 (.34)***
% Both	-1.18 (.89)
Yrs. Teach	.71 (1.18)
YR Single	58.54 (16.66)***
YR Mult.	-8.56 (9.01)
"Concept 6"	-13.94 (5.51)*
R-squared	.767***

What we see from the results here is that in the Los Angeles Unified School district, the district with 77 percent of all "Concept 6" schools in the state, with all other factors held constant, "Concept 6" school have lower API scores than schools on a traditional calendar. Conversely, schools on a year round single-track calendar have higher API scores than schools on the traditional calendar. Although there are only a small number of year-round single track schools in the Los Angeles Unified School District, this finding is notable because unlike the other types of year-round calendars, the year-round single track calendar is not used to increase enrollment but is primarily intended to increase student achievement by minimizing the learning

gap over the summer months. The year round single track calendar also increases the possibility for remedial and enrichment classes during intersessions. While this model tests how students are organized into facilities rather than directly testing the quality of the facilities, we see from this analysis that, as with the quality of the teaching staff, facilities are not inconsequential to school performance.

Conclusion

To a certain degree, Raymond and other expert witnesses for the state do agree with at least some of the broad issues involved in the case, more specifically, the importance of highly qualified teachers for student achievement. (Note 23) However, they argue against the creation of policies to help ensure a more equitable distribution of resources across schools. One of their main arguments is that it is very difficult to define "quality" teachers and that the minimum definition proposed by the plaintiffs – a fully credentialed teacher – doesn't measurably affect student achievement.(Note 24) Moreover, Raymond and other state experts further argue that the cost of decreased local control outweigh any possible educational effects of state policies mandating that schools and districts hire fully credentialed teachers and provide current textbooks and adequate facilities for their students. Finally, Raymond also asserts that these proposals are not only fiscally unreasonable given California's current budget crisis but would also have the effect of "disenfranchising parents" because they would "remove the option for parents to be co-creators of the educational programs that best meet the needs of their children" (17).

It is difficult to imagine how ensuring that most, if not all teachers are credentialed would disempower parents. While information regarding teachers' credentials is currently made available to parents in a standard reporting format through state-mandated school accountability report cards, not only is there a significant lag in the information (i.e. information about the prior school year is reported in the report card published the following school year) but all of the information is aggregated at the school level. As a result, it is difficult for a parent to use this information to advocate for her/his child; the best remaining options, then, are direct inquiry at the school or the word-of-mouth networks that exist among parents. (Note 25) Given these conditions, it could be argued that ensuring that most, if not all classroom teachers are fully credentialed would actually empower parents because they can be assured that all of their children's teachers meet the criteria established for teachers by the California Commission on Teacher Credentialing and can thus use their time and resources advocating for their children in other arenas.

The results of these analyses suggest that short of desegregating schools by socioeconomic status, increasing and equalizing the percentage of fully credentialed teachers is an "input" that is not only relatively amenable to change through state and local policy–and certainly much easier than building additional facilities to ease overcrowding–but also contributes to school performance. (Note 26) Addressing the disparities documented here might entail creating pay and other incentives (e.g. increased autonomy) that would encourage experienced teachers to work in high-poverty schools. (Note 27) And, given that the API is essentially an average of student scores, while the magnitude of the effect on *schools* is subject to debate, such a policy could make a large difference for the academic achievement and life chances of *individual students*. Even if we accept the argument that a more equitable distribution of teachers has a relatively small but positive effect on achievement, we might also consider whether or not the goal of increasing equity in

public education – which this analysis suggests can be done without sacrificing school performance – is an important and desirable end in itself. As we consider the issues in this case, it is important not to let statistical arguments about the determinants of school achievement and the valorization of local control distract us from the larger issues of justice and educational opportunity for *all* students that are at the heart of this case. For those of us who are comfortably middle class or higher, why should we expect poor and minority students to settle for anything less than the schools we want and often demand for our own children?

Notes

1. For a synthesis of the plaintiffs' experts' reports see Oakes (2002a).

2. A May 2003 *San Francisco Chronicle* newspaper story reported that the state has spent approximately $18 million fighting the case (Asimov, 2003).

3. Similar arguments were made by Hanushek (2003), Hoxby (2003), and Philips (2003). I focus specifically on Raymond's report here because of her use of API data to create what she describes as "econometric models of educationally challenged schools in California." In contrast, Philips (2003) discussion of the API is a secondary analysis of reports using API data.

4. Raymond indicates that she substituted API scores for prior years in the case of two schools without 2002 API. When I selected out the cases listed in Raymond's Table 1, I found a third school was missing API information. It is also worth noting that two of the remaining 36 schools were not listed as plaintiff schools on the Williams v. California website; however I included these in the analysis for the purpose of reconstructing Raymond's analysis as precisely as possible. The Williams case website (www.decentschools.org) lists a total of 72 plaintiff schools, but it is unclear from Raymond's narrative why her analysis focused on the group of schools in listed in Table 1 of her report.

5. An additional school was missing information on the parent education variable and was omitted from the analysis.

6. I also restricted the sample to schools with less than 30% missing information on the teacher credentials and experience variables.

7. The 1999 API excluded alternative schools and schools with fewer than 100 students. In the 2000-2002 files used here, there was a fourth school type indicating if the schools was a small school, i.e. the school only had between 11 and 99 valid tests available to calculate its API. Since none of the Williams schools fell into this category, I omitted the approximately 39 schools designated as small schools for these three years.

8. Hoxby also argues that "Good Research" utilizes extensive controls for family background, measures that are either unavailable, or as I will detail below, in the case of the parental education variables in the API unreliable. However, it is also worth noting that Hoxby's analysis of the effects of centralization on state performance on NAEP provided in her report does not appear to control for students' family background.

9. For the 4-year state sample, the 2002 Growth API was less than 8 points lower

than the state sample in Table 2 above. The sample means on all other variables were within a percentage point of each other.

10. 30 of these schools have 2002 Base API scores with a mean for this group is 529.13, which is roughly comparable. As with the state sample, the means on the other variables using the 2002 Base API sample are all within a percentage point or two of those presented for 2002 in Table 4.

11. If we look at the percentage of fully credentialed teachers in the Williams schools for 2000 and 2001, both years with less missing information than 2002, there trend is somewhat inconsistent but tends more towards a decreasing percentage of fully credentialed teachers. In 2000, the average Williams school had 68.91 percent of the teachers were fully certified (7.96 percent missing information). In 2001 the same figure was 70.16 (9.41 percent missing information).

12. According to 2001-2002 figures available on EdData (http://www.ed-data.k12.ca.us/welcome.asp), in 2001-2002, 44.2 percent of all California public school students were Latino. Of the 25.5% English Learners in California's public schools, 21.2% were Spanish speakers.

13. Subtracting the tolerance statistic from 1 gives us the R-squared from the regression of all the other independent variables on the independent variable of interest (R^2_j where j=the variable of interest). Fox (1991) recommends taking the square root of R^2_j, noting that when this figure approaches .9, collinearity becomes a serious problem for the estimation of regression coefficients. In this case, I obtained R_j of .88 for the variable % Minority for three of four year of the analysis (this figure was only marginally lower for the 1999 model).

14. See, for example the analysis by Phillips, Brooks-Gunn, Duncan, Klebanov, and Crane in Jencks and Phillips (1998). .

15. I could have just as easily chosen the 1999 figures for *% Reduced/Free Lunch* and *% English Learners* as they correlate at r> .95 for both variables, reflecting the relative stability of these variables over time.

16. Rogosa (2001, 2002) notes that this pattern is in part a result of how the API is constructed. Because the API is constructed by using a percentile rank metric, students in the highest scoring schools can't raise school scores because they have "topped out" the index.

17. 10 of the Williams schools are elementary schools, 7 are middle schools and 12 are high schools.

18. A schools API target or the amount its API should increase from one year to the next is determined by taking 5 percent of the difference between the school's API in a given year and 800 which is the target API for all schools set by the state. Rogosa (2002) has argued that a difference in API of 5 points or fewer is not significant and is approximately equivalent to about half of the students answering an additional question on the SAT-9 test correctly. Rogosa (2000) has also estimated that if every student increased their percentile rank on each test by one point, the school's API would increase by 8 points (1). I frame the results in terms of the growth targets set by the state because irrespective of the educational consequences of a rise and fall

in API, whether or not schools reach their targets has important political consequences for schools because it is one of the criteria for determining whether or not a school is labeled as performing adequately or inadequately by the state.

19. Even we look at the range of API between these two groups as Rogosa (2001) suggests, although there is substantial overlap in the middle, the lowest boundary for the Williams schools is substantially lower than the minimum for the "Abundant" schools. Likewise, the maximum value among the "Abundant" school is much higher than the maximum value for the Williams schools. I also chose a second comparison group by matching each school by school type and *% Reduced/Free Lunch* and *% Fully Credentialed*. On average, the API scores of this second comparison group were slightly higher than the Williams schools, but the difference in means was not statistically significant.

20. The 129 schools in this model are most likely predominantly middle schools and high schools because one of the variables in the model that Raymond describes as Number of Core Classes missing information for most schools in the full dataset. However, this variable appears to be mislabeled on the CDE website. For all of the other API datasets with the exception of the 2002 Base API Raymond used in her analysis, this variable indicates the Average Size Class, Core Classes, which include the following subject areas: English, Foreign Languages, Math, Science, and Social Science (California Department of Education, Policy and Evaluation Division, 2003: 8). Given this definition, it is not surprising that so many schools were missing information on this variable; if one examines the data by looking at the distribution of the variable by school type (elementary, middle, high) 87.5% of the 2302 schools with information on this variable were middle and high schools. As a result, it is also likely that many of the Williams schools were not included in this model. Interestingly, Raymond uses this model – with an incorrectly interpreted coefficient – as evidence for her assertion that other types of improvements will increase student achievement more than increasing the percentage of fully certified teachers.

21. In contrast, well over 100 districts have schools following various types of year round multi-track calendars other than the Concept 6 calendar.

22. I use the 2002 API Base Data because here, unlike the prior analyses, understanding changes over time is less important and the 2002 Base data with its greater incorporation of the CST, is currently the most politically salient for schools. However, the results I obtained here are consistent with the results using the 2002 Growth API data, which is not surprising because the 2002 Growth API and the 2002 Base API correlate at .999.

23. Raymond (2003) writes: "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" (11). Similarly, Philips (2003) writes: "Though inconvenient, students can share books, use copied materials, or internet resources, wear coats in a cold classroom, or use a restroom on another floor. But if a classroom teacher is not able to effectively focus instruction on the state content standards, for the subject area of the class, disadvantaged children may be ill-equipped to learn the material on their own" (75).

24. To some degree, this is a moot point since No Child Left Behind requires that all teachers of core subject areas be "highly qualified" by the 2005-2006 school year. Teachers on emergency permits, waivers, or pre-intern certificates do not meet the criteria for "highly qualified." A June 2003 memo from the State Superintendent of Public instruction directs districts, counties and charters schools to focus their current hiring and recruitment efforts on teachers that meet the NCLB requirements (O'Connell, 2003).

25. A similar argument can be made about textbooks. Even if classroom teachers use textbooks differently, wouldn't it be more empowering to parents to know that the current textbooks are available in their child's classroom for the teacher to use at her/his discretion?

26. Kahlenberg (2001) argues that the economic integration of schools could also be a strategy to insure a more equal distribution of teachers across schools (78-80).

27. As noted above, some of the changes generated in the wake of NCLB address the issue of credentialing. However, the results of an earlier analysis focusing on the Los Angeles Unified School District and the San Diego Unified School District suggested that the gains made from hiring more fully credentialed teachers could be offset by a loss in more experienced teachers as this will in all likelihood emerge as a source of inequality between schools once the disparities in credentials are equalized (Powers, 2003).

References

Asimov, N. (2003, May 5, 2003). Bitter Battle over Class Standards. State Spends Millions to Defeat Students' Suit. *San Francisco Chronicle*.

California Department of Education. (2000). *Fact book 2000: Handbook of education information*. Sacramento: Author.

California Department of Education. (2001). Year-round education calendars." Sacramento: Author. [accessed May 15, 2003]

California Department of Education Office of Policy and Evaluation. (2000). *1999 base year academic performance index (API)*. Sacramento: Author.

California Department of Education Office of Policy and Evaluation. (2003). *Explanatory notes for the 2002 Academic Peformance Index base report*. Sacramento: Author.

Darling-Hammond, L. (2002). "Access to quality teaching: An analysis of inequality in California's schools." *Decent Schools for California: Williams v. State of California*. (Available online: http://www.decentschools.org)

Fetler, M. (1999). High school staff characteristics and mathematics test results. *Education Policy Analysis Archives*, *7*(9).

Fox, J. (1991). *Regression diagnostics*. Newbury Park, CA: Sage Publications.

Glass, G. V., & Hopkins, K. D. (1996). *Statistical methods in education and psychology (3rd Ed.)*. Boston: Allyn and Bacon.

Hoxby, C. M. (2003). Achievement, efficiency and centralization in California's public schools. *Decent Schools for California: Williams v. State of California*. (Available online: http://www.decentschools.org)

Laczko-Kerr, I., & Berliner, D. C. (2002). The effectiveness of "Teach for America" and other under-certified teachers on student academic achievement: A case of harmful public policy. *Education Policy Analysis Archives*, *10*(37).

Lewis-Beck, M. (1980). *Applied regression: An introduction*. Beverly Hills, CA: SAGE Publications.

Oakes, J. (2002). Education inadequacy, inequality and failed state policy. A synthesis of expert reports prepared for Williams v. State of California. *Decent Schools for California: Williams v. State of California*. (Available online: http://www.decentschools.org)

O'Connell, J. (2003, June 30). Update on California's plan of highly qualified teachers (Letter to district supervisors, county superintendents, and charter school administrators: . Sacramento: California Department of Education. [Available online at www.cde.ca.gov]

Powers, J. M. (2003). An analysis of performance-based accountability: Factors shaping school performance in two urban school districts. *Educational Policy, 17*(5), 558-586.

Purdum, T. S. (2000, May 18). Rights Groups Sue California Public Schools. *New York Times*.

Raymond, M. E. (2003a). Williams v. State of California report by Margaret E. Raymond . *Decent Schools for California: Williams v. State of California*. (Available online: http://www.decentschools.org)

Raymond, M. E. (2003b). The performance of California charter schools . Palo Alto, CA: CREDO, Hoover Institution.

Rogosa, D. (2001). *Year 2000 update: Interpretive notes for the Academic Performance Index*. Palo Alto, CA: Stanford University.

Rogosa, D. (2002). A further examination of student progress in charter schools using the California API. Palo Alto, CA: Stanford University.

Russell, M. (2000). Summarizing change in test scores: Shortcomings of three common methods. *Practical Assessment, Research and Evaluation*, *7*(5).

About the Author

Jeanne M. Powers
Division of Educational Leadership and Policy Studies
Arizona State University
P. O. Box 872411
Tempe, AZ 85287-2411
Phone: 480.965.0841

Email: jeanne.powers@asu.edu

Jeanne M. Powers is an Assistant Professor in the Division of Educational Leadership and Policy Studies, Arizona State University. She received her PhD in Sociology from the University of California, San Diego. Her current research examines the equity implications of education policy with a focus on high-stakes testing and accountability, and school choice.

Appendix

Table A1
Correlations between Student Background Variables and Teacher Qualifications Variables Over Time

Bivariate Correlations: % Emer. Cred and…	**% Minority**	**% Reduced/Free Lunch**	**% English Leaners**
1999	.589	.457	.469
2000	.566	.447	.429
2001	.510	.378	.378
2002	.457	.347	.340
Bivariate Correlations: Avg. Yrs Exper and …	**% Minority**	**% Reduced/Free Lunch**	**% English Learners**
1999	-.272	-.279	-.284
2000	-.325	-.306	-.308
2001	-.332	-.293	-.318
2002	-.343	-.302	-.315

The World Wide Web address for the *Education Policy Analysis Archives* is **epaa.asu.edu**

Editor: Gene V Glass, Arizona State University

Production Assistant: Chris Murrell, Arizona State University

General questions about appropriateness of topics or particular articles may be addressed to the Editor, Gene V Glass, glass@asu.edu or reach him at College of Education, Arizona State University, Tempe, AZ 85287-2411. The Commentary Editor is Casey D. Cobb: casey.cobb@unh.edu.

EPAA Editorial Board

Michael W. Apple David C. Berliner

University of Wisconsin

Greg Camilli
Rutgers University

Sherman Dorn
University of South Florida

Gustavo E. Fischman
Arizona State Univeristy

Thomas F. Green
Syracuse University

Craig B. Howley
Appalachia Educational Laboratory

Patricia Fey Jarvis
Seattle, Washington

Benjamin Levin
University of Manitoba

Les McLean
University of Toronto

Michele Moses
Arizona State University

Anthony G. Rud Jr.
Purdue University

Michael Scriven
University of Auckland

Robert E. Stake
University of Illinois—UC

Terrence G. Wiley
Arizona State University

Arizona State University

Linda Darling-Hammond
Stanford University

Mark E. Fetler
California Commission on Teacher Credentialing

Richard Garlikov
Birmingham, Alabama

Aimee Howley
Ohio University

William Hunter
University of Ontario Institute of Technology

Daniel Kallós
Umeå University

Thomas Mauhs-Pugh
Green Mountain College

Heinrich Mintrop
University of California, Los Angeles

Gary Orfield
Harvard University

Jay Paredes Scribner
University of Missouri

Lorrie A. Shepard
University of Colorado, Boulder

Kevin Welner
University of Colorado, Boulder

John Willinsky
University of British Columbia

EPAA Spanish and Portuguese Language Editorial Board

Associate Editors for Spanish & Portuguese

Gustavo E. Fischman
Arizona State University
fischman@asu.edu

Pablo Gentili
Laboratório de Políticas Públicas
Universidade do Estado do Rio de Janeiro
pablo@lpp-uerj.net

Founding Associate Editor for Spanish Language (1998-2003)
Roberto Rodríguez Gómez

Universidad Nacional Autónoma de México

Adrián Acosta (México)
Universidad de Guadalajara
adrianacosta@compuserve.com

Teresa Bracho (México)
Centro de Investigación y Docencia Económica-CIDE
bracho dis1.cide.mx

Ursula Casanova (U.S.A.)
Arizona State University
casanova@asu.edu

Erwin Epstein (U.S.A.)
Loyola University of Chicago
Eepstein@luc.edu

Rollin Kent (México)
Universidad Autónoma de Puebla
rkent@puebla.megared.net.mx

Javier Mendoza Rojas (México)
Universidad Nacional Autónoma de México
javiermr@servidor.unam.mx

Humberto Muñoz García (México)
Universidad Nacional Autónoma de México
humberto@servidor.unam.mx

Daniel Schugurensky(Argentina-Canadá)
OISE/UT, Canada
dschugurensky@oise.utoronto.ca

Jurjo Torres Santomé (Spain)
Universidad de A Coruña
jurjo@udc.es

J. Félix Angulo Rasco (Spain)
Universidad de Cádiz
felix.angulo@uca.es

Alejandro Canales (México)
Universidad Nacional Autónoma de México
canalesa@servidor.unam.mx

José Contreras Domingo
Universitat de Barcelona
Jose.Contreras@doe.d5.ub.es

Josué González (U.S.A.)
Arizona State University
josue@asu.edu

María Beatriz Luce(Brazil)
Universidad Federal de Rio Grande do Sul-UFRGS
lucemb@orion.ufrgs.br

Marcela Mollis (Argentina)
Universidad de Buenos Aires
mmollis@filo.uba.ar

Angel Ignacio Pérez Gómez (Spain)
Universidad de Málaga
aiperez@uma.es

Simon Schwartzman (Brazil)
American Institutes for Resesarch–Brazil (AIRBrasil)
simon@sman.com.br

Carlos Alberto Torres (U.S.A.)
University of California, Los Angeles
torres@gseisucla.edu

EPAA is published by the Education Policy Studies Laboratory, Arizona State University