Evaluation Study of the Immediate Intervention/Underperforming Schools Program of the Public Schools Accountability Act of 1999

Final Report

Submitted to:

Kim Wells, Contract Monitor
Evaluation Unit
Policy and Evaluation Division
California Department of Education
1430 "N" Street
Sacramento, CA 95814

Project Leadership:
Tom Parrish, Principal Investigator
Catherine Bitter, Project Director
Maria Pérez, Director of Achievement Analysis
Raquel González, Project Coordinator

Report Authorship:
Catherine Bitter
Maria Pérez
Tom Parrish
Raquel González
Miguel Socias
Lizzie Salzfass
Kassie Chaney
Paul Gubbins
Priyanka Anand
Karl Dawson
Vicky Yu
Danielle Delancey
Phil Esra

September 30, 2005

Acknowledgments

We wish to extend our appreciation to the district and school administrators, teachers, SAIT members, and all interview respondents participating in this evaluation. We attribute the success of this evaluation effort to the valuable input provided by the participants.

We also wish to acknowledge the invaluable assistance of key project staff who contributed to this study. In addition to the authors of this report, we wish to thank Roger Levine, Yu Zhang, Heather Quick, Jennifer O'Day, Amy Merickel, and Jean Wolman for their insights at various points in the project. We would like to thank Cecilia Speroni and Amie Haas for providing support to the student achievement analysis. We also wish to thank Estherlyn Juanitas, Lee Carlson, Diana Doyal, and Connie Conroy for providing administrative assistance throughout the project.

Finally, we wish to acknowledge the guidance and input provided by the project Advisory Group, as well as Kim Wells, Rachel Perry, and other staff of the CDE. We want to thank Jeff Zettel, Cliff Rudnick, Fred Balcom, Wendy Harris, and Laura Wagner for providing data and answering questions for this evaluation.

Evaluation Advisory Board Members

Holly Covin Jacobson (California School Boards Association)

Jacque Roberts (Office of the Secretary)

Paul Warren (Legislative Analyst's Office)

Lynette Nyaggah (California Teachers Association)

Pete Cervinka /Lenin Delcastillo (Department of Finance)

Lisa Hall (1st grade teacher at Glenwood Elementary School in Robla School District)

Executive Summary

Over the past decade, the standards-based reform movement in education has focused attention on the responsibility of schools and school systems to produce results in student learning. Nationally this focus is most evident with the implementation of accountability provisions of the federal No Child Left Behind (NCLB) Act of 2001. However, even prior to NCLB, many states and districts had instituted their own performance-based accountability programs. In 1999, the California legislature approved the Public Schools Accountability Act (PSAA), which incorporated three central components designed to hold schools accountable for improving student outcomes. The initial PSAA components were the Academic Performance Index (API), the Immediate Intervention/ Underperforming Schools Program (II/USP), and the High Performing/Improving Schools Program (also known as the Governor's Performance Award (GPA)). Later, the High Priority Schools Grant Program (HPSGP) was incorporated into the PSAA legislation. The state now faces the challenge of integrating these components with the federal NCLB Act.

In June 2003, AIR completed the first legislatively mandated study of the PSAA, with support from Policy Analysis for California Education (PACE) and EdSource.[1] In November 2004, AIR was contracted by the California Department of Education to conduct a continuation study of the II/USP component of PSAA. The II/USP provided funds to low-performing schools in the state to develop and implement an Action Plan for school improvement, with the assistance of a state-approved External Evaluator. Schools subsequently had two to three years to implement the Action Plan, and are subject to sanctions at the end of this implementation period if they did not improve student performance.

Although II/USP has essentially ended, examining this policy retrospectively provides important information that can be used to shape current and future accountability programs and legislation in California and other jurisdictions. With the implementation of NCLB, it is critical for California policymakers to understand what has worked well, and not as well, within the state system as they consider ways to continue implementation of the state accountability system and to resolve differences between the state and federal programs.

The Approach

Building on the findings from the 2003 PSAA evaluation, this study was designed to provide follow-up information on the overall impact of the II/USP program and on the factors and strategies that have contributed to growth in selected low-performing schools. In addition, since II/USP has evolved to incorporate the School Assistance and Intervention Team (SAIT) process, we examined this component of the policy specifically.[2]

[1] The final report from this study can be found at http://www.air.org/publications/pubs_ehd_school_reform.aspx

[2] II/USP schools that did not make significant growth in both of the two implementation years were required to enter into a contract with a SAIT. SAITs are teams of educational consultants who work with and monitor schools to improve student achievement.

The AIR team used a mixed-method approach that allowed us to look across cohorts and to triangulate perspectives and data from multiple sources. First, we analyzed statewide achievement trends using data from all II/USP schools and relevant comparison schools across the state to determine the impact of II/USP in the aggregate. Second, we conducted telephone interviews with administrators and teachers in schools that had made substantial API growth during and after II/USP implementation, in II/USP schools that made minimal API growth, and in similar non-II/USP schools that demonstrated high API growth. We also interviewed district-level administrators associated with our sampled schools and interviewed a small sample of state-level respondents. To explore the SAIT process, we included a subset of SAIT schools in our telephone interview sample, and interviewed the SAIT leader assisting the sampled schools.

Overall Conclusions Regarding II/USP

Below we discuss our findings across four main areas: the overall impact of II/USP on student achievement, factors contributing to or hindering growth within II/USP schools, the role of the district in school improvement efforts, and the impact of NCLB implementation on the focus and implementation of the state accountability program.

1) Overall, the impact of II/USP participation on student achievement has been negligible. Any small advantage experienced by II/USP schools relative to comparison schools during program participation dissipated before or soon after program completion.

While we observed high overall growth in student outcomes among *all* low-performing schools, we found only modest differences in student outcomes between II/USP and comparison schools. In some cohorts, in some grade levels, II/USP schools experienced slightly larger achievement gains in comparison to similar low-performing schools in the state during II/USP planning and implementation years. More often, however, there was no observed effect, and trends were inconsistent across cohorts and grade levels. Even where an II/USP advantage occurred and was statistically significant, the effect was so small (approximately .02 standard deviations) as to be educationally inconsequential. Moreover, any advantage observed during the period of program participation appeared to dissipate after program completion. Indeed, we found a small, but statistically significant, negative impact of II/USP in the years after the program ended for groups of schools that had made gains relative to non-II/USP comparison schools in the planning or implementation years.

Given the focus in this study on the longer-term effects of II/USP, particular attention must be paid to the decline in student achievement growth in II/USP schools, in relation to similar non-II/USP schools, that is found in later years. The change in relative growth may result from increased growth among non-participating schools, as policies such as NCLB place pressure on other low-performing schools in California. However, the change in relative growth could also be a result of a decline in II/USP schools' growth after the program ends. Our interview data suggest that despite optimistic expectations of long-term growth, many schools had to reduce or eliminate programs implemented using II/USP funds. This calls into question the ability of schools to sustain reforms and growth after program completion.

2) Despite the lack of an overall program effect, there is evidence that II/USP participation contributed to growth in <u>some</u> schools. Respondents in these schools identified specific factors and strategies they believe led to their improvement.

Even though we did not find an overall impact of II/USP on student achievement, some previously struggling schools were able to make substantial progress during the time of II/USP implementation, and some attributed this growth to participation in the program. Several essential factors for growth were identified by schools that made consistent and/or high growth in student achievement during II/USP. These included school capacity, instructional coherence, and systematic assessment and data-based decision-making. A subset of these schools spoke highly of the contribution of II/USP, citing the funding and the focus provided through planning as most important.

However, many others did not experience high growth, nor had positive experiences with the program. Factors reported to hinder the effectiveness of II/USP included late distribution of funds, limited guidance on the use of the funds, limited communication among II/USP schools and between the schools and the state, mixed experiences with External Evaluators, limited follow-up support and monitoring, and the focus on negative labeling and sanctions. Some schools reported that they saw II/USP as a wake-up call, while others found it de-motivating to be labeled as a failing school. Some also noted the seeming irony of receiving supplemental funds while failing, then subsequently having the funds removed upon success.

Of particular note is that while the subset of schools that made substantial progress generally felt that they were successful in improving student learning, it was evident that few schools were communicating with each other to share these effective improvement strategies. Several respondents mentioned that they would have liked to have more opportunities for communication with other schools participating in II/USP to learn more about what was working in other schools.

3) Local districts were found to influence the achievement trends in low-performing schools. While district supports were reported as key to some schools' improvement, these supports did not appear to be present in all schools.

As in the 2003 PSAA evaluation, we again observed a significant district effect on the achievement trajectories of low-performing schools in the state. That is, while II/USP and comparison schools performed in similar ways *within* districts, the pool of low-performing schools varied considerably in performance *across* districts (controlling for the characteristics of the student population and other factors). These findings, coupled with our interview data, suggest that districts serve as intermediaries between state-level policy and school-level implementation. Additionally, districts can institute policies and supports designed to improve the work of all their low-performing schools, irrespective of program participation. These include technical assistance and professional development, particularly around systematic assessment and data use, as well as the targeting of resources to low-performing schools. Districts can also provide a focus for schools' improvement efforts. Such supports and focus did not appear to be present in all districts or realized by all schools.

4) While possibly intensifying the focus on accountability overall, the presence of differing state and federal accountability systems has diffused the attention schools are paying to PSAA.

The implementation of NCLB has created an additional layer of achievement targets and expectations for schools in California. Schools are expected to meet both API and AYP targets, and they face sanctions from both the state *and* federal governments if they fail to do so. As discussed earlier, by focusing efforts on all schools, NCLB has possibly raised the focus on accountability among schools *not* participating in II/USP.

Some respondents did *not* report the additional layer of accountability as a major distraction, saying that their plans to improve student achievement overall should meet the needs of both programs. Many others, however, at both the school and district levels, reported that having two separate accountability systems in place has "confused" or "overwhelmed" school and district staff. Respondents referred to conflicting expectations among the two policies. They generally reported that the AYP targets associated with NCLB were more challenging to meet than API targets. In addition, they reported that the focus on *growth* in the state system was more appropriate to the improvement goals of low-performing schools than meeting a set proficiency target. Finally, respondents reported an increased level of attention toward the AYP targets, due to perceptions of stricter outcome requirements and more severe sanctions.

Recommendations for Ongoing Accountability Efforts

Because II/USP has essentially come to an end, the recommendations below are not specifically directed at this program but instead more generally at the state's ongoing efforts to improve low-performing schools. The recommendations focus on systemic changes and specific strategies we believe will be needed to enhance the likelihood that accountability interventions like II/USP will succeed in the future.

1) The state should consider the resources needed for sustained academic success in low-performing schools and ensure that they are present and will be sustained in applicant schools and their districts.

The state should specify ongoing resource standards to address the needs of the state's most challenging, highest-poverty schools, and ensure they are allocated effectively by districts to schools, as a precondition for programs like II/USP. Given that schools reported being unable to sustain key strategies implemented through II/USP, sufficient resources to *sustain* improvement efforts must be an important component of any program that provides focused resources for change. Since the district was found to be a key intermediary between state-level policy and school-level implementation, the state must ensure that districts have the resources to provide the necessary assistance and support to their schools, and that they allocate them to low-performing schools as needed. For example, they might be required to ensure a teaching force in these schools that at least equals, or exceeds, the district average on such characteristics as experience and degree level.

2) The full power and potential of districts and other intermediate agencies should be brought to bear on the problem of shoring up low-performing schools. Districts, and other relevant agencies, should be expected to play a key role in providing support, vision, and guidance for school-level improvement, and should be held accountable by the state when this does not occur.

Given the high number of underperforming schools in II/USP, and in California overall, the state will need all the help it can get in providing the intensive support and monitoring required for schools undergoing improvement efforts. This study demonstrates that districts can make a substantial impact on improving low-performing schools when they focus their attention, energy, and resources on this effort. Districts are clearly an important vehicle to ensure proper and effective implementation of programs such as II/USP, and to ensure the growth of the lowest-performing schools in their jurisdiction.

Within such as a system, districts should be held accountable for providing additional focused assistance for the schools most in need, and for providing long-term and sustained support to maintain progress made through programs like II/USP. To hold districts accountable for providing this assistance, the state should explore ways to provide a system of incentives, guidance, support, and sanctions to districts that fail to make progress with their low-performing schools. Since the effectiveness of district-level sanctions is yet to be determined, the state should regularly examine data to assess the effectiveness of any such system put in place.

3) The state should acknowledge schools that improve within programs like II/USP or are successful with high percentages of students in poverty.

The vast majority of low-performing schools in the state and across the nation are those with high percentages of students in poverty. The odds against success in these schools are formidable, and yet some schools are able to defy these odds to make substantial academic gains and to sustain this growth. The findings from this study suggest that garnering attention by focusing only on schools that are failing has the potential to de-motivate schools.

To foster proper acknowledgment of success, the state should ensure that continued sufficient resources are available to schools that make growth. At the least, high-poverty schools that are beating the odds should receive comparable *public acknowledgement* to those who are failing. This acknowledgment can potentially serve both as a continuing motivator for growth, as well as a means to identify models for other high-poverty schools that are not succeeding.

4) The state should work with districts to develop vehicles whereby schools in need of improvement can be linked with and learn from schools that have been successful in improving outcomes with comparable populations of students.

In light of the limited communication reported and evident among schools in our sample, we recommend that the state and districts work in tandem to better facilitate this communication by establishing structures for schools to learn from each other. Districts should be assisted in learning how to pair struggling and successful schools, in cases where this is possible within a given district. The state should further assist in facilitating pairings of schools across districts as needed. The state should also work with districts to determine useful techniques for enhancing knowledge transfer across the two schools. The state should track the performance of schools that have been paired with others in an attempt to assess when these

pairings have been successful and to make adjustments as needed based on results. The first step in developing such a structure is to establish clear criteria for success and to review data in light of these criteria to identify schools. The state should allocate personnel to review data yearly to this end.

5) Similarly, the state should identify *districts* that have successfully improved student outcomes in their low-performing schools, and facilitate opportunities for other districts to learn from them.

Given that districts can play a key role in improvement efforts at their schools, we recommend a similar effort as above, but targeted at districts. In this case, the state would again set simple and straightforward criteria for high-growth districts that can be reviewed on a yearly basis. Using these data, we recommend that the state provide opportunities, either through regular conferences and meetings, or through a more intensive partnering program, to facilitate learning across districts. Attention should be paid to the characteristics of districts to ensure that districts are matched with others that have similar student populations.

6) The state should foster the use of benchmark data as a feedback mechanism for informing instruction and identifying students for extra support, as described in this report.

A key factor for growth identified through this study was the systematic analysis and use of data to inform instruction. Respondents emphasized the importance of frequent (e.g., monthly or every 6 weeks) benchmark assessments tied to the school's curriculum and tied to the state standards. An important component of data use was communication and collaboration around data among grade-level and cross-grade-level colleagues. This was commonly mentioned as a key factor for turning around schools in our study, as well as in the general literature regarding effective practices. Therefore, we recommend that systems be fostered and disseminated statewide that encourage and enable schools to utilize data on a regular and ongoing basis to inform instruction.

7) The state should work with districts, and other agencies, to incorporate long-term guidance and monitoring into assistance programs for underperforming schools.

Given the reportedly mixed experience with External Evaluators and the concerns raised by respondents about the limited support and monitoring received during the implementation years, we recommend that the state accountability system (e.g., HPSGP) incorporate sustained support and monitoring, while holding external providers accountable for assisting and realizing results within their schools. This recommendation is bolstered by the finding in our achievement analyses that, for some cohorts, a decline in growth relative to comparison schools occurred after the planning year.

8) Analogous to what was reported for the school level, the *state* should use data on an ongoing basis to identify the extent to which state-level programs make an impact, and use these data to inform and alter policy and programs in support of low-performing schools as needed.

As we set expectations for schools to regularly use data, we suggest the same process for the state. Above we recommended that the state allocate personnel to review data on school-level progress on a yearly basis. Here we suggest that these data be used to examine state-level policies in regard to holding districts accountable for their schools' performance and in

support of struggling schools to see what is working well and what needs to be revised. Given the high-stakes environment and urgency to improve student outcomes, the state itself should establish more mechanisms to review policies regularly, assess what components of its policies are on the right track, and adjust policies on an ongoing basis as needed.

9) The state should consider methods to better align the state and federal accountability systems. The state should take into consideration that stakeholders in low-performing schools generally report that while they consider the API to be a better outcome measure, they feel pressure to address AYP targets.

Given the conflict and confusion associated with two overlapping accountability systems, we recommend that the state focus further on their alignment. We recognize this is not an easy task and should be done carefully to preserve the most effective aspects of the state system. Potential actions the state can take include examining better ways to align the expectations and associated sanctions of the state model with the federal model. For example, the state could focus on having the *same* schools identified under *both* programs for similar sanctions. Additionally, the state can continue to press the federal government for greater incorporation of the state API into the AYP measure. Since there appears to be greater buy-in to a growth model and to the API measure in general, this seems an important aspect of the state model to preserve.

SAIT Process Findings and Recommendations

At this early point, we are unable to assess the actual effectiveness of the SAIT process on improving student outcomes. While a substantial percentage of SAIT schools met their growth targets in the first year of participation (2003-04), we have minimal evidence at this point to confirm a link between these outcomes and the SAIT process. Findings from our investigation of SAIT implementation include:

1. While the SAIT process was reported to provide a basic infrastructure for school improvement, respondents noted that the process did not always provide the intensity and focus on factors such as leadership and instructional quality necessary for systemic change within the schools. In addition, several key factors were reported as missing from the essential program components (EPCs), including a focus on English Learners and students with disabilities. In addition, some SAIT providers reported that the exclusion of a school climate/culture component was problematic in some schools, and others discussed the complications associated with the intervention requirements for secondary schools.

2. The level of service provided by SAITs appeared to vary broadly. Many SAITs reported providing intensive coaching and other supports, while a few focused only on monitoring the implementation of the nine EPCs.

3. Several SAIT providers reported that their ability to assess and monitor progress within the SAIT process was hindered by their lack of access to classrooms.

4. Despite clear expectations for a large district role within the SAIT legislation, participation of districts in the process was reported as variable. While some districts appeared to participate actively in supporting the SAIT in their work, others were reported to create barriers to change.

5. School-level respondents were unclear about the sanctions they would face if they did *not* improve through the SAIT process. These unclear expectations were reported to create increased anxiety and confusion surrounding the process, and in some cases, decreased motivation to improve.

Recommendation: **The state should engage in ongoing assessment of the most successful and effective methods for realizing school improvement within the state monitoring process.**

It is still too early to know whether the current SAIT process will be effective in the aggregate in improving schools that failed through the II/USP process. Regardless, SAITs are the critical "next step" after II/USP in attempting to turn around these continuously low-performing schools. School respondents' perceptions were mixed regarding the helpfulness of SAITs. In addition, it is reasonable to expect that whether SAITs are largely successful in the aggregate or not, some SAITs will be more successful than others, and some strategies will be more effective than others. In light of the importance of this work, we recommend an ongoing and systematic data collection and assessment of the effectiveness of individual SAIT providers. In addition, the state should regularly assess the extent to which SAIT schools are improving.

We also recommend that the state consider several additional issues in relation to SAIT implementation for further research. In particular, the EPCs should be examined further to consider the addition of components that address the instruction of English learner and special education populations and the enhancement of school climate and culture. In addition, we recommend that the requirements for intervention program participation be closely examined for adverse impact on enrollment in core courses at the secondary level. We also suggest that the list of EPCs be bolstered with specific strategies and guidelines that enhance the quality of implementation of the EPCs, and that can be used in conjunction with the components to lead to greater success.

In addition, we recommend that the state reconsider allowing classroom observations for SAIT members to better enable assessment of schools' progress. Finally, we advise the state to clarify the next steps to address schools that fail to improve through the state monitoring process. These plans should then be clearly communicated to reduce confusion and anxiety surrounding the reform efforts.

Table of Contents

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

References

Appendices

List of Exhibits

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 1. Study Overview

Introduction

Over the past decade, the standards-based reform movement in education has focused attention on the responsibility of schools and school systems to produce results in student learning. Nationally this focus is most evident with the implementation of accountability provisions of the federal No Child Left Behind (NCLB) Act of 2001. However, even prior to NCLB, many states and districts had instituted their own performance-based accountability programs that aimed to improve student learning, particularly in low-performing schools, and to provide incentives (both positive and negative) for schools and districts to improve student outcomes. In 1999, the California legislature approved the Public Schools Accountability Act (PSAA), which incorporated three central components designed to hold schools accountable for improving student outcomes. The initial PSAA components were the Academic Performance Index (API),[1] the Immediate Intervention/Underperforming Schools Program (II/USP), and the High Performing/Improving Schools Program (also known as the Governor's Performance Award (GPA)).[2] Later, the High Priority Schools Grant Program (HPSGP) was incorporated into the PSAA legislation. The state now faces the challenge of integrating these components with the federal NCLB Act.

In June 2003, AIR completed the first legislatively mandated study of the PSAA, with support from Policy Analysis for California Education (PACE) and EdSource.[3] In November 2004, AIR was contracted by the California Department of Education to conduct a continuation study of the II/USP component of PSAA. This report describes this study and presents its findings. Although II/USP has essentially ended, examining this policy retrospectively provides important information that can be used to shape current and future accountability programs and legislation in California and other jurisdictions. In addition, with the implementation of NCLB, it is critical for California policymakers to understand what has worked well, and not as well, within the state system as they consider ways to continue implementation of the state accountability system and to resolve differences between the state and federal programs.

In this chapter we provide an overview of our findings from the first evaluation, an overview of the research questions and purposes of this continuation study, background on the PSAA legislation, and a brief summary of findings. In the second chapter, we provide the methodology for the study. In Chapters 3 through 5 we present our primary findings, starting with the impact of II/USP on student achievement (Chapter 3). In Chapter 4 we assess factors that led to student achievement growth in selected II/USP schools, and in Chapter 5 we delve into findings to date regarding the School Assistance and Intervention Team (SAIT)

[1] The API is a composite scale used to measure the academic performance and growth of schools. The API is calculated from individual student test scores, and will be discussed in greater depth in the section entitled *PSAA – Legislative and Programmatic Background*.

[2] The GPA program provided rewards for schools meeting their API growth targets.

[3] The final report from this study (which we refer to in this report as the "2003 PSAA Evaluation Report") can be found at http://www.air.org/publications/pubs_ehd_school_reform.aspx

process. In Chapter 6, we conclude with a discussion of major findings from the study and their implications for policy and further research.

The 2003 PSAA Evaluation

The first legislatively mandated study of PSAA, concluded in June 2003, addressed the impact of II/USP and the GPA, as well as factors that contributed to or hindered achievement growth in participating schools. This study utilized a multi-level, mixed-method approach that involved the analysis of achievement trends from II/USP and relevant comparison schools across the state; the administration of surveys to district administrators, external support providers, and teachers and principals of II/USP, GPA, and comparison schools; and in-depth data collection at 21 case study schools across the three cohorts of II/USP. This first study provided the basis from which the current continuation study grew. The conceptual framework developed during the first study served as a foundation for the design, data collection, and analysis efforts for the follow-up study presented in this report. Thus, it is important to start by reviewing the primary findings from the first evaluation:

- We found that PSAA successfully focused district and school personnel's attention on student achievement outcomes and low-performing schools. However, there was a reported tendency to neglect subject areas and student developmental needs not captured by the API.

- Against the backdrop of very large increases in student achievement scores in the state, the direct additional contributions of II/USP and GPA to mean achievement across participating schools were negligible. The most consistent pattern was a small positive "bump" in growth for some groups of II/USP schools relative to non-II/USP counterparts in the planning year of the program, which then dissipated after the first year. In addition, we found no significant effect of II/USP participation on a school's likelihood of meeting API growth targets, nor any impact of GPA awards on subsequent API scores.

- Local districts significantly influenced instructional practice and achievement trends in low-performing schools—both II/USP and non-II/USP—and appeared to mediate the effects of II/USP participation. Our case studies revealed that this influence came in large part through instructionally-related policies for all underperforming schools.

- A school's ability to develop a coordinated and coherent instructional program was a key factor in its ability to meet and surpass academic growth targets. Our case study data indicated a strong association between instructional coherence and growth in student achievement, while both case study and survey data revealed substantial variation among schools in their ability to develop a coherent instructional program.

- Neither the threatened severe sanctions of II/USP nor the potential awards promised through the GPA program were salient among school personnel. More important for motivating response were the expected public scrutiny for low performance and (where applicable) direct monitoring from the district.

- Although External Evaluators, districts, and schools implemented the planning year provisions specified in the PSAA, school improvement planning did not necessarily lead to instructional coherence or improved achievement for II/USP schools. It seems likely that this is because the planning process was divorced from implementation in many respects.

In this continuation study, many of these themes still ring true. Through telephone interviews with school and district staff and through analyses of achievement data, we have been able to further confirm and generalize many of these findings, as well as to further examine factors that contributed to growth within selected II/USP schools.

Research Questions

This continuation study involved additional evaluation activities that built on the first PSAA evaluation study. While this study extends the scope of the 2003 PSAA evaluation (as discussed further below), the original six research questions continued to guide our work:[4]

1. What are the impact on, and benefits to, students from a school's participation in II/USP based on:
   - Results of assessments used to determine whether or not schools have made significant progress towards meeting their growth targets per the PSAA law?
   - Results of disaggregated pupil performance data for each of the following subgroups, as specified in the PSAA law? These subgroups include
     - English language learners
     - Pupils with exceptional needs
     - Pupils that qualify for free or reduced price meals and are enrolled in schools that receive funds under Title I, A of the Improving America's Schools Act (IASA) of 1994.
2. What factors contribute to schools meeting or not meeting growth targets under PSAA?
3. How effectively did participating schools, school districts, and other agencies implement the API, the API for alternative schools, and II/USP components of the PSAA law?
4. What gains in student academic performance are realized from the investment of PSAA resources in the II/USP schools relative to comparable non-II/USP schools?
5. What has been the overall impact of PSAA on school and district personnel, parents, community members, and on school and district organization and practices?
6. What unintended consequences have resulted from the implementation of the PSAA?

Overall Focus of the Continuation Study

While the prior study took more of a macro view of II/USP in regard to its overall efficacy and effect on participating schools, this continuation study added a focus on lessons learned in regard to II/USP implementation. For example, we learned from the first study that the overall impact of II/USP on mean achievement scores across participating schools was negligible. We also learned that districts significantly affected achievement trends in low-performing schools, sometimes for the better and sometimes not. In this study, we not only conducted a summative assessment of the effects of II/USP on student achievement, but also attempted to go beyond this assessment by focusing more intently on schools in which positive results were realized during and after II/USP implementation, as well as on locations where this did not seem to be occurring.

[4] Note: the original set of research questions included reference to the effectiveness of GPA. Since we did not study this program further, we have tailored the questions accordingly.

The theory driving this approach is that while the impact of the program may have been negligible in the aggregate, considerable variation in growth underlies these averages. In the previous study we found that some schools were able to seize this opportunity to make local change in ways that resulted in substantially enhanced student performance, while others did not. Beyond knowing that the program did not have a substantial impact in the aggregate, the current study is designed to learn from sites that appeared to prosper during the time of this program. What factors appear to differentiate growth? How might these factors be used to inform the design and implementation of future state policy? Even if the II/USP had no substantial impact in the aggregate, what can we learn from this vast II/USP experiment to inform and assist schools, districts, and the state in their future endeavors to provide sustainable growth among low-performing schools?

To get beyond isolated stories of success and unique circumstances, it was necessary to broaden our sample to a sufficient number of schools to allow the possibility of detecting the trends underlying growth and to distinguish sites making substantial progress during the period of program participation from those that were not. For this reason, we elected to use telephone interviews, which allowed us to include a broader array of schools than would have been possible had we simply repeated the more in-depth site visits used in the first study. Also, rather than randomly selecting sites, sample selection was focused on "growth" schools (schools that consistently met API growth targets or saw high average API growth during and after II/USP participation) and schools that consistently did *not* meet growth targets or had low average API growth during and after II/USP participation. This sample allowed us to differentiate what was occurring in these two sets of II/USP participants.

We also extended our telephone interviews to non-II/USP schools demonstrating consistent growth during the program period whose student and prior performance characteristics were similar to those participating in II/USP. The purpose of these comparison sites was to help us assess the extent to which the observed successful interventions were related to II/USP participation.

We also wanted to learn more about how the School Assistance and Intervention Team (SAIT) process[5] was working for the purpose of informing policies regarding the development of future SAITs and the procedures they might use to assist underperforming schools. In considering the overall telephone interview sample, we assumed considerable overlap between SAIT participants and our sample of low-growth schools. The fact that this alignment was not perfect raised questions about the extent to which SAITs are intervening in the schools needing it the most, and also necessitated a separate cohort of telephone interviews to SAIT schools.

Regarding the district role in II/USP implementation, we were particularly interested in understanding more about district contributions to schools that made progress within II/USP. Although II/USP is largely school-focused, it seems reasonable that districts can be, and arguably should be, major change agents in regard to bolstering performance in II/USP schools. The potential district impact on II/USP participants was also a featured finding of our prior II/USP study. Thus, district telephone interviews were included to learn more about the district role in participating schools experiencing relatively high levels of growth. Exactly what were these districts doing to support their schools during this process and what might

[5] II/USP schools that did not make significant growth in both of the two implementation years were required to enter into a contract with a SAIT, which would "investigate and provide intensive support and monitoring." This process is discussed further in the section entitled *PSAA – Legislative and Programmatic Background* and in Chapter 5.

other districts do to realize similar improvement? What might the state do to bolster district capacity in this regard?

The telephone interview protocol for this study was designed to allow for substantial qualitative responses (i.e., to allow respondents to describe what occurred in their own words), while also being conducive to quantification. For example, if some schools reported that a certain form of instructional approach was a primary factor in their growth, exactly how many schools of what type said this and what was the perceived impact in relation to other things they were doing at this time? Thus, while allowing respondents the latitude to tell their story, we also wanted to be able to conduct trend analyses of exactly what was reported as making a difference. For this reason, we focused on recording and coding the single most important factor in enabling, and the most important factor in inhibiting, successful improvement during the time of II/USP participation. Beyond this, we wanted to know the specific role of II/USP participation in regard to the successful implementation of these strategies, if any.

In short, the telephone interview component of this study was designed to uncover the most valuable lessons from the considerable funds, time and effort spent on II/USP. While II/USP in its current form has ended, the issue it was designed to address (i.e., how to bolster sustainable increases in student performance in underperforming schools) has not gone away, nor is it likely to disappear as a major issue anytime soon. Given this, it seems most helpful to learn as much as possible about this overall experience by disaggregating it into its predominant successes and failures. What can be learned from these schools to assist in forming future education policy in the state?

Additional Purposes of the Continuation Study

Given our overarching focus for the study, we outlined four specific goals or purposes for this continuation study, in addition to the six primary research questions presented earlier. These purposes helped to focus our study design and analysis plans:

1. **Extend analyses begun in the first study in order to observe implementation and achievement patterns over a longer period of time.** Our analyses in the previous study were limited by the timeframe of the II/USP intervention. Two cohorts were just beginning to implement, or were in the process of implementing, their II/USP reforms, preventing us from adequately assessing the outcome of these reforms. By examining these schools two years later, we were able to assess the longer-term impact and sustainability of the reforms that schools implemented through II/USP.

2. **Build on our findings from the first study to deepen our understanding of major influences on instruction and student learning.** To this end, we focused on identifying factors that contributed to or hindered a school's growth within II/USP. This focus addressed the larger question of how the state can better assist and bolster underperforming schools in general (for example, through the High Priority Schools Grant Program). We explored in greater depth our findings from the previous evaluation about factors that affected the improvement of schools within II/USP. Three of these findings stood out as requiring further attention: the district role in school improvement, the influence of instructional coherence, and internal school capacity.

3. **Study new aspects of II/USP.** Since we initiated the first evaluation in December 2001, several modifications were made to the implementation of II/USP. One such

modification was the formation of School Assistance and Intervention Teams. As part of this evaluation study, we examined the implementation and effectiveness of these teams in the school improvement process. This goal was particularly pertinent given our finding in the first study that many schools lacked monitoring or assistance during the implementation years.

4. **Examine areas of conflict and consistency among II/USP and other accountability programs in California.** One challenge the state faces today is to integrate its school accountability program with the federal NCLB Act of 2001. In addition, many schools are currently affected by several California-based accountability programs, including II/USP, HPSGP, and the comprehensive school reform (CSR) program, as well as district-level accountability programs. In our continued study of II/USP we have examined these overlapping efforts and the methods the state and districts have used to resolve the simultaneous implementation of these programs. We have focused primarily on the overlapping implementation of NCLB.

PSAA—Legislative and Programmatic Background

The Public Schools Accountability Act of 1999 marked a critical point in the standards-based reform efforts in California that began back in the 1980s.[6] Prior to the PSAA, the state had taken many steps toward developing and adopting statewide academic standards and related assessments in English/language arts (ELA), mathematics, history/social science, and science. In the fall of 1998, the adoption of these standards was complete. Also in 1998, the State Board of Education selected the SAT-9 (to be taken by all public school students in grades 2 to 11) as the assessment on which schools would be held accountable for student performance. This test was subsequently changed to the CAT/6 in the 2002-03 school year, and the state completed the transition to standards-based assessments in the form of the California Standards Test (CST) in 2002. In this same timeframe, the adoption of aligned instructional materials in all four core subject areas was completed, and well over $1 billion in state funds were allocated for school districts to use to purchase the materials. It was within this environment that the PSAA was passed in 1999. Below we outline the major components of this legislation.

PSAA Components

PSAA originally included three major components: the Academic Performance Index, the Immediate Intervention/Underperforming Schools Program, and the High Performing/Improving Schools Program. The API provides the basis for the growth targets and the identification of schools for programs like II/USP. The API is a numeric index assigned to each school, ranging from 200 to 1000. Initially based solely on the results of the norm-referenced SAT-9 portion of the Standardized Testing and Reporting (STAR) program, calculation of the API has now incorporated the California High School Exit Exam (CAHSEE) and the California Standards Tests in ELA, mathematics, science, and history/social science, and has increased the weight assigned to these standards-based measures (the exact weight depends on the grade span of the school and, to a lesser degree, on the number of valid scores).

[6] For a full description of the political and legislative history associated with standards-based reform and accountability, see the 2003 PSAA Evaluation Report.

The Board of Education set an interim performance API target of 800 for all schools to achieve. This goal has dictated the basis for determining yearly individual school API targets. For a school with an API score below 800, the annual performance target is to grow by five percent of the difference between its base API score and 800. For a school with an API score of 800 or above, the target is to maintain a score of 800 or above. For a school to reach its target, it must also show comparable improvement for all numerically significant ethnic and economically disadvantaged subgroups.

The II/USP was first implemented in the summer of 1999 when schools scoring in the bottom half of the state's schools on the SAT-9 for two consecutive years (1998 and 1999) were invited to submit an application to participate in the program. Cohort 1 included 430 schools, representing a range of grade levels, SAT-9 deciles,[7] and geography. Cohort 2 included an additional 430 schools in the fall of 2000, and 430 were included as Cohort 3 in the fall of 2001. These Cohort 2 and 3 schools had API scores in the lower five deciles and had not met their API growth targets in the previous year.

Schools that participated in the II/USP made the explicit trade-off of additional resources over three years for potential consequences at the end of this period should those resources not result in improved student performance. More specifically, II/USP schools received funds to create and implement an Action Plan for school improvement. II/USP schools were given $50,000 for the first year to develop the Action Plan with the required assistance of a state-approved External Evaluator.[8] They then received funding at a level of $200 per pupil per year to implement the Action Plan. Implementation funding for each school came from one of two different sources: state funds appropriated for II/USP ("Action Plan Schools"[9]) and funds from the federal Comprehensive School Reform Demonstration (CSRD) program, which is now referred to only as the "Comprehensive School Reform" (CSR) program. Additionally, schools were required to fully match program funds with other new or existing monies.

State-funded schools had the following two years to implement their Action Plan, while CSR-funded schools had the following three years to implement a research-based school reform model. Cohort 1 consisted of 350 state-funded schools and 80 CSR-funded schools. The CSR-funded schools in this cohort started implementation in the first year of funding; hence they did not have an External Evaluator. All Cohort 2 schools, however, received state-funded planning grants for the first year, with 47 schools receiving CSR implementation funds in following years and 383 schools receiving state implementation funds. Cohort 3 followed the same process as Cohort 2. (See Exhibit 1.1 for the specific number and distributions of participating schools, and Exhibit 1.2 for the timeline for II/USP cohorts (both state-funded and CSR).) Legislative modifications were made throughout the implementation of the three cohorts. See Appendix B-1 for a table outlining these changes.

[7] Each decile represents 10 percent of all schools. The "first" decile refers to the lowest performing 10% of schools in terms of API. The "tenth" decile refers to the highest performing 10% of schools.

[8] External Evaluators were state-approved individuals or teams from private and public organizations that schools contracted with to develop a plan for improvement, or Action Plan.

[9] This term was coined by the CDE to refer to state-funded II/USP schools, since they were all required to complete an Action Plan in the first year of funding.

Exhibit 1.1: Distribution of schools selected for II/USP by grade level

Cohort	Funding Source	Elementary	Middle/ Junior High	High	Other*	Subtotals	Totals
Cohort 1	CSR	56	13	10	1	80	430
	State-funded	241	65	43	1	350	
Cohort 2	CSR	33	10	4	0	47	430
	State-funded	224	92	67	0	383	
Cohort 3	CSR	10	1	5	0	16	430
	State-funded	283	50	70	11	414	

*Other includes small schools

Exhibit 1.2: Timeline for II/USP Cohorts (State and CSR-funded)

Group	1999-2000	2000-2001	2001-2002	2002-2003	2003-2004	2004-2005
II/USP Cohort 1	CSR	CSR	CSR	(State)		
	Planning	State	State	(State)		
II/USP Cohort 2		Planning	CSR	CSR	CSR	(State)
		Planning	State	State	(State)	
II/USP Cohort 3			Planning	CSR	CSR	CSR
			Planning	State	State	(State)

Notes: 1) "CSR" and "State" represent years that the cohort received implementation funds from either the CSR or state-funded programs, respectively. 2) Schools that did not meet all growth targets were given an additional year of funding from the state, as indicated in parentheses above. 3) Shading indicates the year in which data collection occurs for this evaluation study.

With II/USP resources, schools were expected to identify barriers to student improvement, devise strategies to remove them, and ultimately to show improvement in student achievement. Initially, schools were scheduled to be subject to sanctions if they did not meet their growth targets for both II/USP implementation years. However, those schools that showed some positive growth in either implementation year—but did not meet their growth targets—were given funding for an additional implementation year (a third year of implementation for "Action Plan" schools), and subsequently categorized as "under watch."

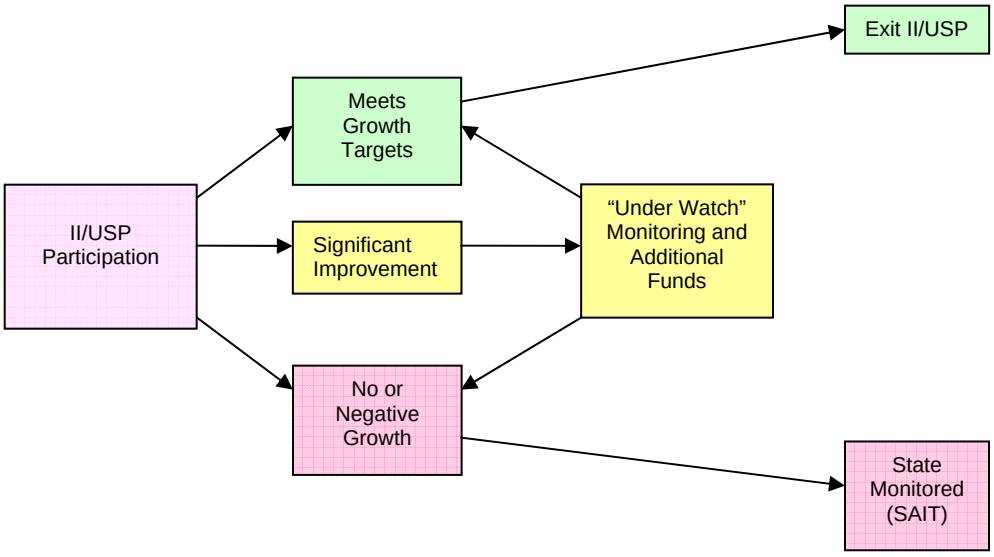
Schools that showed no growth in both of the two implementation years became "state-monitored" schools and were required to enter into a contract with a School Assistance and Intervention Team.[10] SAITs are teams of educational consultants—often retired educators, and other individuals from private companies, county offices of education, and nonprofit organizations—who work with and monitor schools to improve student achievement. The SAIT first assesses whether a school has the "essential program components" (EPCs)[11] necessary for student achievement. In schools where these components are missing, the goal

[10] Schools could also enter state-monitoring if they made some growth during the two implementation years (but did not meet growth targets), and negative or no growth in a year thereafter.

[11] The EPCs vary by grade level, but in general include components such as the adoption of state-board adopted curricula, AB 75 training for principals, and the implementation of an assessment system to monitor student progress. The nine components are discussed in greater detail in Chapter 5.

is to implement them, focusing on resource allocation and benchmarks for student achievement. The graphic below in Exhibit 1.3 shows the various paths schools can take based on their outcomes during the implementation years of II/USP.

Exhibit 1.3: Implementation Outcomes of II/USP Schools



In addition to II/USP, the PSAA included the High Performing/Improving Schools Program, also known as the Governor's Performance Award. This was an incentive program awarding schools that met their growth targets, showed comparable growth among all significant ethnic and economically disadvantaged subgroups, and satisfied testing participation rate requirements. Since no funds have been appropriated for awards since 2002, the GPA program was not further addressed in this continuation study.

Though not directly within the scope of this evaluation, the state's High Priority Schools Grant Program affects some schools that currently participate in the II/USP. Established in 2001 as a part of PSAA, HPSGP targets schools with API scores in Decile 1 (including II/USP Decile 1 schools in Cohorts 1, 2 and 3) and provides planning and implementation funds for schools that apply and are selected for the program. The funding is to be used for raising student academic achievement. II/USP schools selected for HPSGP in 2002 were considered jointly funded and received $200 from II/USP and $200 from HPSGP, for a total of $400 per pupil per year. Jointly funded schools were limited to a maximum of three years of implementation funds. In addition to the years in which they received II/USP funds, Cohort 1 jointly funded schools received one year of implementation funds under HPSGP, Cohort 2 schools received two years of implementation funds under HPSGP, and Cohort 3 jointly funded schools received all three years of implementation funds under HPSGP.

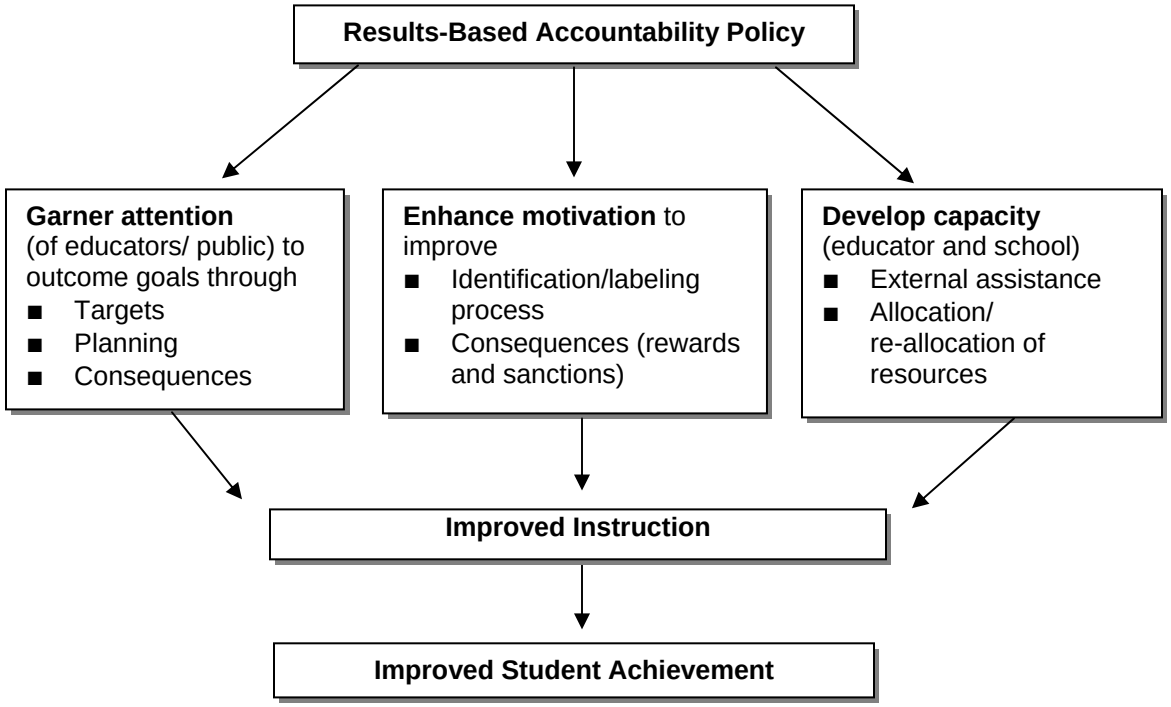
II/USP Theory of Action

Despite programmatic changes and adjustments, the basic elements and underlying assumptions associated with the II/USP legislation have stayed constant through the past six years. As explained in the 2003 PSAA Evaluation Report, II/USP reflects a general model of results-based school accountability. These elements include the following:

- The policy identifies the school as the primary unit of accountability.
- II/USP holds schools accountable for producing specified results on tests of student achievement.
- II/USP provides extrinsic incentives for producing those results.
- The policy extends assistance and/or resources to schools that lack the capacity to do so.

In the 2003 PSAA Evaluation Report, we outlined this general "theory of action" (Argyris and Schon, 1978). A theory of action is a set of related assumptions about the mechanisms and causal relationships through which they work (O'Day, 2002; Finnigan and O'Day, 2003). Exhibit 1.4 shows a simplified graphic for this theory.

Exhibit 1.4: Simplified accountability theory of action model[12]



In our previous report, we also outlined some of the specific details associated with the PSAA model, including the following:

- Since the pool of qualifying schools for II/USP drew from the bottom half of the state, a large number of schools were identified as underperforming—a much greater proportion of schools than has been the case in most prior state systems. The assumption and hope

[12] This graphic was also shown in the 2003 PSAA Evaluation Report.

were that the potential for identification will encourage improvement efforts across this broad range of schools.

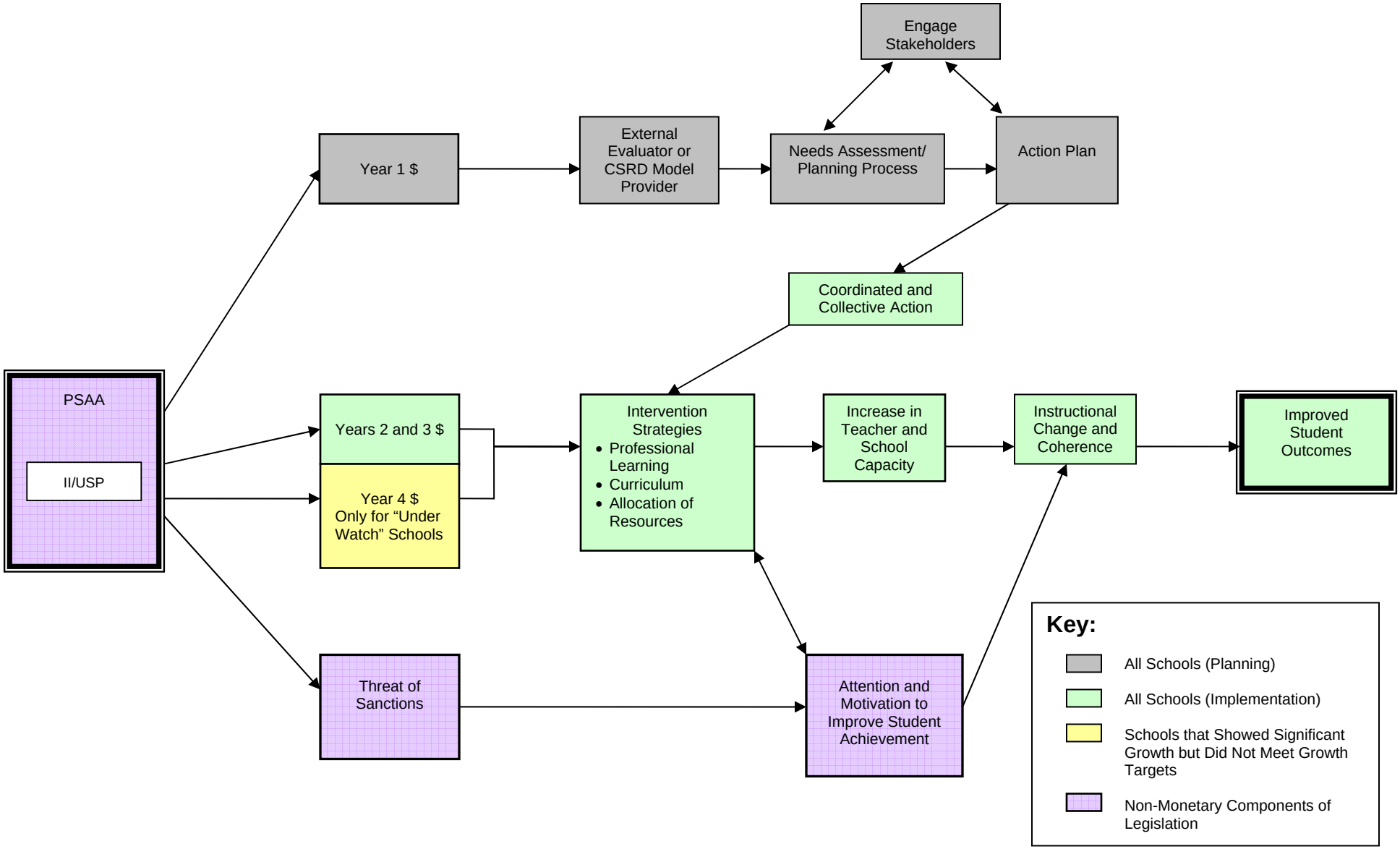
- The API growth targets and subgroup targets will motivate schools to improve student outcomes overall each year, and to do so across multiple subjects and for all students, particularly the lowest-performing students.[13]

- The II/USP program was voluntary. By applying for the program, schools were expected to trade off the possibility of additional resources for the potential for sanctions down the road should those resources not produce the desired improvement.

Other features of the policy that merit particular attention include the heavy emphasis on the role of planning and the expected contributions of "External Evaluators." Taken together, the assumptions behind these and other components of the policy form the overall theory of action shown in Exhibit 1.5.[14]

[13] The API formula rewards growth from the bottom upward more heavily than growth from the middle upward, creating an incentive for schools to provide the most help to pupils with the lowest scores.
[14] This graphic is a slightly modified version of that displayed in the 2003 PSAA Evaluation Report.

Exhibit 1.5: II/USP Theory of Action



As the II/USP has evolved, the policy has incorporated specific consequences for growth or lack thereof during participation in II/USP. Schools that met growth targets throughout II/USP implementation exited the program, schools that made "significant" API growth[15] but did not meet growth targets remained in the program with one extra year of funding, and schools that made zero or negative growth were placed under state monitoring. These consequences are based on several additional assumptions:

- Schools will have implemented sustainable strategies for growth during II/USP participation. Schools that exit the program do not receive additional funds. It is expected that the large influx of funds during the implementation years will be used toward reforms that build capacity and coherence and sustain growth beyond the initial years of implementation.

- Schools that have made some growth, but not sufficient growth, need additional resources and time to improve. A fourth year of funding provided to this subset of schools can be used to provide additional capacity for growth. In addition, the "under watch" status allows for delayed growth that can result *after* the implementation of reforms.

- Schools that are not growing need additional resources, external monitoring, and further assistance and guidance (through the state monitoring process) to improve.

There are also several assumptions behind the state monitoring process itself. The process involves a state-approved team conducting an audit based on the nine essential program components, including the adoption of state-board adopted curricula, AB 75 training for principals, and the implementation of an assessment system to monitor student progress. The nine essential components were developed by the California Department of Education (CDE) and were based on research and literature on school reform. Schools are expected to improve by implementing each of the components. This process is based on several assumptions:

- Schools that are not succeeding are lacking one or more essential components. This is one form of a "gap" model, assuming that once schools fully implement each component, and fill each gap, student results will improve.

- A focus on the instructional program and teacher/principal knowledge and skills will improve student learning. The nine essential components do not include non-academic components such as school climate or facilities.

- Consistent and regular monitoring of progress on the nine essential components by an informed external entity will ensure proper and full implementation of these components.

A detailed theory of action graphic for the state monitoring process is shown in Appendix B-2.

[15] "Significant growth" was defined as at least one API point.

Overview of Study Design

The design for this study was derived from the research questions posed by the CDE, the additional purposes specified for the continuation study, our understanding of the policy theory of action, and a conceptual model we developed to investigate the factors that might influence variation in program implementation and outcomes (discussed in Chapter 2).

We utilized a mixed-method approach that enabled us to examine student achievement across three cohorts of II/USP schools and retrospectively examine factors that led to schools' meeting or not meeting their growth targets while in II/USP. We used quantitative data to assess the relationships between program implementation and outcomes in the aggregate and as a basis for selecting the telephone interview sites. We used qualitative data to explain quantitative findings, identify issues not captured through the quantitative approaches, and uncover strategies and other factors that led to or hindered successful improvement in selected II/USP schools.

To design this study, we took into account the evolution of the PSAA programs, the complexity of the state and district context, and the multiple factors that can affect the implementation of a complicated accountability system with numerous components. Our design included analysis of achievement data as well as responses from key stakeholders associated with the implementation of II/USP, including state policymakers, school administrators, teachers, district staff, and SAIT leaders. The primary design components are summarized further below. Specific methods used are discussed in greater detail in Chapter 2.

Statewide Achievement Trend Analyses

Analyses of student achievement built on our models and findings from the previous study. We used similar analysis models, incorporating data from the 2002-03 and 2003-04 school years, as well as new assessments such as the CAT/6 and CST. The purposes of these extended analyses were as follows:

- To extend the student achievement analyses to include two additional years of data. These analyses provided information on outcomes of II/USP schools after the implementation years. Since we had insufficient data to fully examine Cohort 3 in the previous study, this enabled us to further explore the outcomes for this last cohort.

- To further explore the gap in achievement between various subgroups of students, such as English learners, special education students, and students in poverty.

- To further explore districts' ability to affect student outcomes. A primary finding in our previous study was the influence of the district context, which often overwhelmed any of the influences found from II/USP. We expanded the statewide API and student-level test data analyses by incorporating additional districts and new analysis methods.

Telephone Interviews

In addition to conducting analyses of achievement trends, we selected schools that had made consistent or high growth during II/USP ("growth" schools), and a set of schools that made minimal growth during II/USP. In addition, we selected a small sample of non-II/USP

schools that had made considerable growth during the time of II/USP implementation (non-II/USP "growth" schools) and a subset of schools being assisted by a SAIT.[16] We conducted interviews with various respondents associated with this sample of 40 schools, as outlined below.[17]

School-Level Respondents: We conducted telephone interviews with administrators and teachers within the sampled schools. This data collection effort allowed us to do the following:

- ***Identify factors facilitating growth.*** This provided greater understanding of implementation at the school level, and of the factors that led to schools meeting or not meeting their growth targets.

- ***Examine changes over time***. We sought additional information regarding how schools adjusted to changes in the accountability system over the past two years. In addition, in some cases we were able to explore how schools dealt with the loss of II/USP funds upon completing implementation.

- ***Conduct more in-depth data collection on key findings.*** The telephone interviews provided us with more in-depth information related to factors that contributed to schools' growth under PSAA, as well as the effectiveness of II/USP implementation and its overall impact.

District-Level Respondents. We conducted a series of interviews with 14 administrators from districts associated with our sampled schools. These interviews allowed us to gain additional information on the following:

- The assistance districts provide to low-performing schools and the implementation of district policies and requirements for low-performing schools. We also gathered information on what role the district played in the school improvement process, particularly at II/USP schools.

- How administrators are resolving multiple accountability systems within their districts. Many schools participate in several programs, such as II/USP, HPSGP, and Title I.

School Assistance and Intervention Team Respondents. We explored the implementation of SAITs by including in our telephone interview sample a subset of 10 state-monitored schools that were being assisted by a SAIT. We conducted telephone interviews with the lead SAIT member assisting these schools to gather information about the processes undertaken at these schools, the assistance provided by the teams, and any perceived outcomes resulting from the teams' assistance.

State-Level Respondents. We included interviews with a small set of nine state-level policymakers who were familiar with PSAA and with other accountability programs operating in the state. Through these interviews we gained a stronger understanding of the state's goals in terms of accountability and school improvement. For example, the interviews allowed us to better understand how the state is integrating multiple accountability systems.

[16] We discuss our sampling methods in greater depth in Chapter 2.

[17] We were unable to conduct interviews with all 40 schools. Response rates ranged from 84% to 100% depending on respondent type (see Chapter 2).

Summary of Findings

In Chapter 6, we outline our major conclusions for the study, and provide policy implications and recommendations for ongoing state accountability efforts. Below, we provide a brief summary of these findings:

1. **Overall, the impact of II/USP participation on student achievement has been negligible. Any small advantage experienced by II/USP schools relative to comparison schools during program participation dissipated before or soon after program completion.** While we once again observed high overall growth in student outcomes among *all* low-performing schools, there were only modest differences between II/USP schools and comparison schools. While for some cohorts, in some years, there was a statistically significant but relatively small *positive* effect of II/USP participation, we found statistically significant but relatively small *negative* effects of II/USP participation after program implementation ended, which calls into question the sustainability of the small gains sometimes observed in association with this program.

2. **Despite the lack of an overall program effect, there is evidence that II/USP participation contributed to growth in <u>some</u> schools. Respondents in these schools identified specific factors and strategies they believe led to their improvement.** We found that factors such as school and staff capacity, a coherent instructional program, and systematic use of data to inform instruction contributed to growth among sample "growth" schools. In a limited number of these schools, II/USP was specifically mentioned as contributing to these factors and related strategies by providing additional funding and by facilitating the development of a focus and goals. However, the effectiveness of II/USP was reported to be limited by factors such as late dispersal of funds, and lack of guidance and monitoring during implementation.

3. **Local districts were found to influence the achievement trends in low-performing schools. While district supports were reported as key to some schools' improvement, these supports did not appear to be present in all schools.** As we saw in the 2003 PSAA Evaluation, analyses of achievement scores among several large districts in California indicated that districts influence achievement trends in low-performing schools (both II/USP and similar non-II/USP schools), in that they clearly make better progress as a group in some districts than others. Through our interviews, we found that while district support and policies played important roles in schools' improvement efforts in some districts, there was substantial variation in the level and type of support provided. Some districts were clearly cited as being more helpful than others.

4. **While possibly intensifying the focus on accountability overall, the presence of differing state and federal accountability systems has diffused the attention schools are paying to PSAA.** While some respondents did not report this as a major distraction, saying that their plans to improve student achievement overall should meet the needs of both programs, others reported that the additional layer of achievement targets and expectations associated with NCLB has "confused" or "overwhelmed" staff. The sometimes conflicting measures reportedly prevented schools from having clear targets to work towards. While respondents generally reported that the growth model associated with the state accountability system is a good one, and that API targets were more reasonable for low-performing schools (in comparison to AYP), they often reported focusing more on AYP targets due to perceptions of stricter outcome requirements and more severe sanctions.

In our investigation of the SAIT process, we identified the following trends:

1. While the SAIT process was reported to provide a basic infrastructure for school improvement, respondents noted that the process did not always provide the intensity and focus on factors such as leadership and instructional quality necessary for systemic change within the schools. In addition, several key factors were reported as missing from the EPCs, including a focus on English Learners and students with disabilities.

2. While SAIT providers adhered closely to state requirements for completing the audit and monitoring corrective actions, the role of the SAITs varied among schools. Many SAITs reported providing intensive coaching and other supports, while a few focused only on monitoring the implementation of the nine EPCs.

3. Several SAIT providers reported that their ability to assess and monitor progress within the SAIT process was hindered by their lack of access to classrooms. They reported that their inability to observe instruction limited their ability to assess key issues that may hinder growth.

4. We found variation in the district role within the SAIT process. While some districts appeared to participate actively in supporting the SAIT in their work, others were reported to create barriers to change.

5. School-level respondents were unclear about the sanctions they would face if they did *not* improve through the SAIT process. These unclear expectations were reported to create increased anxiety and confusion surrounding the process, and in some cases, decreased motivation to improve.

Chapter 2. Methodology

This chapter provides an overview of the methodology for the study, including the conceptual framework that guided our instrument development and qualitative data analysis. The methodology for the student achievement analyses is presented in greater detail in Chapter 3 and Appendix A.

Introduction

As described in Chapter 1, this study built on the methods and findings from the 2003 PSAA Evaluation report. Since no new schools have been identified for II/USP since 2001, this continuation study built in additional years of data now available for the first three cohorts of sites. We sought retrospective views regarding implementation as well as the factors and barriers to success, and primarily focused on lessons learned about II/USP and school improvement strategies that could be applied to current and future accountability and school improvement programs.

We had several goals in mind in determining the methodology used in this evaluation. First, we continued analyses from the 2003 PSAA Evaluation to determine the effectiveness of II/USP in the aggregate. However, we also sought a method that would go beyond such a summative analysis of II/USP to learn from sites that experienced positive results from II/USP and from others that did not. This approach was intended to provide information on the factors and specific strategies that had led to success in selected low-performing schools, and the extent to which this success appeared to have been affected by II/USP participation. The policy objective was to attempt to incorporate these factors and strategies for success, as well as strategies for removing barriers, in future state accountability policy. In addition, since II/USP has evolved to incorporate the School Assistance and Intervention Team (SAIT) process, we wanted to specifically examine this component of the policy.

To address these goals, we used a mixed-method approach that allowed us to look across cohorts and to triangulate perspectives and data from multiple actors. Our solution was the approach outlined in Chapter 1, which included the analysis of statewide achievement trends using data from all II/USP and relevant comparison schools across the state. In addition, we conducted telephone interviews with administrators and teachers in schools that had made high and low levels of API growth during and after II/USP implementation, as well as in similar non-II/USP schools that made substantial API growth. We also interviewed district-level administrators associated with our sampled schools and interviewed a small sample of state-level respondents. To explore the SAIT process, we included a subset of SAIT schools in our telephone interview sample, and interviewed the SAIT leader assisting the sampled schools.

Exhibit 2.1 summarizes the data collection methods used and links them with the six research questions that were outlined in the introduction, as well as with the additional evaluation purposes of this continuation study.

Exhibit 2.1: Overview of the data collection strands

Original Research Questions	Student Achievement Analyses	Achievement Analyses – Sub-populations	Telephone Interviews: School	Telephone Interviews: District Level	Telephone Interviews: SAIT	Telephone Interviews: State Level
1. What are the impact on, and benefits to, students from a school's participation in II/USP based on: a. Results of assessments used to determine whether or not schools have made significant progress towards meeting their growth targets per the PSAA law b. Results of disaggregated pupil performance data for each of the following subgroups: i. English language learners ii. Pupils with exceptional needs iii. Pupils that qualify for free or reduced price meals and are enrolled in schools that receive funds under Title I, A of the IASA.	X	X				
2. What factors contribute to schools meeting or not meeting growth targets under PSAA?	X	X	X	X	X	
3. How effectively did participating schools, school districts, and other agencies implement the API, the API for alternative schools, and II/USP (especially the External evaluator provision)?			X	X	X	
4. What gains in student academic performance are realized from the investment of PSAA resources in the II/USP schools relative to comparable non-II/USP schools?	X	X				
5. What has been the overall impact of PSAA on school and district personnel, parents, community members, and on school and district organization and practices?			X	X	X	X
6. What unintended consequences have resulted from the implementation of the PSAA?			X	X	X	
Additional Purposes of Continuation Study						
1. Extend analyses begun in the 2003 PSAA Evaluation in order to observe implementation and achievement patterns over a longer period of time.	X	X	X	X		
2. Build on findings from the 2003 PSAA Evaluation to deepen our understanding of major influences on instruction and student learning.			X	X	X	
3. Study new aspects of II/USP (e.g., SAITs).			X	X	X	X
4. Examine areas of conflict and consistency among II/USP and other accountability programs in California.			X	X	X	X

Overview of the Conceptual Framework

In the previous study, we developed a conceptual framework that guided the development of data collection instruments and guided our data analysis. We continued to use this conceptual framework in developing our data collection instruments and analyses in this study. The conceptual framework was based on II/USP's theory of action (as outlined in Chapter 1) and on relevant literature on school improvement and accountability. The framework summarized the linkages among major constructs we believed were important to explore in the study. Our goal was to understand the research basis for the policy, as well as to identify factors that might attenuate its effects. In this report, we briefly summarize and update the relevant literature.

Relevant Literature

Effectiveness of Accountability Systems

Because results-based accountability systems are now mandated through federal provision, there is a considerable amount of research assessing the effectiveness and implications of these high-stakes systems. The fact that accountability systems are not implemented "in a vacuum" (Bassok, Davis & Raymond, 2004), and coexist as a set of simultaneous systems and reform efforts at multiple levels (federal, state, district, etc.), complicates our ability to assess their effectiveness. While some argue that tests inherently reduce the possibility of reforming our education system (Gunzenhauser, 2003), others argue that they are necessary in order to have some measure of accountability and measurement (Russell, Higgins and Raczek, 2003).

More specific to high-stakes accountability, some evidence has been collected suggesting that these reform efforts are working. Grissmer et al. (2000) assert that NAEP scores validate state evidence of improved achievement in Texas and Maryland, states that, along with North Carolina, established testing and accountability programs in the 1990s. Raymond and Hanushek (2004) also reported that, in the long run, accountability systems implemented in the 1990s that have consequences attached to outcomes have positively influenced student achievement.

While some literature presents cautious optimism about the effectiveness of high-stakes accountability systems, there is also evidence that there is variable impact, depending on prior achievement levels and schools' internal norms of accountability (DeBray, Parson, and Woodworth, 2001; Elmore, 2001) and on initial organizational capacity (Gwynne and Easton, 2001; O'Day, 2002). In addition, evidence is limited on the effects of state-level accountability systems on retention and progression in high school (Carnoy and Loeb, 2002), and concerns have been raised that they initiate inappropriate changes in instruction (Noddings, 2004). This evidence, combined with concerns about the implementation and design of accountability systems, both nationwide and in California, (Mintrop, 2004; Russell, Higgins and Raczek, 2003) raises important policy questions. While some growth in student achievement has been observed, have accountability mechanisms as currently implemented been effective in promoting academic success?

Contributors to Academic Success

The fact that accountability systems have had a variable impact on student achievement comes as no surprise when viewed in the context of school reform. Multiple studies have examined the characteristics of schools that reported high student achievement despite challenging (i.e., high-poverty, high-risk) populations and found several resounding factors that contributed to academic success. When viewed collectively, studies of these effective schools consistently highlight five school characteristics that correlate with high performance as defined by student achievement. These are as follows:

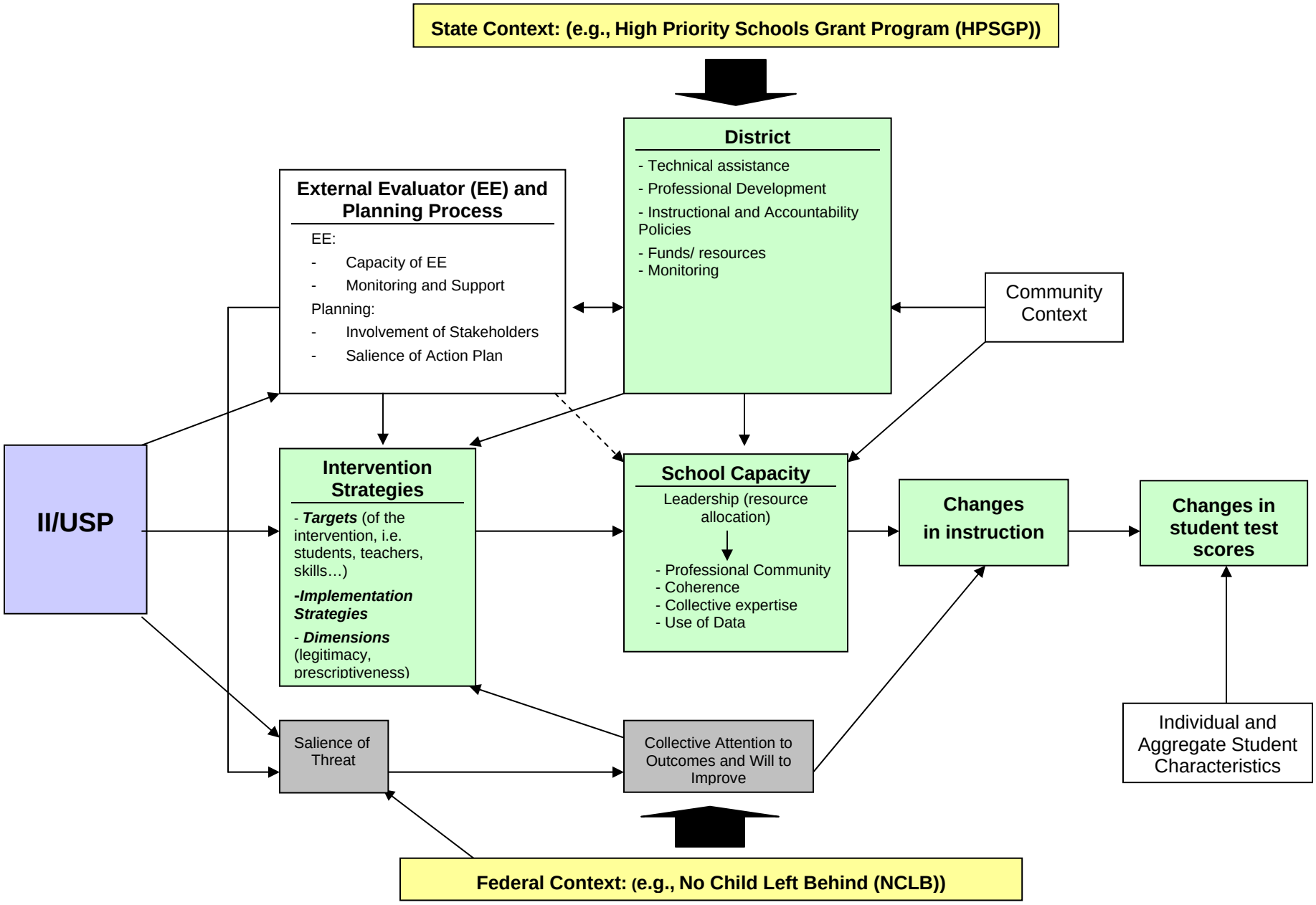
- Monitoring of academic progress (Neumann, 1996; Kannapel and Clements, 2005)
- Strong instructional leadership (Terry, 1996; Muijs, Harris, Chapman, Stoll, and Russ, 2004)
- Shared goals within the professional community (Hoy and Hunnum, 1997; Darling-Hammond, 1996)
- A positive and academically focused school climate (Terry, 1996; Zigarelli, 1996)
- Involved parents in school governance and at home (Desforges and Abouchaar, 2003).

In addition, a recent study in Kentucky (Kannapel & Clements, 2005) had a similar focus to what we attempted to accomplish in this evaluation through our focus on learning from success. They note that "nearly all the worst-performing schools across the nation are high poverty. However, there are enough schools that defy this trend to prove that the background of the student body does not have to determine achievement results." They attempted to "add to the growing evidence on high-performing high-poverty schools by looking closely at the practices" of such schools in Kentucky. They identified elements very similar to those listed above as making the difference for the high poverty, high success schools they visited.

The consistency of findings in these areas must be noted, as it indicates a substantial body of knowledge exists about "what works" to create effective instructional environments. We know less, however, about the exact strategies that underlie these main points, as well as exactly what states and districts need to do to ensure that the strategies are being implemented.

Synthesizing this review of the literature and applying it to our earlier delineation of the revised II/USP theory of action, our research team developed a revised conceptual framework that can be seen in Exhibit 2.2. This conceptual framework guided our qualitative analysis efforts for II/USP. Using this framework, we focused our efforts on delineating the strategies that underlie factors (like those listed above) for "growth" among II/USP schools, in order to inform policy recommendations on how to foster these factors through state and local policy. In addition, we specifically look at the relationship of II/USP to these known elements of success.

Exhibit 2.2: Conceptual Framework



Overview of Methodology

We now turn to an overview of the more specific methods used throughout this study. We have organized this discussion into two main parts. The first outlines the student achievement analyses, and the second describes the methods employed in the telephone interviews.

Student Achievement Analyses

We conducted analyses of student achievement to examine statewide and school-level achievement trends resulting from the implementation of the II/USP program. Our analyses used statewide Stanford Achievement Test (SAT-9) results in reading and mathematics from 1998 through 2002, California Achievement Test (CAT/6) results in reading and mathematics from 2003 to 2004, and California Standards Test (CST) results in English language arts and mathematics from 2001 to 2004. We also used school-level API scores from 1999 and growth scores from 2000 through 2004 to create "synthetic" APIs for analysis. We analyzed the data to look for significant impacts of II/USP participation for three cohorts of II/USP elementary, middle, and high schools. We pursued five primary analytic approaches for the analysis of II/USP schools. The methodologies are described briefly below and are outlined in greater detail in Chapter 3 and in Appendix A.

Statewide API Analyses

A key analytic challenge to investigating trends in API scores over time is the change in API formulation over the years of STAR administration. After including only SAT-9 scores for the 1999 and 2000 APIs, the state then began including data from the CAT/6 and the CST. In any given year, two APIs were calculated by the state—one for growth from the previous year (using the previous year's API formulation) and another to serve as a baseline for the next year (using the new API formulation). Because of these changes, the API was not comparable across the five consecutive years of interest. To overcome this problem, we developed a synthetic API score based on the 1999 API and the growth scores for each year. (This technique is explained in greater detail in Appendix A-2.) We then utilized hierarchical linear modeling (HLM), controlling for school-level variables known to contribute to achievement results (such as average level of parent education), to analyze performance trends as measured by these synthetic APIs in II/USP and comparison schools (see Appendix A-1 for a description of the methodology).

Statewide Student-Level Analyses

To obtain greater precision in contrasting II/USP and comparison schools, a second analysis was conducted using the STAR tests, which allowed us to control for variation in factors known to contribute to achievement (such as parent education) at the student level (rather than just the school level). We utilized a similar HLM approach to analyze the trends in mathematics and reading scores in II/USP and comparison schools. A significant challenge for this evaluation was posed by changes in the available test scores. The SAT-9 was replaced by the CAT/6 in 2002-03 (see more about the change in the test in the *Analysis Strategy* section of Chapter 3), and the CST was introduced in 2001. Once again we looked for significant impacts of participation in II/USP for each of the three cohorts of schools at the elementary, middle, and high school levels. We used the same comparison group as that used for the API analyses.

Statewide Student-Level Analyses by Sub-population

As specified in the research questions above, further analyses of student-level scores were conducted using disaggregated data for the following sub-populations: English learners (ELs), special education students, and students eligible for free or reduced price lunch in Title I schools. We looked for:

- Significant differences between the performance of these subsets of students and all other students over time within II/USP and non-II/USP schools (i.e. trends in the achievement gaps within each category of school).
- Significant effects that participation in II/USP had on these groups of students in comparison to all other students in II/USP schools (i.e., the effect of II/USP on changes in the gaps).

Analysis of Time Needed to Reach CDE Growth Targets

We analyzed the number of years it takes low-performing schools to achieve their growth targets as established by the CDE. To reach the specified target, the school as a whole must score at or above the specified API level, with sub-populations meeting comparable improvement growth targets.[1] For those schools not meeting their target during the planning year, we considered how many years later they met their growth target.

Analysis of District Influence on School Performance

In this analysis, we explored two important questions associated with potential school district influence on low-performing schools. First, we analyzed whether II/USP and non-II/USP schools showed different test score trajectories within school districts to assess whether II/USP participation appeared to make a difference in some districts, but not in others. And second, we analyzed differences in performance trajectories *across* districts to assess whether some districts appear to do a better job than others in improving the academic performance of their low-performing schools over time.

Telephone Interviews

Our second primary methodology was a set of telephone interviews with administrators, teachers, SAIT members, and district staff from a sample of 40 schools. Our purpose was to explore strategies and factors that contributed to or hindered growth in underperforming schools, to expand on findings from the previous evaluation regarding the role of II/USP, and to learn about the SAIT process. More specifically, this data collection effort allowed us to do the following:

- To learn from sites participating in II/USP, and comparable sites not in the program, that experienced a high level of achievement growth, as well as II/USP sites not achieving growth. This effort allowed us to identify strategies and factors that appeared to differentiate success, the perceived effect from II/USP, and to consider how these strategies and factors could inform the design and implementation of

[1] The following subgroups must achieve at least 80 percent of the schoolwide target: African American or Black (not of Hispanic origin), American Indian or Alaska Native, Asian, Filipino, Hispanic or Latino, Pacific Islander, White (not of Hispanic origin), and socioeconomically disadvantaged. However, while the school as a whole must meet the specified target, this lesser standard for these subgroups will likely only apply to schools in which they are relatively small in number.

future state accountability policy. Examining both schools that made high and low levels of growth allowed us to differentiate what was occurring in these two sets of II/USP participants.

- To learn more about how the School Assistance and Intervention Team (SAIT) process was working in order to inform future state efforts.
- To understand more about the district role in relation to underperforming schools.

Sample Selection Criteria

In order to explore factors that contributed to schools meeting or not meeting their growth targets, we used historical data on schools' API scores as the primary criteria for school selection. Specifically, we chose three subsets of schools:

- Schools that had consistently met their growth targets and/or made high average growth since II/USP implementation ("growth" schools)
- Schools that had failed to meet their growth targets and/or made low average growth since II/USP implementation ("low-growth" schools)
- Schools that were participating in the state monitoring process, guided by a School Assistance and Intervention Team (SAIT schools)

In addition, since we saw evidence of a district effect on school progress in the previous II/USP evaluation, we limited the sample to schools within 16 districts, which enabled us to conduct interviews with district staff associated with most of the sampled schools. To ensure that the sample was representative of the pool of II/USP schools as a whole, we used additional primary criteria for selection, including II/USP cohort and grade level. We also considered several secondary criteria, including geographic location, urbanicity, and participation in CSR and HPSGP. We describe each of the primary criteria in greater depth below.

API Growth: Schools were selected for the phone interview sample primarily based on whether or not they met CDE growth targets from the first II/USP implementation year and beyond. We first chose a sample of "growth" schools that met CDE schoolwide and comparable improvement API growth targets every year after entering II/USP. Preference was given to schools that had not met CDE growth targets prior to entering II/USP, in order to target schools that experienced a change in performance levels corresponding to entrance into II/USP. With Cohort 1 schools, very few schools met the CDE growth target criteria for all years after entering II/USP, and therefore six additional Cohort 1 schools (3 elementary, 3 secondary) were selected due to high average API growth since entering II/USP. Within this sample, we also included four non-II/USP growth schools. This allowed us to look at factors for growth that could apply to all schools, not just those participating in II/USP.

The low-growth school sample consisted of schools that failed to meet both CDE growth targets each and every year after entering II/USP. Two of these schools were participating in the SAIT process.

We then selected an additional eight SAIT schools, targeting those that were in the lower deciles and had made minimal growth during II/USP participation. All SAIT schools in our

sample began the SAIT process in either 2003-04 or 2004-05. We constrained our sample in this way to ensure that we were examining the SAIT process in its current form.[2]

We constrained our sample of growth schools to those in deciles 1-3 at the time they entered II/USP. In addition, our SAIT school sample was constrained to schools in deciles 1-3 at the time they began the SAIT process. These constraints allowed us to look at factors for growth and challenges faced among the schools most in need of improvement.[3]

Cohort: Inclusion of schools from each of the three cohorts provided perspectives on three different stages of II/USP implementation. This resulted in a richer understanding of to what extent schools have sustained, and plan to sustain, the reforms intended by the PSAA. In addition, since modifications were made to II/USP administrative procedures each year (see Appendix B-1), it was important to include schools that represented each iteration of the II/USP selection, planning, and implementation processes.

Grade level: We also included elementary, middle, and high schools within each cohort of case study schools. Since curricular programs, among other factors, differ considerably across levels of schooling, it was important to include representative schools from each category. We included a larger number of elementary schools than middle and high schools to reflect the larger number of elementary schools in the state and in the program.

Sample Characteristics

The sample included 40 schools selected from 16 school districts, proportionately stratified for region and urbanicity according to the distribution of II/USP schools in the state of California. Similar consideration was applied to stratify the sample by school type and II/USP cohort, including CSR schools.

Twenty-two II/USP growth schools, six II/USP low-growth schools, eight (additional) SAIT schools, and four non-II/USP growth schools were selected. Fifteen of the growth schools were selected based on meeting their API growth targets consistently after entering II/USP, and seven were selected due to making high average API growth during and after II/USP participation. In some of the categories in Exhibit 2.3 below, we were unable to locate schools that met our criteria, either for growth or low-growth samples. Thus, the numbers below are not consistent across cohorts. The exhibit shows the numbers of schools in each selection category.

It is important to note that two of the low-growth schools also participated in SAIT. They therefore satisfied both the criteria for schools that made low growth during II/USP, and the criteria for the SAIT sample. For simplification, we classify them as low-growth schools in Exhibit 2.3 below, but show the full number of SAIT schools in parentheses.

[2] In 2003-04, the SAIT process changed to incorporate a focus on nine essential program components. The CDE requested that our sample incorporate only schools undergoing this current SAIT process.

[3] The sample of II/USP schools was selected from a candidate pool of 1,001 II/USP schools in the state of California. While there are a total of 1,290 II/USP schools across all three cohorts, 289 Cohort 1 schools were removed to ensure consistent eligibility criteria across all cohorts. In other words, schools were removed that had met estimated growth targets in 1998-99 based on our synthetic API, which was the additional eligibility requirement for Cohorts 2 and 3.

Exhibit 2.3: Sampled Schools

Elementary Schools (22 schools)

	Cohort 1	Cohort 2	Cohort 3	Non-II/USP	Total
Growth Schools	4	4	4	2	**14**
Low-Growth Schools	-	2	1	-	**3**
SAIT Schools	-	5	-	-	**5**
Total	4	11	5	2	**22**

Middle/High Schools (18 schools)

	Cohort 1	Cohort 2	Cohort 3	Non-II/USP	Total
Growth Schools	4	3	3	2	**12**
Low-Growth Schools	1*	1*	1	-	**3***
SAIT Schools	1 (2*)	2 (3*)	-	-	**3 (5*)**
Total	6	6	4	2	**18**

* The Cohort 1 and Cohort 2 Low-Growth Schools are also SAIT schools, and are therefore present in both cells.

The actual sample selection occurred through several iterations. Gaining access to schools and districts proved to be a challenge due to the pressures districts and schools were experiencing with the competing demands of internal projects, research studies, testing schedules, and other activities. We substituted two schools from one district that declined participation, and we also substituted two schools that declined participation early in our data collection. We selected each substitute school carefully, attempting to ensure comparable program status, school level, demographics, geography, and urbanicity.

The final sample of 40 schools included the following:

- 5 CSRD, 17 HPSGP schools
- 22 elementary, 9 middle, 9 high schools
- 10 Cohort 1, 17 Cohort 2, 9 Cohort 3 schools
- 10 northern, 6 mid-, 24 southern California schools
- 8 different SAIT organizations, 9 different SAIT lead members
- 29 urban, 9 suburban, 2 rural schools
- 23, 7, 4, 3, and 3 schools from Deciles 1 through 5, respectively (based on year that school began participation)[4]

Data Collection

Interviews were conducted in spring 2005, with the intention of including a principal and a teacher at each of the sampled schools. We also interviewed a district staff member at each of the districts, and the SAIT leader for each SAIT school. After receiving permission from each

[4] For non-II/USP growth schools these rankings were based on 1999 deciles.

district, we contacted the administrators directly to set up the interviews. To choose teachers for the study, each data collector asked the interviewed principal for the names of four teachers who had been at the school for at least three years, to ensure that the teacher would have knowledge of the reforms implemented during II/USP participation. One teacher was then randomly chosen from the list at each school and contacted by the data collector to set up an interview.

Each interview was approximately 45 minutes to one hour long. In order to ensure that we could gather consistent and comparable data across sites, in addition to open-ended questions, we also incorporated questions that could be quantified during analysis.

The interview for principals and teachers primarily focused on factors that led to or hindered growth in the schools. Using our conceptual framework as a guide, we developed a framework of possible responses. Using these potential responses, we developed sets of more detailed probing questions for each response type. To develop additional questions about II/USP implementation, the SAIT process, and the district role in reform, we specified key variables at district, school, and classroom levels, reviewed existing protocols, and pilot-tested the instruments with a small sample of three principals and three teachers. Appendix C includes copies of all instruments developed for this study.

Separately, we interviewed one to two district administrators knowledgeable about II/USP and the state monitoring process in each district in the sample. In addition, we interviewed the lead SAIT provider at each SAIT school in the sample. Finally, we interviewed a small set of state-level respondents with extensive knowledge of II/USP.

All interviewers used the developed protocols to guide interviews with school and district staff and (with permission) audio-taped interviews to ensure accuracy of notes. Respondents participated voluntarily and received a modest honorarium for participation.

Interview Response Rates

To calculate response rates for each respondent type, we divided the total number of interviews completed by the total number of interviewees contacted for an interview. This information is shown in Exhibit 2.4. Response rates were over 80 percent for all categories of respondents.

Exhibit 2.4: Response rates for phone interviews by interview respondent type

	# Interviewed	# in Sample	Response Rate
Principals	34[†]	40	85%
Teachers	32	38*	84%
District Administrator	14	16**	88%
SAIT Leader	9	9	100%
State Respondent	9	9	100%

[†] At four schools we interviewed another school administrator instead of the principal. These school administrator interviews are counted in the tally of principal interviews.
*This response rate was calculated based on the number of teacher contact names we received (38). At two schools, the principal refused to provide teacher names.
** One district was a small district where the principal also served as a district-level administrator. This interview is counted in the tally of principal interviews.

Although we were unable to interview all respondents, these response rates were considerably higher than response rates typical of written surveys. We found during repeated calls to set up

interviews that respondents are overburdened with research studies and excessive demands in general. In particular, schools in II/USP tend to be very busy with additional work focusing on school improvement.

Exhibit 2.5 breaks out the response rates by school level. It shows that 78 percent of middle and high schools had at least one interview, compared with 95 percent for elementary schools.

Exhibit 2.5: Percentage of schools where at least one respondent was interviewed, by type and level of school

	Total Number of Schools	% of Elementary Schools Interviewed	% of Middle Schools Interviewed	% of High Schools Interviewed
Growth School*	26	93%	60%	71%
Low-Growth School**	6	100%	100%	N/A
SAIT School	8	100%	100%	100%
Total	**40**	**95%**	**78%**	**78%**

*Growth schools include II/USP and non-II/USP schools.
**There are two schools that were both low-growth and SAIT schools. For the purposes of this table we have counted them as low-growth schools.

Analysis of Interview Data

The study's conceptual framework guided our analysis of interview data. Using detailed notes and audio tapes from each interview, the data collectors completed an analysis matrix by filling in information related to the primary constructs in the conceptual framework. Summarizing the data in matrix form allowed the research team members to identify patterns and evidence along the dimensions of the conceptual framework. To facilitate the identification of themes and patterns, we held ongoing analysis meetings with the data collection team to discuss individual constructs. We used the student achievement analysis results and prior research in the field of school accountability to inform our discussions and to frame these analyses.

Chapter 3. Analysis of Student Achievement in II/USP Schools

Overview

In this chapter we examine statewide and school-level achievement trends associated with the implementation of II/USP. The relevant research questions for this section are:

1. What are the impacts on, and benefits to, students from a school's participation in II/USP based on:

    a. Results of assessments used to determine whether or not schools have made significant progress towards meeting their growth targets per the PSAA law?

    b. Results of disaggregated pupil performance data for each of the following subgroups?

        i. English language learners (ELs)

        ii. Pupils with exceptional needs (students enrolled in special education)

        iii. Pupils that qualify for free or reduced price meals and are enrolled in schools that receive funds under Title I, Section A of the IASA.

2. What gains in student academic performance are realized from the investment of PSAA resources in the II/USP schools relative to comparable non-II/USP schools?

Analysis Strategy

An overview of our methodology and new analytic challenges for the student achievement analyses were outlined in chapter 2. In this section, we discuss additional analytic challenges that are comparable to those discussed in the 2003 PSAA Evaluation Report. These include:

- What sample of II/USP and comparison schools should be the focus of the analysis?
- What achievement outcome measures should be used for the analysis?
- How should the effects of II/USP be determined?
- How should missing data be taken into account?
- What school and student background characteristics should be controlled statistically in the analysis?

In the following sections, we discuss each of these issues and explain how our current methodology that was described in chapter 2 compares to that used in our 2003 PSAA Evaluation. (Appendix A-1 provides further details on the methodological approach used to estimate the II/USP effect.)

Identifying the Sample of II/USP and Comparison Schools

Three cohorts of elementary, middle, and high schools participated in II/USP over the course of the program. Details regarding the application procedure and the years in which each cohort participated in II/USP can be found in Chapter 1. It is important to keep in mind the change in eligibility requirements across cohorts. While Cohort 1 schools qualified for II/USP by scoring in the bottom half of the state's schools on the SAT-9 in both 1998 and 1999, Cohort 2 and 3 schools had API scores in the bottom five deciles and failed to meet their growth targets in the year before they applied. The more strict eligibility criteria for Cohorts 2 and 3 resulted in fewer schools being eligible and applying for II/USP than in Cohort 1.

In addition, when selecting the sample of II/USP schools, it was important to account for those schools that concurrently participated in other funding programs such as the Comprehensive School Reform (CSR) program[1] and the High Priority Schools Grant program (HPSGP). These schools differ from II/USP "Action Plan" schools because aside from having additional funding sources (in the case of HPSGP) and different program requirements, the selection criteria for these programs were different than for II/USP. As a result, instead of removing these schools, we have controlled for CSR participation across all II/USP cohorts. In the case of Cohort 3 II/USP, we also controlled for HPSGP participation.

In order to evaluate the effect of II/USP on the academic achievement of a school, it is important to have a comparison group that estimates how II/USP schools would have performed had they not participated in the program. The ideal group would exhibit the exact same characteristics as II/USP schools except for their participation in the program. As was done in the 2003 PSAA Evaluation, we decided that the most appropriate group to use was those schools that had applied for the program but were not selected to participate. In this way, we could avoid any bias arising from a school's decision to apply for the program.[2]

Because the number of schools that applied for Cohort 1 was significantly larger than the number that could be supported by the program, we were able to use the non-participant applicant pool as the comparison. However, for Cohorts 2 and 3, nearly all the schools that applied were accepted; as a result, the excluded group of schools was too small to be used as a comparison group.[3] Thus, it was necessary to relax our criteria and use the larger pool of eligible schools as the comparison groups for

[1] Formerly known as the Comprehensive School Reform Demonstration (CSRD) program

[2] This lack of bias only pertains if random selection was used to determine program participants among the applicants. We have been assured by CDE that selection was random subject to two constraints set by the legislation: PSAA required participating schools be distributed equally across the five deciles, and it set limits on the number of elementary, middle, and high schools to be served. Also, some districts reported identifying certain schools within the eligible pool in their district to apply for II/USP. This may introduce some bias in the participant pool.

[3] For Cohort 2, a total of 505 schools were used as a comparison. From those, 98 schools had applied to II/USP. For Cohort 3, 833 schools were used as comparison, from those, 318 had applied to II/USP.

Cohorts 2 and 3. This distinction between cohorts in the selection of comparison schools is important to keep in mind. The II/USP effect may be biased upward in Cohorts 2 and 3 because the comparison group includes schools that may have characteristics associated with not applying for II/USP, such as less motivation.

Another consideration when selecting the comparison group is that some of these schools subsequently participated in II/USP. Leaving these schools in the comparison group would likely bias the comparative results downward because these schools were now receiving additional funding and support not received by other comparison schools. However, removing these schools from the comparison group might also bias the results because these schools, by definition, were poor performers in the year that made them eligible for II/USP. Thus, eliminating them would potentially underestimate the II/USP effect. In the 2003 PSAA Evaluation, these schools remained in the comparison group because an estimation procedure to adjust the achievement outcomes for these schools showed that there was little impact on the size or significance of the relationship between II/USP and school outcomes.

However, given the additional years of data that are now available and included in our analyses, leaving these schools in the comparison group presents a more serious problem. For example, a comparison group for II/USP Cohort 1 could have started participation in II/USP Cohort 2 in 2000. This school would be considered a comparison school based on the 2003 PSAA report methodology even though they now will have received II/USP funds for three years, spending only one year without II/USP funds. As a result, we decided to remove a school from the comparison group in the year that they begin to receive II/USP funds. This method minimizes the bias by not removing the poor performers from the sample until the year that they begin receiving II/USP funds and support, at which time we cease to include them in our comparison group.

Exhibit 3.1: Number of Comparison Schools by Year by Cohort

	1997-98	1998-99	1999-2000	2000-01	2001-02	2002-03	2003-04
Cohort 1	1,000	1,000	1,000	783	602	602	602
Cohort 2	505	505	505	505	424	424	424
Cohort 3	833	833	833	833	833	833	833

As shown in Exhibit 3.1, the number of schools in the comparison group grows smaller in the later years of II/USP for Cohorts 1 and 2. The number of comparison schools for Cohort 3 remains constant. It is possible that the schools that participated in II/USP and were subsequently removed from the comparison group have systematically different characteristics than comparison schools that never participated in II/USP; therefore, removing these schools from the analysis provides a more homogenous comparison group.

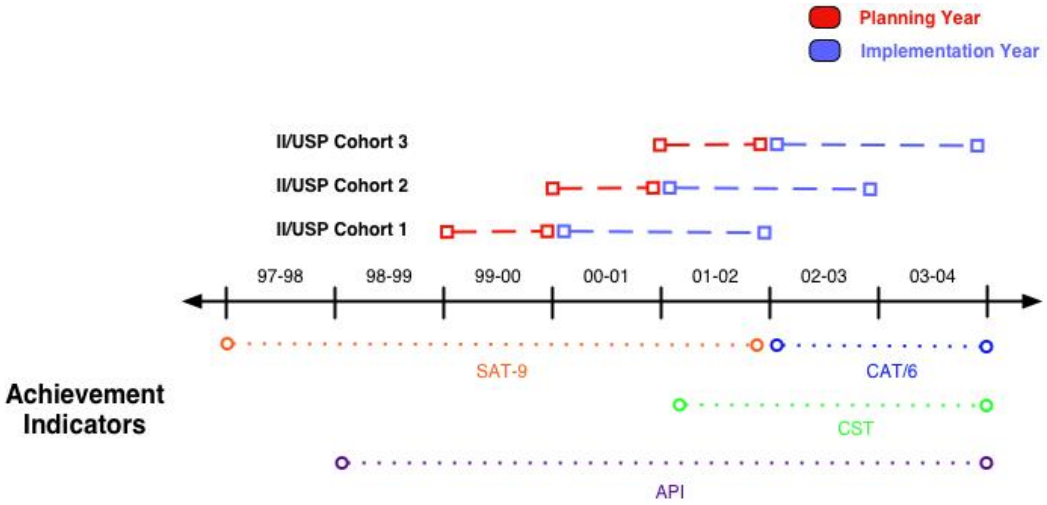
Estimating the Effect of II/USP Participation

In order to estimate the effect of II/USP, we compared the trajectories of student achievement at II/USP schools before, during, and after participation[4] to trends at the comparison schools. Exhibit 3.2 displays the years each cohort participated in

[4] This is true for Cohort 1 and Cohort 2 II/USP schools. In the case of Cohort 3 II/USP schools we do not observe post-II/USP participation data.

II/USP, as well as the outcome measures that are available for each year. The outcome measures will be discussed in further detail in the next section.

Exhibit 3.2: Timeline of II/USP Cohorts and Achievement Measures



Cohort 1 II/USP schools received their initial funding in 1999-2000, which means that there are two years of achievement data from before participation, three years during II/USP, and two years post-participation. In the case of Cohort 2 II/USP schools, there are three years of data before participation, three years during II/USP, and one year post-participation. The availability of post-participation data is unique to this year's analysis and allows us to examine achievement trends after the funding ends for Cohort 1 and Cohort 2 II/USP schools. In the case of Cohort 3 II/USP schools, there are four years of data from before participation and three years during II/USP.[5]

Most accounts report that CDE selection of II/USP schools was random. As a result, we anticipated that II/USP schools and comparison schools would have similar achievement growth during the years before II/USP schools received their initial funding. An examination of Exhibits 3.8 to Exhibit 3.15 in the following section shows that the achievement trajectories prior to II/USP awards do not appear to be different. If II/USP had a positive effect on achievement, we would expect greater growth during the participation years. Furthermore, the post-participation trajectories will indicate whether the achievement trends observed in II/USP schools during the participation years are sustainable after the program ends.

Determining Achievement Outcome Measures

The three outcome measures used in this evaluation are school-level API scores, student-level SAT-9 and CAT/6 standardized test scores, and student-level California Standards Test (CST) scale scores. All these measures meet the requirement of Research Question 1, which states that the impact should be evaluated based on the "results of assessments used to determine whether or not schools have made

[5] Some schools received a fourth year of II/USP funding, but since the primary focus on planning and implementation took place in the first three years, we considered three years to be the standard duration of II/USP participation for our analyses.

significant progress towards meeting their growth targets according to the PSAA law."

As shown in Exhibit 3.2, API scores are available for each year from 1999 to 2004. As explained in the "Data and Methodology" section in chapter 2, we calculated a set of transformed API scores that were based on the official scores for 1999 to 2004, and SAT-9 scores for 1998. These "synthetic" API scores provide a consistent basis to measure growth in achievement, and make it possible to include the 1998 school year in our analyses.[6] See Appendix A-2 for the "synthetic" API calculation.

The second measure used in this analysis is the STAR norm-referenced test at the student level. This measure is available for all seven years included in these analyses (1998 to 2004) for grades 2 to 11 and is accompanied by student background characteristics. However, as mentioned in chapter 2, a significant challenge was posed by the change in the STAR norm-referenced test. This measure changed in 2002-03 from the SAT-9 to the CAT/6 test. The scale scores of these two tests cannot be linked, which prohibits conducting longitudinal analysis of student-level scale scores as was done in the 2003 PSAA Evaluation. To overcome this challenge, we calculated *standardized* test scores for the math and reading SAT-9 and CAT/6.[7] Although standardization allows us to make valid comparisons against a specific benchmark (i.e., the average performance of the state), it does not permit measurement of absolute growth since the average performance of the state is set to zero every year. More details on how to interpret the results of standardized test scores can be found in the results section of this chapter.

Finally, we analyzed student-level math and English language arts (ELA) CST scale scores for Cohort 3 schools. CST scores are available for 2002, 2003, and 2004, and are accompanied by student level background information.

Accounting for Missing Data

There are quite a few schools that are missing API data for at least one year (the exact number of schools that are missing API data is listed in Appendix A-3). Cohort 3 has the highest percentage of schools with missing data, with approximately 20 percent of II/USP schools and 18 percent of comparison schools missing API scores for one or more years. For Cohort 2, 12 percent of II/USP schools and 17 percent of comparison schools were missing API scores. And for Cohort 1, 18 percent of II/USP schools and 12 percent of comparison schools were missing II/USP scores for one or more years.

There are many reasons why a school may not have an API score in a given year. For example, it may not have been included in the testing program due to small size or other school characteristics, or there may have been an irregularity in test administration at the school. In the 2003 PSAA evaluation, schools were excluded if they were missing API data for at least one year. However, this methodology would result in many more schools being excluded in the current analyses given that more years are included. As a result, in this year's evaluation, schools are dropped from the

[6] No statewide test score existed prior to 1997-98.

[7] This procedure allows us to quantify the distance (measured in standard deviations) between the mean of a subgroup of schools (II/USP schools, for example) from the average performance of the state.

analysis only in the year that they are missing API data, which means that the treatment group may vary slightly from year to year.

Controlling for School and Student-Level Background Characteristics

Given that there are many factors that influence student achievement, the evaluation statistically controls for an extensive set of school-level and student-level characteristics in order to isolate the effect of II/USP on student achievement (see Appendix A-4 for a list of the control variables used). These student and school background characteristics accompany the outcome variables in the STAR and API databases.

Descriptive Analysis

In this section the status of the different II/USP school cohorts and important demographic characteristics of II/USP and comparison schools are analyzed. Examining the status of the different cohorts provides important information that can be used to inform and interpret the achievement analysis results that will be presented in the next section, *Analysis of the Impact of Program Participation*. Analysis of the demographic characteristics of II/USP and comparison schools provides contextual information that is critical when analyzing the achievement analysis results.

II/USP Status

The current status of all II/USP schools in 2004-05 is presented in this section. Appendix A-5 presents the same information broken out by school level (i.e., elementary, middle, and high school). As shown in Exhibit 3.3, about 75 percent of Cohort 1 II/USP schools have exited the program, and about 15 percent of the schools in this cohort are currently monitored by the state. When looking at the status of II/USP Cohort 2 schools, about 54 percent of the schools have exited the program, about 26 percent are watched, and about 18 percent are monitored. Across all cohorts about 2.5 percent of the schools have been closed. When analyzing the status across cohorts it is important to keep in mind the varying stages of implementation the cohorts are in. For example, when looking at the status by school level (see Appendix A-5), 65 percent of elementary-level Cohort 2 II/USP schools have exited the program, compared to about 26 percent of high schools.

Exhibit 3.3: II/USP Schools Current Status (n = 1,290) as of 2004-05

II/USP Cohort	Exited		Watched		Monitored		Closed	
	N	%	N	%	N	%	N	%
Cohort 1	319	74.5	33	7.7	66	15.4	10	2.3
Cohort 2	230	53.7	111	25.9	75	17.5	12	2.8
Cohort 3	102	24.5	300	71.9	5	1.2	10	2.4
Total	**651**	**51.1**	**444**	**34.9**	**146**	**11.5**	**32**	**2.5**

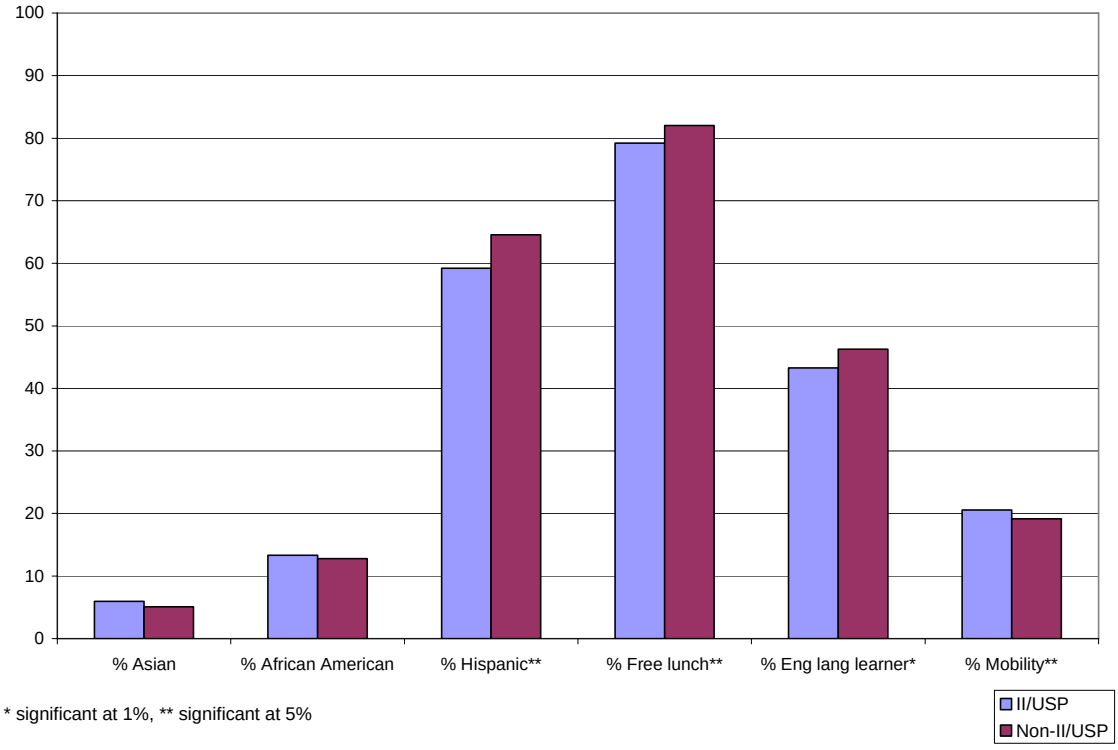
Demographic Characteristics of II/USP Schools

Many factors can influence student achievement every year, in addition to II/USP participation. Given that II/USP schools' performance is compared against similar schools, it is important to know if there are significant differences in the demographic

characteristics of these two groups, i.e. II/USP versus comparison schools[8], when interpreting achievement results. In Exhibits 3.4 to 3.6 only statistically significant differences in at least two demographic characteristics are shown. Because the demographic characteristics for middle and high schools across all three cohorts are generally similar, the exhibits below only feature significant differences among elementary schools. (See Appendix A-6 for the demographic characteristics of middle and high schools).

Exhibit 3.4 shows that Cohort 1 II/USP elementary schools enrolled significantly fewer Hispanic students, students eligible for free or reduced lunch, and English learners. II/USP Cohort 1 elementary schools also had a significant, but small, difference in the percentage of students who are in their first year of attendance at their current school (listed as mobility below). It is possible that there is a bias in favor of the treatment group because Cohort 1 II/USP elementary schools have fewer disadvantaged students than comparison elementary schools,. However, pre-II/USP performance trajectories were similar between Cohort 1 II/USP and comparison elementary schools when controlling for demographic characteristics.

Exhibit 3.4: Demographic Characteristics of Cohort 1 II/USP and Comparison Elementary Schools



* significant at 1%, ** significant at 5%

Exhibit 3.5 shows that Cohort 2 II/USP elementary schools enrolled more African American students and English learners. This difference suggests a possible bias in favor of the comparison schools. However, as in the case for Cohort 1 elementary

[8] CSR schools are also included in this analysis.

schools, the pre-II/USP performance trajectories for Cohort 2 II/USP and comparison elementary schools were fairly similar.

Exhibit 3.5: Demographic Characteristics of Cohort 2 II/USP and Comparison Elementary Schools

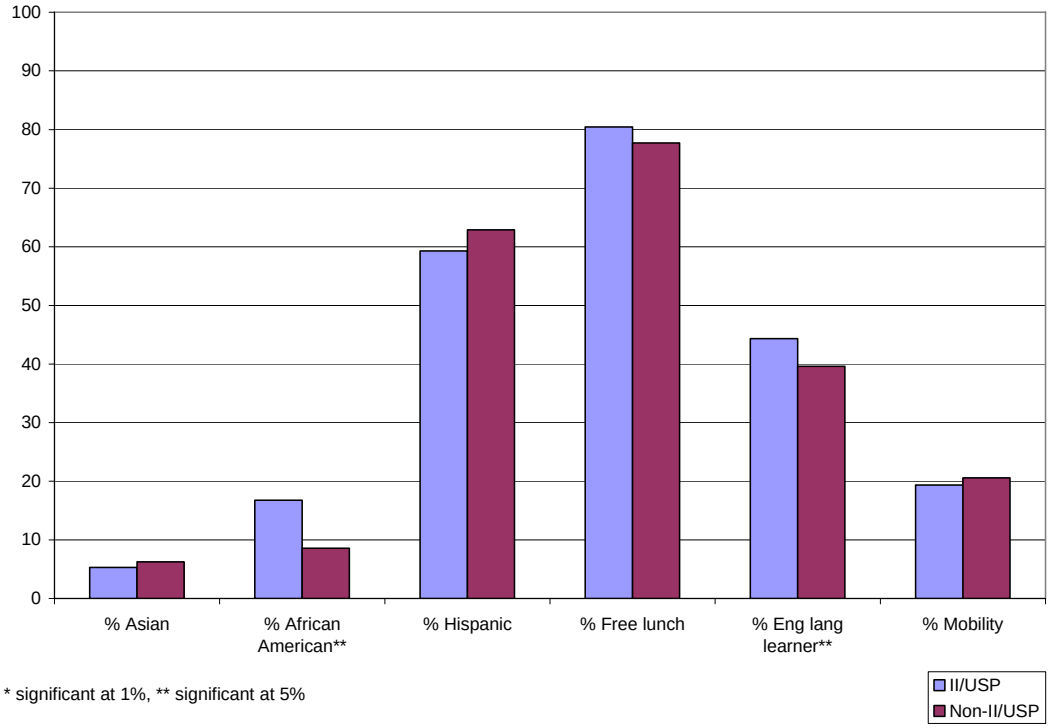
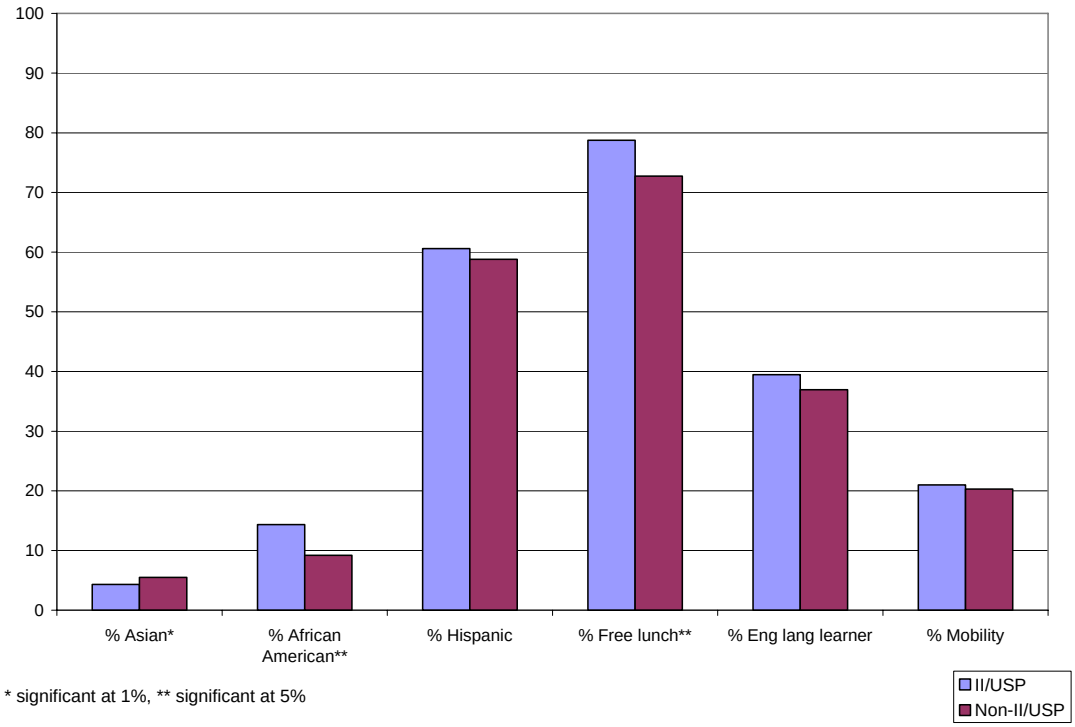


Exhibit 3.6 shows Cohort 3 II/USP elementary schools enrolling more African Americans and students eligible for free and reduced lunch. These differences in demographic characteristics may introduce a downward bias when estimating the II/USP effect on school performance. However, as will be observed in the analysis results, no statistically significant differences are observed between II/USP and comparison schools prior to participation in II/USP after controlling for demographic characteristics.

Exhibit 3.6: Demographic Characteristics of Cohort 3 II/USP and Comparison Elementary Schools



* significant at 1%, ** significant at 5%

Analysis of the Impact of Program Participation

We begin this section with an overview of the main findings regarding the effect of II/USP on schools' academic performance. We will then elaborate on these findings by school type and cohort.

Overview

Exhibit 3.7 summarizes the API, CST, and standardized SAT-9 and CAT/6 reading and math results for II/USP elementary, middle, and high schools by cohort (see Appendix A-7 for effect sizes). The clearest longer term effect of the program can be observed for Cohorts 1 and 2, which are the only two cohorts for which post-II/USP data are available. Cohort 1 and 2 II/USP elementary schools showed a statistically significant but relatively small improvement in student achievement in comparison to the control group during the planning year. Subsequently, they showed either a significant "negative" effect (i.e. less growth than the comparison sites) or no significant effect during the implementation years. In the post-II/USP years, the comparison schools consistently outperformed the treatment sites for these two cohorts.[9]

[9] It is important to keep in mind that a "negative" effect size does not necessarily mean that II/USP school performance decreased relative to the state average; rather, it means that II/USP schools

II/USP middle and high schools showed some statistically significant but relatively small gains in the planning and implementation years of II/USP, but did not outperform the comparison sites after the program ended. The significant negative effect sizes in the post-II/USP years for middle and high schools ranged from -0.026 to -0.06 standard deviations and were typically larger than the positive effect sizes in the implementation years, which ranged from 0.012 to 0.026 standard deviations.[10] This trend indicates that II/USP middle and high schools were not able to sustain the gains made during the program after the funding ended. Overall, no II/USP school, on average, showed a statistically significant positive effect in post-II/USP years.

had smaller gains relative to comparison schools. It is possible that II/USP schools reduced the gap to the state average, but not as much as comparison schools.

[10] The effect size is measured in standard deviations: .25 standard deviations or lower indicates a small effect size, between .25 and .40 is a medium effect size, and .40 standard deviations or above is a large effect size. The literature considers medium or large effect sizes "educationally significant," indicating that meaningful numbers of students are affected (see Cohen, 1969, p. 278-279).

Exhibit 3.7: Summary of Performance Trajectories for Elementary, Middle, and High Schools, by Cohort

	Elementary			Middle			High		
	Cohort 1	Cohort 2	Cohort 3	Cohort 1	Cohort 2	Cohort 3	Cohort 1	Cohort 2	Cohort 3
Planning Year	+	+	No Clear Effect	+	No Clear Effect	No Clear Effect	No Clear Effect	+	No Clear Effect
Implementation Years	-	No Clear Effect	+	+	No Clear Effect	No Clear Effect	+	No Clear Effect	+
Post-II/USP Years	-	-		-	No Clear Effect		No Clear Effect	No Clear Effect	

One of the key findings from the 2003 PSAA Evaluation was that II/USP schools experienced a planning year advantage over comparison schools in Cohorts 1 and 2. This pattern is, to a certain extent, still seen in this year's data, with four of the six Cohort 1 and 2 II/USP school-type groups (elementary, middle, and high schools in Cohorts 1 and 2) showing a statistically significant but relatively small positive effect for at least two of the outcome measures during the planning year. (The effect sizes range from 0.015 to 0.135 standard deviations.) However, none of the Cohort 3 groups of schools demonstrate a planning year advantage over comparison schools. (Although in Cohort 3 elementary and high schools gains were seen in the implementation years.)

One possible explanation for the different pattern seen in Cohort 3 schools is that there may be a selection bias. It is possible that the more responsive schools had already entered the program in Cohorts 1 and 2, and as a result the more recalcitrant schools were left in Cohort 3. Alternatively, Cohort 3 schools had already been operating in a high-stakes accountability environment for several years prior to II/USP participation. Thus a smaller increase in attention may have resulted from being identified as a low-performing school in comparison to the earlier cohorts. In other words, there may be less attention paid to becoming an II/USP school since the II/USP was no longer a "new" program. Nevertheless, the relative optimism associated with a possible planning year effect expressed in the first report has to be somewhat mitigated based on these Cohort 3 findings.

Summary of Findings

Rising performance is observed among the state's low performing schools during the period of II/USP implementation. Thus, some general effect may be occurring in relation to the state/federal accountability movement. There is still some indication of a planning year "bump" as reported in the 2003 PSAA Evaluation Report, but the evidence for this effect is weaker. Cohort 3 schools do not show this difference, and in Cohort 1 the effect is not significant for all measures. Modest II/USP effects are also found in the implementation years. The II/USP effects that are found in the data are statistically significant but very small. Most education research would not consider them educationally significant.[11] In addition, any gains that are made by II/USP schools in relation to comparison schools dissipate when II/USP comes to an end.

II/USP Effect by School Type and Cohort

The following section provides greater detail regarding the overall performance trends observed for II/USP and comparison schools measured by API, CST, SAT-9, and CAT/6, by level of schooling (elementary, middle, and high), and by cohort.[12]

For the API and CST scores, the overall and relative gains are fairly self explanatory;[13] positive gains mean that on average, II/USP schools increased their API and/or CST scores over time. For changes in SAT-9 and CAT/6 standardized scores, however, it is important to keep in mind that the methodology accounting for the substantial change in test does not allow the measurement of gains in performance over time. Standardized scores can only

[11] The literature considers medium or large effect sizes "educationally significant," indicating that meaningful numbers of students are affected (see Cohen, 1969, p. 278-279).

[12] Cohort 3 is the only group of schools that have CST scores for all years of II/USP program participation. A dummy variable is included in the regressions to control for participation in CSR.

[13] API scores are on a linear continuum between 200 and 1000, and CST scaled scores range from 150 to 600.

measure the relative performance of II/USP schools in comparison to a benchmark (e.g., the state average in the case of overall gains and comparison schools in the case of relative gains). For example, a positive overall gain means that II/USP schools are improving their performance in relation to the state average, but does not necessarily mean that II/USP schools are improving their absolute performance. One possible scenario is that the overall state performance is improving, so a *negative* performance relative to the state average means that II/USP schools are not improving their performance at the same rate as the state, but they may still be improving in absolute terms.

Relying on the API alone to measure growth over time is not recommended because each year the composition of the API changes, i.e. is constructed using a different weighting structure of test scores. Similarly, only looking at the SAT-9 and CAT/6 scores provides us with limited knowledge because we are only able to measure relative gains and not absolute gains over time. However, an analysis of API, CST, and standardized SAT-9 and CAT/6 scores together allows us to analyze a complete picture of performance trends over time, which is what we attempt to show in the following section.

Elementary Schools

When analyzing performance for elementary schools it is important to note the substantial upward trajectory for both II/USP and comparison schools in all outcome measures across all cohorts. In other words, II/USP and comparison schools appear to be improving their absolute performance over the time period. Cohort 1 II/USP and comparison elementary schools gained approximately 190 API points between 1998 and 2004, while Cohorts 2 and 3 saw gains of approximately 176 and 166 points respectively.

For Cohort 3, we can also look at the gains made on the CST mean scale scores. During the three years of II/USP, Cohort 3 II/USP and comparison elementary schools improved their math CST scores by 19 and 17 points respectively. These increases were much higher than the gains made on the ELA sections of the CST, which were approximately 7 points for both II/USP and comparison schools.

In terms of the SAT-9 and CAT/6, the upward trajectory means that both II/USP and comparison schools are moving closer to the state average. For example, during the 1999-2000 school year, Cohort 1 II/USP elementary schools' standardized scores in reading were 0.54 standard deviations below the state average, but by spring of 2004, this gap was reduced to an average of 0.42 standard deviations.

The analysis below will present a detailed examination of the relative performance of II/USP and comparison elementary schools across all cohorts.

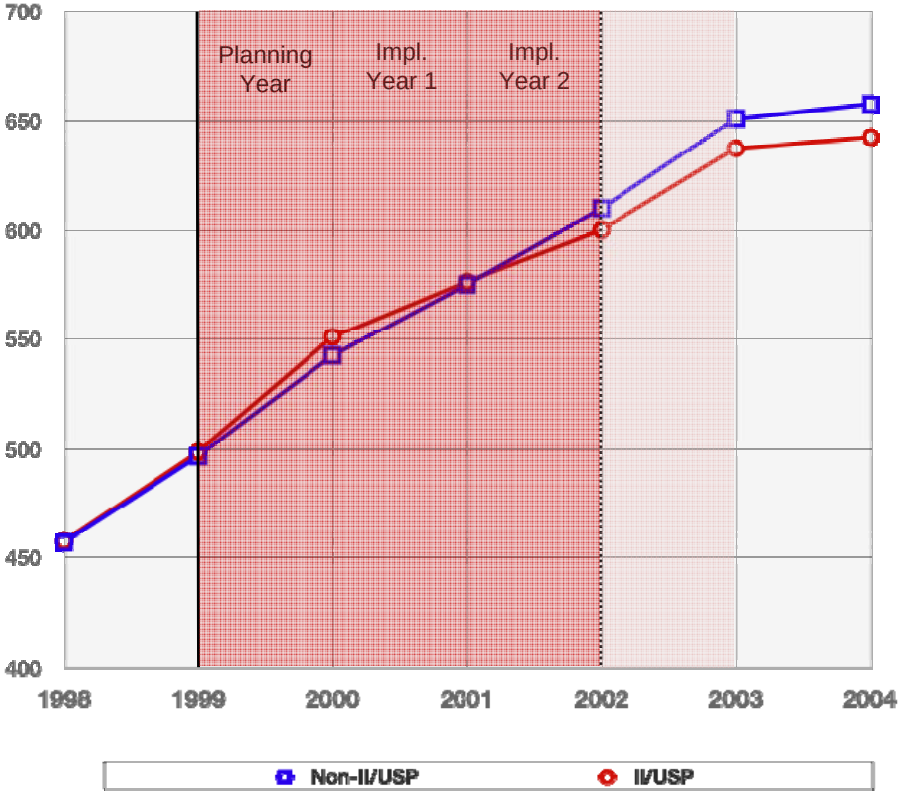
Relative Performance of II/USP and Comparison Elementary Schools

In this section we present the relative performance results for II/USP and comparison elementary schools. We have selected a sample of graphs to show in an illustrative way the patterns observed from the HLM regression results. The full set of results can be found in Appendix A-8.

Cohort 1 II/USP and comparison schools had similar API trends before II/USP began (see Exhibit 3.8 below). However, II/USP elementary schools gained an average of 7 API points more than comparison schools during the planning year (1999-2000), with a statistically

significant effect size of 0.11 standard deviations.[14] Since the average API target for this group of schools in 2000 was 14.4 points, a gain of 7 points is approximately half of that annual target. During the planning year, standardized scores show that II/USP schools had a significant relative improvement in relation to the comparison group in math, but not in reading. After the planning year, II/USP elementary schools were unable to maintain their advantage, and by spring of 2002, comparison schools had caught up to and surpassed their II/USP counterparts in terms of API, SAT-9 and CAT/6 scores.

Exhibit 3.8: Relative API Performance, Cohort 1 Elementary Schools



Cohort 2

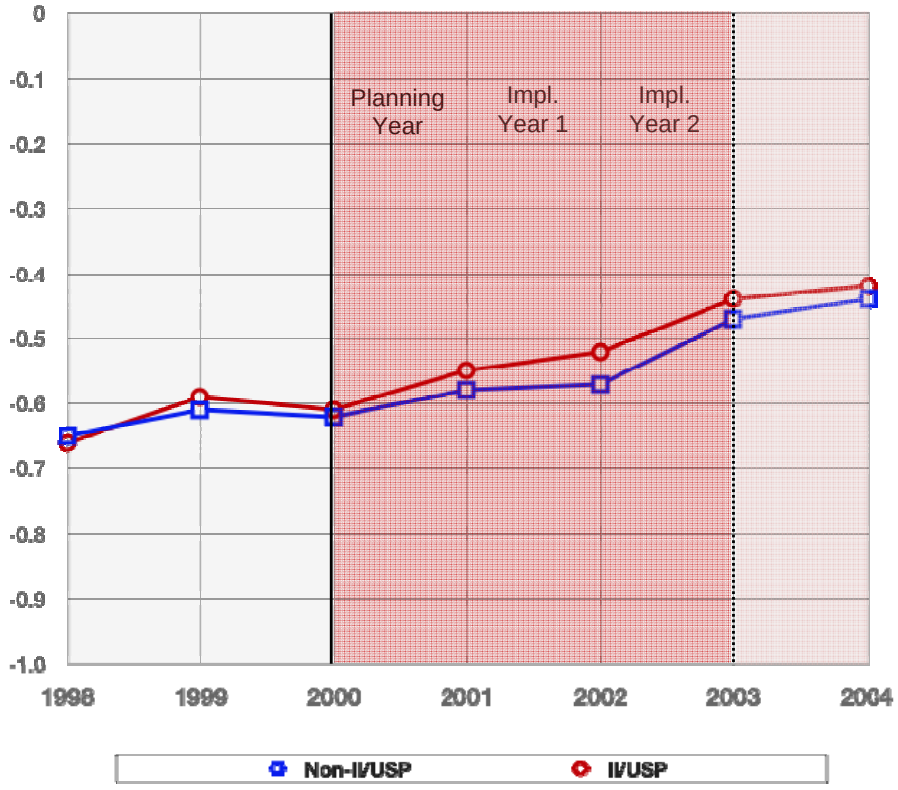
The results suggest that, as with Cohort 1, the II/USP elementary schools in Cohort 2 experienced a significant relative gain during the planning year. Unlike Cohort 1 schools, the gain is statistically significant across all three measures (7.47 API points, 0.015 standard deviations in math, and 0.032 standard deviations in reading). Cohort 2 elementary schools were able to maintain this advantage over comparison schools during the implementation years, though it narrowed after II/USP ended. One possible explanation for the narrowing of

[14] The 2003 PSAA Evaluation found that II/USP Cohort 1 elementary schools gained an average of 8.7 points more than comparison schools during the planning year, with an effect size of 0.11 standard deviations. The findings in this year's report are slightly different because adding additional years of data affects the estimated average gains for all years of the analysis. Furthermore, the treatment and comparison groups may be slightly different since we did not drop schools that were missing API scores for at least one year in the current analysis.

the gap in 2003-04 is that some comparison schools participated in HPSGP and/or Program Improvement, which may have helped improved their performance.[15] Exhibit 3.9 illustrates these patterns using reading test results.

There is also an interesting CSR effect for Cohort 2 elementary schools (not shown in the graph; see Appendix A-8 for the regression results). In the planning year, the reading and math standardized test scores show a significant negative CSR effect; however, this reverses in the first and second implementation year, where both indicators show a significant positive CSR effect. Since this cohort of CSR schools was self-selected from all II/USP schools during the planning year, there could be a systemic difference in their experiences during this year. The positive effect during the implementation years could suggest that the implementation of a whole school reform model positively influenced their growth in relation to other schools.

Exhibit 3.9: Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores),[16] Cohort 2 Elementary Schools



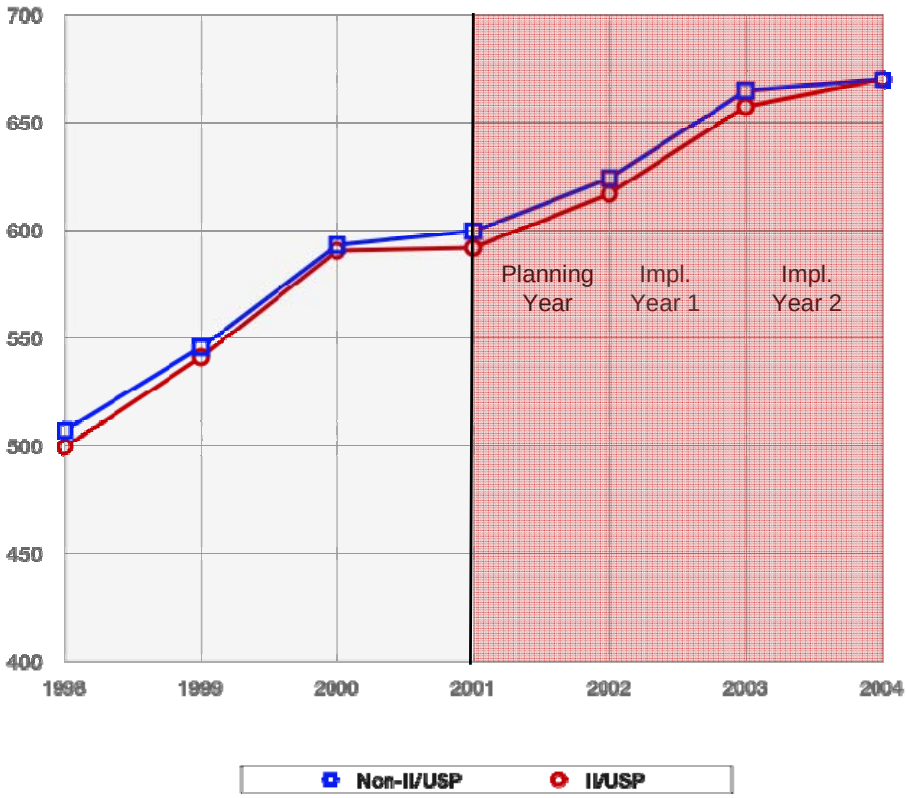
[15] 22 percent of comparison schools participated in HPSGP in 2003-04.

[16] Reading scores are constructed with standardized SAT-9 scale scores from 1998 through 2002, and CAT/6 standardized scale scores from 2002 to 2004.

Cohort 3

Exhibit 3.10 illustrates an example of Cohort 3 elementary school performance. Unlike the other cohorts, Cohort 3 II/USP elementary schools do not show statistically significant relative growth in their planning year, but rather later on during implementation years. The results for API scores indicate that Cohort 3 II/USP elementary schools started with slightly lower scores than comparison schools in the year prior to joining II/USP in 2001. After entering II/USP, Cohort 3 elementary schools began to parallel comparison school growth rates and then even catch up with comparison schools because of a significant increase in growth rates (7.12 API points) during the 2003-04 school year. Similarly, II/USP schools showed significant positive growth in CST scores relative to comparison schools in the last year of the program, in contrast to the planning year, where there was no significant difference. Standardized scores show significant growth (relative to comparison schools) in the first implementation year for math (0.010 standard deviations) as well as the second implementation year for reading (0.013 standard deviations). Examining the period of time from 2001 to 2004, it is interesting to see how Cohort 3 schools recovered from a period of stagnant performance and subsequently raised achievement scores by the end of II/USP.

Exhibit 3.10: Relative API Performance, Cohort 3 Elementary Schools



We also see a significant, positive CSR effect in the planning year for Cohort 3 elementary schools based on the math and reading SAT-9 and CAT/6 scores (see Appendix A-8 for the regression results). This positive effect continues into the first implementation year (shown by the math and ELA CST in addition to the math and reading SAT-9 and CAT/6), but then disappears by most measures in the second implementation year.

Some Cohort 3 II/USP schools also participated in HPSGP, and it is important to examine whether the additional funding and other aspects of this program affected academic achievement. Including an HPSGP variable in the analyses generally show a significant positive effect (see regression results in Appendix A-8). API scores indicate that the HPSGP program had a positive significant effect in the first implementation year of II/USP, whereas the reading and math standardized scores show a positive significant effect of HPSGP in both the planning and first implementation year.

Summary of Elementary School Trends

- There is still some indication of a planning year "bump" as reported in the 2003 PSAA Evaluation Report, but the evidence for this effect is weaker. Cohort 3 schools do not show this difference, and in Cohort 1 the effect is not significant for all measures.

- Overall, any gains that are made by II/USP schools in relation to comparison schools dissipate when II/USP comes to an end. One possible reason for this is that schools may have implemented certain strategies during II/USP that they were unable to sustain when funding ended. We explore this possibility in greater depth in Chapter 4. Another possible reason is that some comparison group schools entered other programs such as Program Improvement or HPSGP during these years. Participation in these programs may have bumped up the comparison group in relation to the II/USP schools, although the number of comparison schools that participated in these programs is small.

Middle Schools

Similar to what was observed for elementary schools, an upward trend in both II/USP and comparison schools is seen for all three cohorts using all three outcome measures. For example, Cohort 1 II/USP middle schools were 0.85 standard deviations below the state average in 1998-1999 (one year before the start of II/USP), and reduced this gap to 0.71 standard deviations in 2003-2004. The API trajectory has improved for both II/USP and comparison schools over time, with Cohort 1 II/USP middle schools gaining an average of approximately 60 API points over the three years of the program (51 points for comparison schools) and 107 points (108 for comparison schools) since 1999 (nearly 2 standard deviations of growth over six years). Similarly, CST scores increased approximately 6 points on the math section and 9 points on the ELA section for Cohort 3 middle schools from 2002 to 2004. It should be noted that these gains are smaller than those for elementary schools.

Relative Performance of II/USP and Comparison Middle Schools

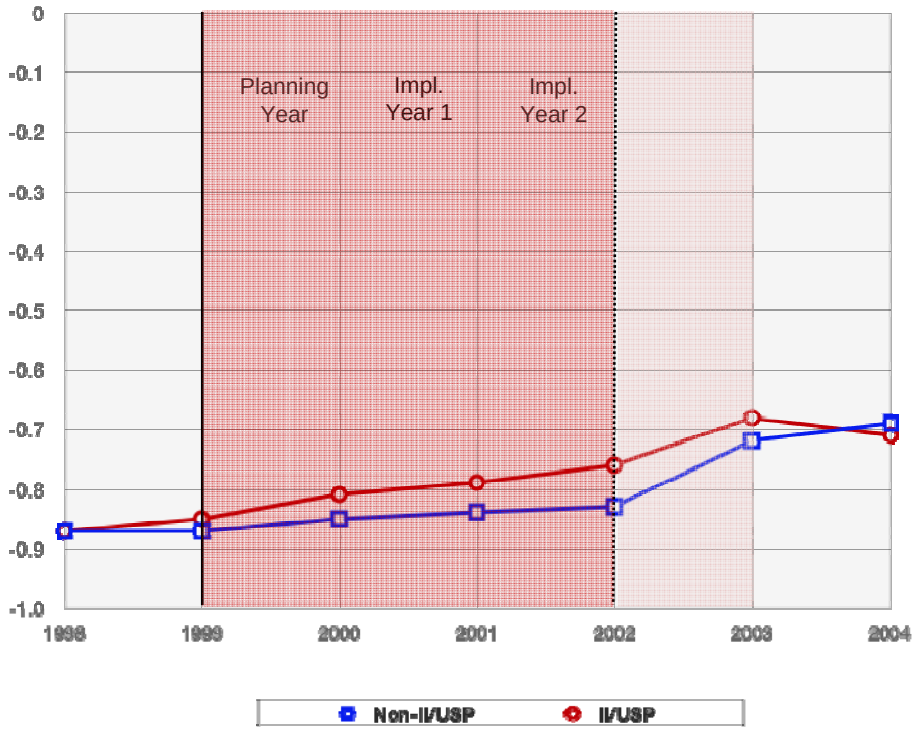
This section presents the relative performance results for II/USP and comparison middle schools. As in the previous section, a selection of graphical examples is shown to illustrate the growth trends.

Cohort 1

The results using API scores and the standardized student-level math scores indicate statistically significant but small growth for Cohort 1 II/USP middle schools in the planning year relative to comparison schools. During the 1999-2000 planning year, II/USP Cohort 1 middle schools increased an average of 9.72 API points more than comparison schools. In the same year, standardized math scores showed a relative increase of 0.036 standard deviations over comparison schools. During the last implementation year (see Exhibit 3.11 below), II/USP schools demonstrated further growth in the math and reading sections of the standardized scores (0.02 standard deviations more than comparison schools). However, we observed that any growth during II/USP was not sustained after the end of the program. In both years after exiting II/USP, treatment schools' standardized scores decreased relative to

comparison schools. The observed decrease of 0.02 standard deviations is statistically significant and is approximately equal to the statistically significant but relatively small gains achieved during each of the IIUSP implementation years.

Exhibit 3.11: Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores),[17] Cohort 1 Middle Schools



Cohort 2

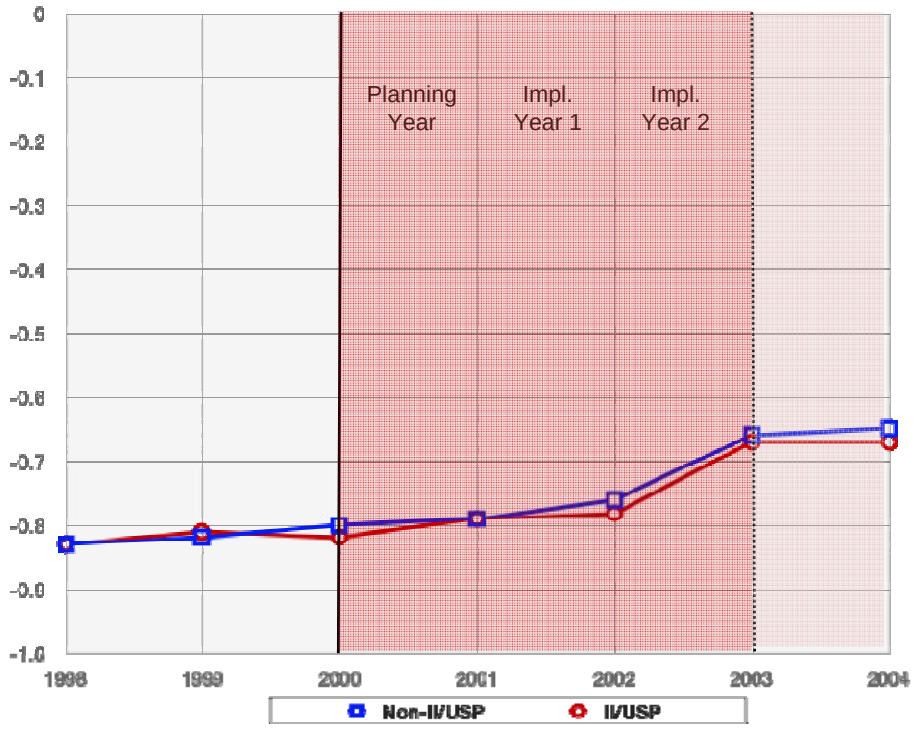
The API scores do not show any statistically significant trends for Cohort 2 II/USP middle schools, but the student-level standardized scores reveal that any gains made during II/USP were not sustained post-II/USP. When measured by student-level test scores, these schools experienced statistically significant but small negative gains in the first year of implementation, and a statistically significant but small positive gain in the second year of implementation. After II/USP support ended, these schools were not able to maintain this gain. In the following school year of 2003-2004, the gap between II/USP and the state average increased by 0.02 and 0.002 standard deviations more than in comparison schools (in reading and math, respectively). Exhibit 3.12 shows the growth trajectories described using reading test scores results.

There appears to be a positive CSR effect in the planning year for Cohort 2 middle schools (see regression results in Appendix A-8.) In the reading and math standardized tests, there are statistically significant but small, positive effects of 0.049 and 0.026 standard deviations

[17] Reading scores are constructed with standardized SAT-9 scale scores from 1998 through 2002, and CAT/6 standardized scale scores from 2002 to 2004.

respectively. In both of the implementation years, however, the CSR effect either becomes negative or not significant.

Exhibit 3.12: Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores),[18] Cohort 2 Middle Schools



Cohort 3

Given that none of the measures shows any significant difference between Cohort 3 II/USP middle schools and comparison schools, no graphs are shown for these schools. However, for II/USP schools that also participate in HPSGP, all of the outcome measures show a significant positive effect (see regression results in Appendix A-8). The reading standardized test scores reveal a significant positive effect in all three years, while the API shows a significant effect in only the planning year and the math scores show a significant effect in the planning year and the second implementation year. Thus, significant differences are observed for middle schools in Cohort 3 participating in both the II/USP and the HPSGP.

Summary of Middle School Trends

- There was a lower overall gain in achievement on all measures among middle schools in comparison to elementary schools. This trend was true for both the II/USP and comparison schools and is in line with historical trends.

[18] Reading scores are constructed with standardized SAT-9 scale scores from 1998 through 2002, and CAT/6 standardized scale scores from 2002 to 2004.

- The indication of a planning year "bump" is weaker for middle schools. It was only evident with Cohort 1 schools.

- As with elementary schools, gains made by II/USP schools relative to comparison schools dissipate after II/USP participation.

- In Cohort 3 there were no significant differences between II/USP and comparison schools, although there was evidence of a positive HPSGP effect for II/USP schools participating in both programs.

High Schools

As was seen for elementary and middle schools, II/USP and comparison high schools also show an upward trend over the years. For example, Cohort 1 II/USP high schools increased an average of approximately 79 API points from 1998-1999 (the year before the start of II/USP) to 2003-2004. However, the upward trend for high schools is noticeably flatter than what was observed for elementary and middle schools.

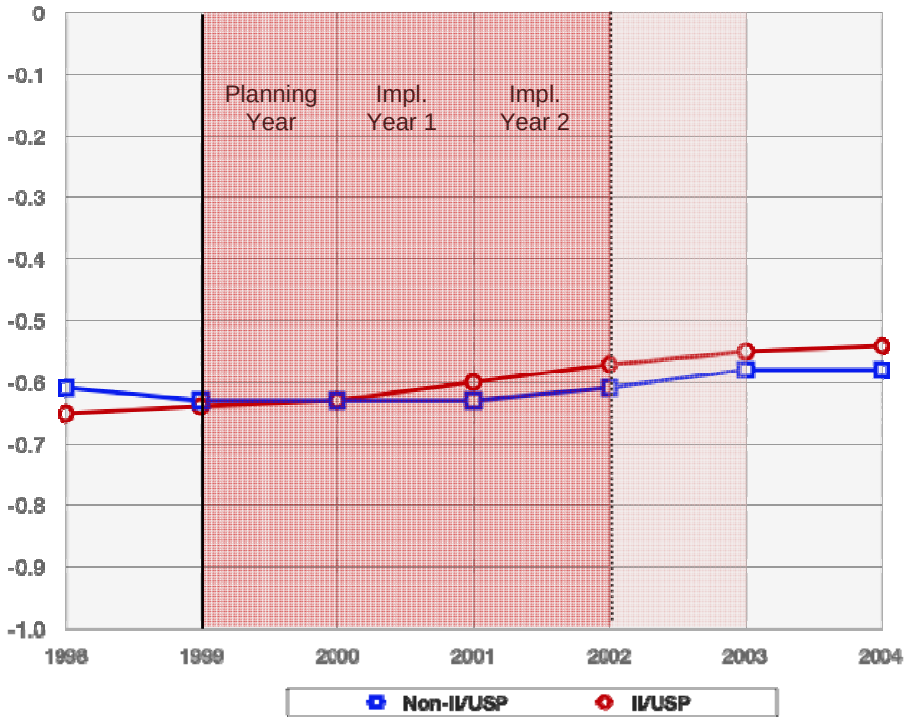
Relative Performance of II/USP and Comparison High Schools

Relative performance results for II/USP and comparison high schools are shown in this section for the three cohorts. As before, a selection of graphs is presented to illustrate the growth trends of these schools.

Cohort 1

The analysis of student-level standardized scores reveals some statistically significant but relatively small differences between II/USP and comparison schools that are not apparent from API trajectories alone. Although Cohort 1 II/USP high schools did not demonstrate the planning year advantage that was observed for Cohort 1 elementary and middle schools, during the two implementation years, they experienced significant gains relative to comparison schools (approximately 0.02 standard deviations in reading and math standardized scores). This advantage was not sustained in the first year after II/USP ended, with Cohort 1 II/USP high schools showing a small, but statistically significant decline in their growth rate than comparison schools in reading and math scores. Exhibit 3.13 shows these growth trajectories for math test score results.

Exhibit 3.13: Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores),[19] Cohort 1 High Schools



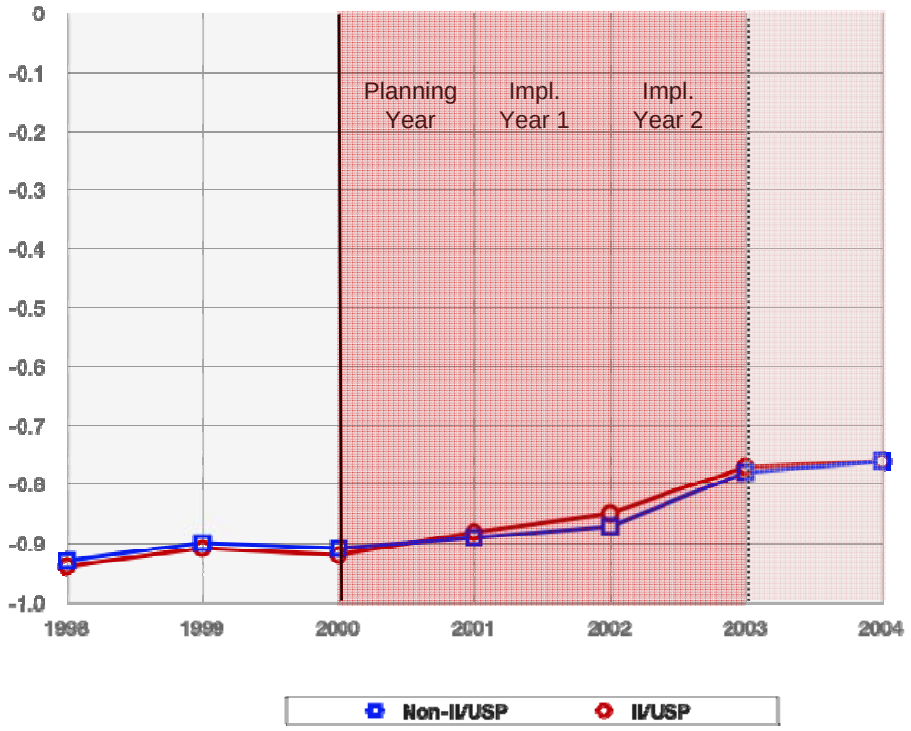
Although Cohort 1 II/USP high schools do not show a planning year effect, we do see a significant but small positive CSR effect of approximately 0.1 standard deviations in both the math and reading SAT-9 and CAT/6. This positive effect does not last throughout the implementation years; the effect either becomes negative or is no longer significant.

Cohort 2

II/USP Cohort 2 high schools exhibit statistically significant but relatively small growth in the planning year of II/USP based on all three outcome measures (see Exhibit 3.14 below). In the planning year, these schools grew an average of about 8 API points more than comparison schools, and also demonstrated significant relative gains on the standardized reading and math scores (0.03 and 0.02 standard deviations, respectively). However, any significant gains made during II/USP were not sustained, and a significant drop-off effect can be seen in the year after the program ended.

[19] Math scores are constructed with standardized SAT-9 scale scores from 1998 through 2002, and CAT/6 standardized scale scores from 2002 to 2004.

Exhibit 3.14: Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores),[20] Cohort 2 High Schools

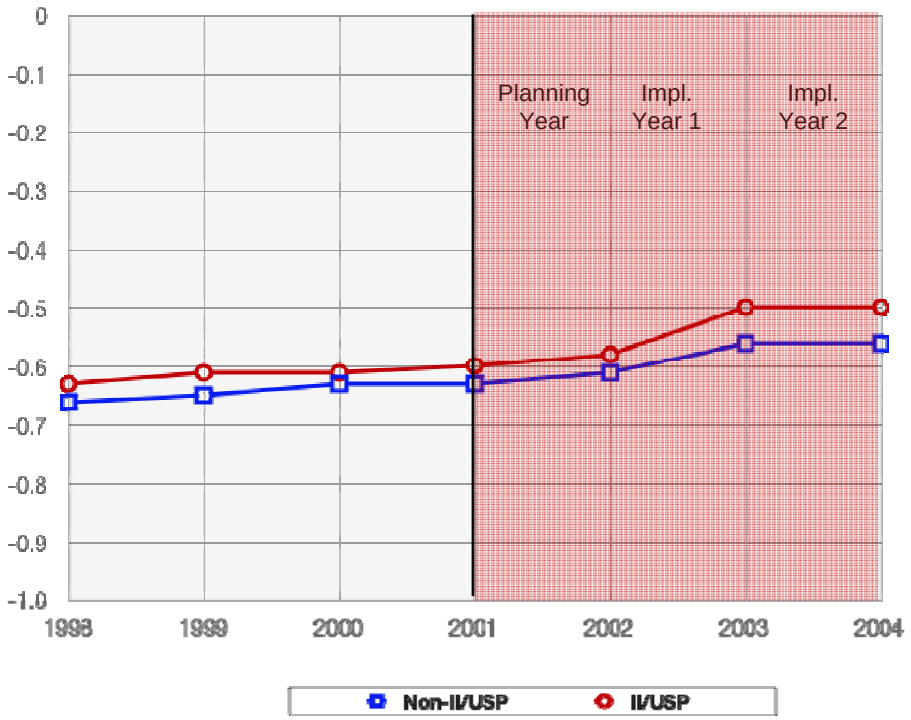


Cohort 3

There were no consistent achievement trends observed for Cohort 3 II/USP high schools. There was a positive statistically significant II/USP effect in the first implementation year, but not in the second year of implementation (see Exhibit 3.15 below). There is also evidence of a significant positive HPSGP effect in the second implementation year (see regression results in Appendix A-8.)

[20] Reading scores are constructed with standardized SAT-9 scale scores from 1998 through 2002, and CAT/6 standardized scale scores from 2002 to 2004.

Exhibit 3.15: Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores),[21] Cohort 3 High Schools



Summary of High School Trends

- The upward trend among II/USP and comparison high schools is noticeably flatter than what was observed for elementary and middle schools.

- Some positive gains were found for II/USP schools in relation to comparison schools at various times during II/USP implementation (the planning year for Cohort 2, and implementation years for Cohort 1).

- Any positive gains during II/USP were lost in the years after II/USP implementation.

[21] Math scores are constructed with standardized SAT-9 scale scores from 1998 through 2002, and CAT/6 standardized scale scores from 2002 to 2004.

Analysis of the Impact of Program Participation, by Subpopulation

We will now shift the discussion from analyzing the aggregate effects of II/USP on student achievement to considering the effect of II/USP on certain subpopulations of students within II/USP schools. As specified in Research Question 1, the three groups of students that will be targeted in this section are:

- English language learners (ELs)
- Pupils with exceptional needs (in special education)
- Pupils that qualify for free or reduced price meals and are enrolled in schools that receive funds under Title I, Section A of the Improving America's Schools Act (IASA) of 1994.

We look at the performance of these subpopulations from two perspectives. First, we will look at whether there are changes over time in the achievement gap between each of these subpopulations of students as compared to all other students in II/USP and non-II/USP schools. Then we will determine whether changes in the achievement gaps differ between schools receiving II/USP support and the comparison schools. These two perspectives will allow us to better understand II/USP participation in relation to the achievement gap between each of these subpopulations of students as opposed to all other students.

The following sections highlight the overall achievement trends for each subgroup in II/USP and non-II/USP schools as measured by the performance of students on the SAT-9 and CAT/6 over time. The results of the analysis are presented in two ways. First, the average relative gains over time for the subpopulation as opposed to other students in both II/USP and comparison schools are presented in a graph. Similar to the last section, it is important to keep in mind that these graphs display the changes in the gap between the different subpopulations and the state average (i.e., including all students) in each year before, during, and after II/USP. For brevity, we only include graphs for Cohort 1 elementary schools, which have the longest trajectory that includes two years of post-II/USP data. In addition to the graphs, we present a summary table of changes in the achievement gap between the subpopulations and all other students (i.e., all students not in the respective subgroup) by school type and cohort. These data are derived from the parameter estimates and significance levels, which can be found in Appendix A-9.

English Learners

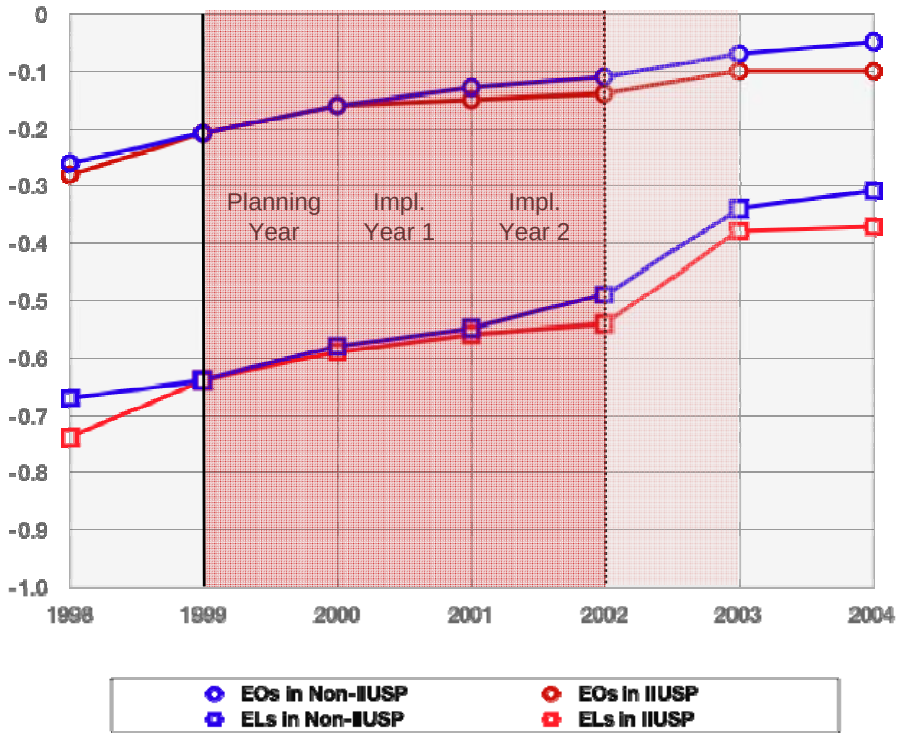
This first section analyzes the academic achievement of English learners (ELs) relative to non-ELs (English-only students, or EOs) in II/USP and comparison schools. Note that, in order to avoid confounding results, this analysis does not include former English learners who have since been reclassified as fluent.

Exhibit 3.16 presents the achievement trends on the standardized SAT-9 and CAT/6 reading scores of ELs and non-ELs in Cohort 1 elementary II/USP and comparison schools. The upward trend from 1997-98 to 2003-04 indicates that all groups of students in II/USP and comparison schools are getting closer to the statewide average. It also appears that the gap between ELs and non-ELs has narrowed slightly over the years as a result of ELs improving their performance faster than non-ELs. In reviewing these changes in achievement gaps between EOs and ELs, it is important to keep in mind the substantial increase in EL participation in the large scale assessments over time. The number of ELs tested in statewide

assessments has increased from about 72 percent in 1997-98 to 97 percent in 2003-04. In contrast, EO participation rates have been virtually constant over time, with a participation rate of 96 percent in 1997-98 and 95 percent in 2003-04 (Parrish et al., 2005).

Another noticeable trend is the sharp reduction in the gap that occurred in 2002-03 for ELs. This change corresponds with the year that California changed its statewide assessment from the SAT-9 to the CAT/6. The CAT/6 differs from the SAT-9 in terms of format, content emphasis, difficulty level, and number of items. It is possible that these differences benefited ELs more than EOs, which would explain the sharp improvement in their performance in the year that the test changed. Parrish et al. (2005) analyzed the performance of ELs and EOs using the CST test results to see if the same sharp reduction in the gap was observable when using an alternative assessment measure. They found that CST performance shows a small gap decrease from 2002 to 2003 (from 0.90 standard deviations in 2002 to 0.84 in 2003). Given this, they argue that it seems likely that at least some of the gap decrease observed in conjunction with the introduction of the CAT/6 is an artifact of the change in test.

Exhibit 3.16: Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores)[22] for ELs and EOs in Cohort 1 II/USP and Comparison Elementary Schools



[22] Reading scores are constructed with standardized SAT-9 scale scores from 1998 through 2002, and CAT/6 standardized scale scores from 2002 to 2004.

Since it is difficult to visually ascertain whether the difference between II/USP and comparison schools is statistically significant, the following Exhibit 3.17 summarizes the achievement gains measured by the reading standardized scores made by ELs. [23] This exhibit also illustrates whether the gap between ELs and EOs in II/USP and comparison schools has increased or decreased over the course of the program participation for each cohort. A positive sign indicates a *decrease* in the gap. It also shows if the decrease or increase in the gap between ELs and EOs in II/USP schools was statistically significantly higher or lower than in comparison schools. A positive sign indicates that II/USP schools closed the gap at a *higher* rate than comparison schools. For effect sizes see Appendix A-7.

ELs have made statistically significant gains since 1999 compared to EOs in II/USP and comparison schools across almost all cohorts and school types. For example, the gap has closed across all cohorts and school types for II/USP schools. However, the effect size of the gap closure in elementary schools is considered very small (lower than 0.1 standard deviations), and only somewhat larger for middle and high schools with effect sizes ranging from 0.1 to 0.15 standard deviations.[24]

Evidence of an II/USP effect varies according to school type. No II/USP effect is observed in elementary schools; with an inconsistent pattern regarding II/USP participation seen in middle schools. For high schools, however, a significant positive II/USP effect is observed for ELs. This seems worthy of note given the consistency of this pattern across all three cohorts.

[23] The math scores for ELs in elementary schools follow a pattern similar to that of the reading scores; however, middle and high school results differ depending on whether one uses math or reading scores. We have chosen to focus on the reading scores since they are a more relevant estimate of the academic progress of ELs.

[24] The effect size is measured in standard deviations: .25 standard deviations or lower indicates a small effect size, between .25 and .40 is a medium effect size, and .40 standard deviations or above is a large effect size. The literature considers medium or large effect sizes "educationally significant," indicating that meaningful numbers of students are affected (see Cohen, 1969, p. 278-279).

Exhibit 3.17: Change in Achievement Gap between ELs and EOs, by Cohort

	Elementary			Middle			High		
	Cohort 1	Cohort 2	Cohort 3	Cohort 1	Cohort 2	Cohort 3	Cohort 1	Cohort 2	Cohort 3
Change in Gap in II/USP Schools[25]	+	+	+	+	+	+	+	+	+
Change in Gap in Comparison Schools[22]	+	+	+	No Clear Effect	+	+	No Clear Effect	+	+
II/USP Effect on Gap[26]	No Clear Effect	No Clear Effect	No Clear Effect	+	No Clear Effect	-	+	+	+

[25] This row indicates whether the gap between EL students and EO students has increased or decreased over the course of program participation for each cohort. A positive sign indicates a statistically significant reduction in the gap and a negative sign indicates a statistically significant increase in the gap. No clear effect means that there is no statistically significant change in the gap.

[26] This row indicates if the decrease in the gap between ELs and EOs in II/USP schools was statistically significantly higher or lower than in comparison schools. A positive sign indicates that II/USP closed the gap at a higher rate than comparison schools.

Special Education Students

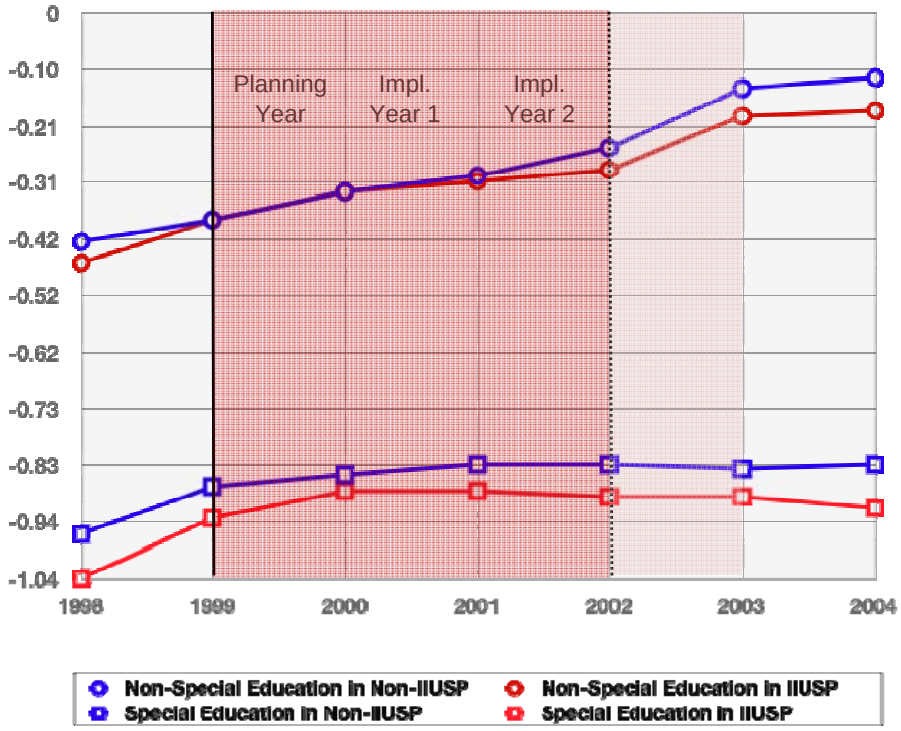
We define students with special needs as those receiving special education services. This group of students makes up approximately 9.59 percent of the students in California. This makes them a far smaller subpopulation than ELs.[27] The results for these students are far less positive than those for English learners presented above. There are many confounding factors affecting the population of special education students that hinder the measurement of an II/USP effect. One of these factors is the huge increase in participation in large scale assessments that has affected special education students.

The No Child Left Behind Act of 2001 (NCLB) mandates that 95 percent of special education students participate in a state's annual assessment of achievement in order to increase school accountability. Since 1997-98, the participation rate of special education students has more than doubled, from 41.6 percent in 1997-98 to 88 percent in 2003-04. As it is likely that students with the most severe disabilities were those previously excluded from testing, we would expect the academic performance of the former excluded special education students to be lower than for previously non-excluded special education students (Harr, Pérez, McLaughlin & Blankenship, 2005).[28] In light of this very large increase in participation, it is not surprising to also observe a reduction in the performance of special education students as a subgroup and therefore a substantial increase in the gap between special education and all other students as shown in Exhibit 3.18 for the period 2001-02 to 2003-04.

[27] The counts of special education students ages 6 through 21 are from the Annual Reports to Congress on the Implementation of the Individuals with Disabilities Education Act (IDEA). Total public enrollment is from the NCES Common Core of Data, 2001-02.

[28] The authors estimated the achievement performance of the former excluded special education students using NAEP scores. They demonstrated that the estimated achievement of the excluded special education students is lower than the average achievement of non-excluded special education students.

Exhibit 3.18: Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores)[29] for Special Education and Non-Special Education Students in Cohort 1 II/USP and Comparison Elementary Schools



The summary of achievement trends shown in Exhibit 3.19 below further substantiates these results. For brevity, we only display the achievement trends for the reading standardized test scores; the math results are very similar to the reading results and can be found in Appendix A-9. The full set of effect sizes can be found in Appendix A-7. In both comparison and II/USP schools, the general tendency is an increase in the gap between special education and non-special education students (i.e., therefore the negative sign in the exhibit).

The overall II/USP effect on the achievement trends of special education students is varied. Elementary and middle schools follow a similar pattern: a positive significant effect for Cohort 1, no effect for Cohort 2, and a significant negative effect for Cohort 3. There is no II/USP effect for Cohort 1 and 3 high schools, and a negative effect for Cohort 2 high schools. It is important to point out the positive II/USP effect that is observable for Cohort 1 elementary schools. This is due to the fact that the gap increase in comparison schools is larger than the one in II/USP schools. These results are sufficiently mixed to result in no clear evidence of II/USP being beneficial to the overall academic trajectory of special education students.

[29] Reading scores are constructed with standardized SAT-9 scale scores from 1998 through 2002, and CAT/6 standardized scale scores from 2002 to 2004.

Exhibit 3.19: Change in Achievement Gap between Special Education Students and Students Receiving Regular Education, by Cohort

	Elementary			Middle			High		
	Cohort 1	Cohort 2	Cohort 3	Cohort 1	Cohort 2	Cohort 3	Cohort 1	Cohort 2	Cohort 3
Change in Gap in II/USP Schools[30]	-	-	-	No Clear Effect	No Clear Effect	-	No Clear Effect	-	+
Change in Gap in Comparison Schools[22]	-	-	-	No Clear Effect	-	-	No Clear Effect	No Clear Effect	No Clear Effect
II/USP Effect on Gap[31]	+	No Clear Effect	-	+	No Clear Effect	-	No Clear Effect	-	No Clear Effect

[30] This row indicates whether the gap between special education students and students receiving regular education has increased or decreased over the course of program participation for each cohort. A positive sign indicates a statistically significant reduction in the gap and a negative sign indicates a statistically significant increase in the gap. No clear effect means that there is no statistically significant change in the gap.

[31] This row indicates if the change in the gap between special education students and students receiving regular education in II/USP schools was statistically significantly higher or lower than in comparison schools. A positive sign indicates that the II/USP schools' gap increased at a lower rate than comparison schools.

Students Eligible for Free and Reduced Price Lunch

Students receiving free or reduced price lunch make up about 50 percent of all students in California, and comprise nearly 80 percent of students in the II/USP and comparison schools. In this section we analyze the academic performance of students receiving assistance from the Free or Reduced Price Lunch Program (FRLP) participating in the II/USP, as compared to the academic achievement of II/USP participants not receiving free or reduced price lunch. The achievement trends in the standardized SAT-9 and CAT/6 reading scores for students in Cohort 1 elementary schools are displayed in Exhibit 3.20. The gap between all of the students in these schools in relation to the state is decreasing. However, while the largest increase relative to state average performance for students not receiving free or reduced price lunch occurred in 1998 and 1999, the largest increase for students receiving free or reduced price lunch occurred in 2002. This is after the II/USP program ended for most schools in this cohort. As a result, the gap between the two groups is largest in the II/USP years.

Exhibit 3.20: Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores)[32] for Students Eligible for Free or Reduced Price Lunch and "Regular" Students in Cohort 1 II/USP and Comparison Schools Elementary

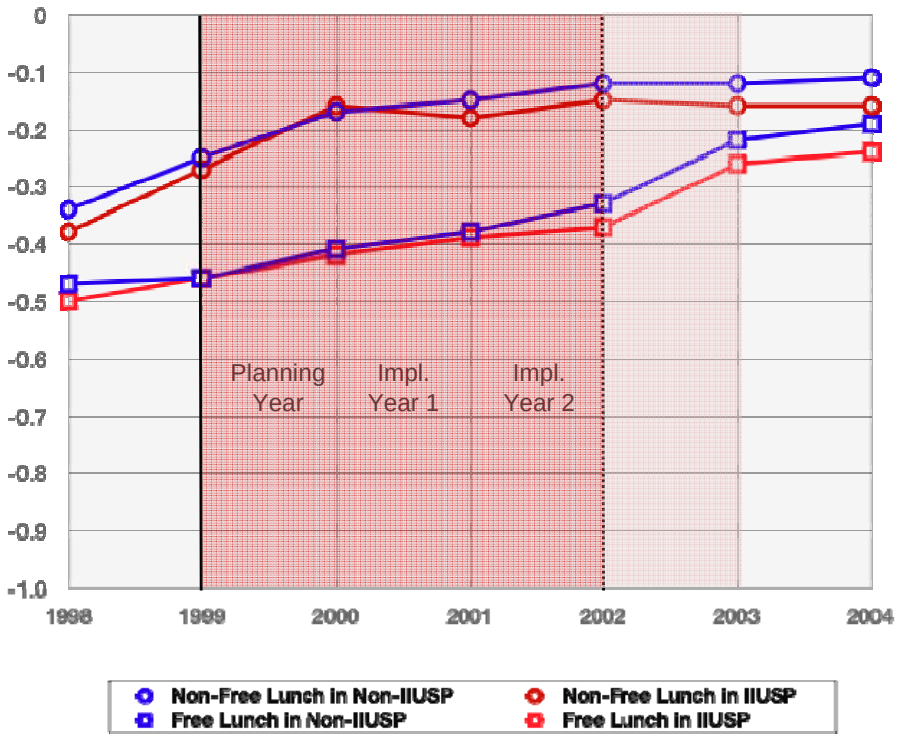


Exhibit 3.21 summarizes the achievement trends for these two categories of students, and estimates an II/USP effect by cohort and school type. Once again, we see that the gap between FRLP students appears to increase from 1998 to 2000, but then decreases from 2000

[32] Reading scores are constructed with standardized SAT-9 scale scores from 1998 through 2002, and CAT/6 standardized scale scores from 2002 to 2004.

until 2004. The II/USP effect is varied across cohort and school type. A positive II/USP effect is observed for Cohort 2 elementary and middle schools, as well as Cohort 1 and 3 high schools. However, a negative II/USP effect is seen in Cohort 1 elementary schools.

These results indicate possible benefit from being in an II/USP school for students eligible for free and reduced price lunch in that the gap between their performance and non-FRLP students is decreasing at a somewhat faster rate than observed in the comparison sites. However, the implications of this finding are hard to interpret in that we see little enhancement in overall school performance as a result of II/USP participation and given that 80 percent of the students in these treatment as well as comparison schools receive FRLP.

Exhibit 3.21: Change in Achievement Gap between Students Eligible for Free or Reduced Lunch Price (FRLP) and non-FRLP, by Cohort

	Elementary			Middle			High		
	Cohort 1	Cohort 2	Cohort 3	Cohort 1	Cohort 2	Cohort 3	Cohort 1	Cohort 2	Cohort 3
Change in Gap in II/USP Schools[33]	-	+	+	No Clear Effect	+	+	No Clear Effect	+	+
Change in Gap in Comparison Schools[22]	No Clear Effect	+	+	No Clear Effect	+	+	-	+	+
II/USP Effect on Gap[34]	-	+	No Clear Effect	No Clear Effect	+	No Clear Effect	+	No Clear Effect	+

[33] This row indicates whether the gap between FRLP students and non-FRLP students has increased or decreased over the course of program participation for each cohort. A positive sign indicates a statistically significant reduction in the gap and a negative sign indicates a statistically significant increase in the gap. No clear effect means that there is no statistically significant change in the gap.

[34] This row indicates if the decrease in the gap between FRLP students and non-FRLP students in II/USP schools was statistically significantly higher or lower than in comparison schools. A positive sign indicates that II/USP closed the gap at a higher rate than comparison schools.

Analysis of the CDE Growth Target Requirements

The previous sections analyzed academic achievement in II/USP and non-II/USP schools using several continuous measures of performance: school-level API, and student-level SAT-9, CAT/6, and CST scores. This section analyzes the question of academic improvement from a different perspective. Here we analyze the number of years it takes low-performing schools to achieve the growth target considered significant/satisfactory by the CDE. This growth rate is measured by the time it took to meet the schoolwide and comparable improvement growth targets. In other words, for those schools that during their planning year did not meet their schoolwide and/or comparable improvement growth target, how many years later did they satisfy these conditions? How many of these schools did not meet these targets in any year of our analysis? Is it possible to observe important differences between II/USP and non-II/USP schools?

Exhibit 3.22 presents the number of II/USP schools that did not meet their schoolwide and/or comparable improvement growth targets. [35] The table shows that 67 II/USP elementary schools did not meet their growth targets during their planning year. Of these, 38 (57 percent) reached their targets in their first implementation year. Of the remaining schools, 10 (15 percent) and 12 (18 percent) satisfied their growth requirements for the first time two or three years after the planning year. This means that seven II/USP schools were not able to grow at the required rate in any year of this analysis. This represents 10.5 percent of the Cohort 1 elementary schools. These percentages are 5.6 for Cohort 2 and 11.5 percent for Cohort 3.[36]

Middle schools present a similar picture to elementary schools. But as we move to high schools, the percentage of II/USP and comparison schools meeting their required schoolwide and comparable improvement growth targets in any year after the planning year drops significantly. For instance, for Cohort 1 II/USP high schools, 37 percent did not meet their required growth rates, while for Cohort 2 and 3 this percent is equal to 28 and 24 percent, respectively. These results tend to suggest that improving academic achievement in low-performing high schools is harder than in elementary and middle schools.

[35] Appendix A-10 separates these schools in two groups. The first group is made up of schools that even though did not meet their growth targets, experienced a positive growth of at least one point in their API (labeled in yellow). The second group is made up of schools that grew less than one API point, or experienced a negative growth rate (labeled in red).

[36] It is important to keep in mind that we are able to observe only two years of achievement results for policy implementation in Cohort 3. This may explain the higher percentage of schools that never reached the required growth rate after their implementation year.

Exhibit 3.22: Number and Percentage of II/USP Schools that Did Not Meet their Schoolwide and/or Comparable Improvement Growth Targets

School Type	Cohort	II/USP or Comparison	Schools that Did Not Meet their Growth Targets During the Planning Year **"A"**	Schools that Never Met their Growth Targets During the Period of Analysis	Percent of Schools of "A" that Never Met their Growth Targets During the Period of Analysis
Elementary	Cohort 1	II/USP	67	7	10.5%
		non-II/USP	159	5	3.1%
	Cohort 2	II/USP	71	4	5.6%
		non-II/USP	86	5	5.8%
	Cohort 3	II/USP	113	13	11.5%
		non-II/USP	205	28	13.7%
Middle	Cohort 1	II/USP	32	3	9.4%
		non-II/USP	86	14	16.3%
	Cohort 2	II/USP	53	8	15.1%
		non-II/USP	65	5	7.7%
	Cohort 3	II/USP	38	4	10.5%
		non-II/USP	120	14	11.7%
High	Cohort 1	II/USP	30	11	36.7%
		non-II/USP	90	34	37.8%
	Cohort 2	II/USP	50	14	28.0%
		non-II/USP	116	23	19.8%
	Cohort 3	II/USP	55	13	23.6%
		non-II/USP	108	27	25.0%

Another important aspect of this analysis is that not all schools that satisfied their growth requirements during their planning year met their targets later on. In other words, there is not only school mobility from below- to above-target status (i.e., from not satisfying to satisfying growth targets), but also from above- to below-target status. Exhibit 3.23 shows that out of the 216 Cohort 1 II/USP elementary schools that met their targets during their planning year, only 41.2 percent (or 89 schools) also met them one year later (i.e., during their first implementation year). For Cohort 2 and 3 II/USP elementary schools this percentage is relatively similar, equal to 51.4 and 49.7 percent, respectively. But as we move to middle and high schools, this success rate drops significantly. The percent of schools that still satisfied their growth target one year after their planning year is, on average, equal to 33.7 and 20.4 percent for middle and high schools, respectively.[37] This empirical finding highlights that in middle, and especially in high schools, it is not only difficult to achieve the required growth targets, but when met, it is hard to sustain those levels of growth.

[37] These percentages are weighted averages of the percentages that appear in exhibit 3.23.

Exhibit 3.23: Number and Percentage of II/USP Schools that Met their Schoolwide and Comparable Improvement Growth Targets During the Planning and First Implementation Year

School Type	Cohort	Schools that Met their Growth Targets During the Planning Year **"A"**	Schools that Also Met their Growth Targets One Year After the Planning Year	Percent of Schools of "A" that Also Met their Growth Targets One Year After the Planning Year
Elementary	Cohort 1	216	89	41.2%
	Cohort 2	177	91	51.4%
	Cohort 3	165	82	49.7%
Middle	Cohort 1	43	18	41.9%
	Cohort 2	47	13	27.7%
	Cohort 3	11	3	27.3%
High	Cohort 1	21	4	19.1%
	Cohort 2	16	2	12.5%
	Cohort 3	17	5	27.4%

When we apply this type of analysis in comparing II/USP and non-II/USP schools, no clear and consistent differences are observable. A higher percent of II/USP than comparison schools that were not able to meet their growth targets during their planning year, were able to meet them later on in the case of elementary school Cohorts 2 and 3, middle school Cohorts 1 and 3, and high school Cohorts 1 and 3. In the other cases comparison schools were more successful. A similarly mixed set of results is observable for II/USP and non-II/USP schools that were able to meet their growth targets during their planning year, but not later on.

District Influences on School Performance

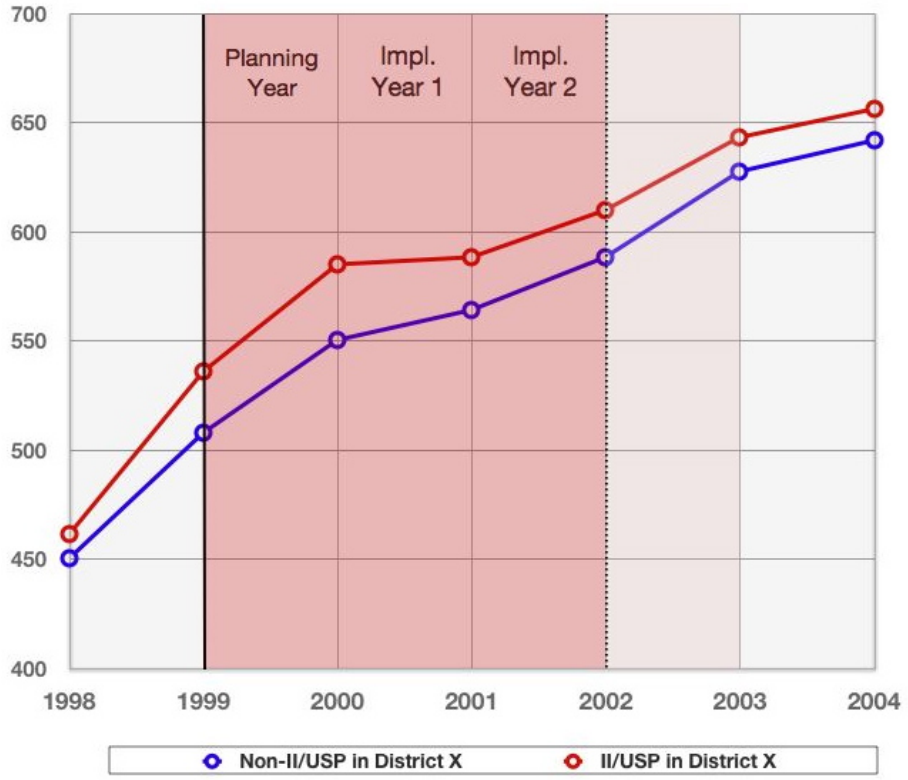
School districts are also an important potential vehicle for school reform, acting as an intermediary between state-level policy design and school-level policy implementation. The implementation of II/USP in low-performing schools in California is clearly not the only factor affecting observed differences in test score trajectories of II/USP and non-II/USP schools. II/USP schools may be more concentrated in school districts that are more successful in increasing student performance in low-performing schools. Or, the lack of substantial changes in improvement between II/USP and comparison schools may be more affected by factors common to both than the program per se. One important factor potentially affecting school performance is the stance of the district, especially in regard to its lowest performing schools. Because of this, it is important to better understand the role of the district in school reform and to attempt to disentangle test score trajectories specific to II/USP schools from general district trends.

The present section analyzes two important questions associated with school districts. The first is whether II/USP and non-II/USP schools show different test score trajectories within school districts. This type of analysis assists in attempting to assess a possible II/USP effect by holding a broad range of district factors constant. The second question for this section is whether there are test score trajectory differences across districts in regard to their II/USP and non-II/USP schools. This second type of analysis focuses on the districts' role in improving student achievement in low-performing schools in general. In response to the first question, we found no significant differences between II/USP and non-II/USP schools within districts. At the same time, some districts appear to be doing a better job than others of increasing student performance across their low performing schools.

Within-District Differences

In order to explore the first question—whether it is possible to observe significant differences in test score trajectories of II/USP and non-II/USP schools within districts—we expanded our HLM models to include variables designating selected individual districts. For this analysis, we chose four districts with relatively large numbers of II/USP and comparison schools. The full results of these analyses appear in Appendix A-11. In line with our 2003 PSAA Evaluation Report, we found that II/USP schools tend to perform fairly similarly to non-II/USP schools within individual districts. As an example, Exhibit 3.24 presents the test score trajectories of II/USP and non-II/USP schools in a given district. No significant differences are observable.

Exhibit 3.24: Average API Scores for Cohort 1 II/USP and Comparison Elementary Schools in District X



Across-District Differences

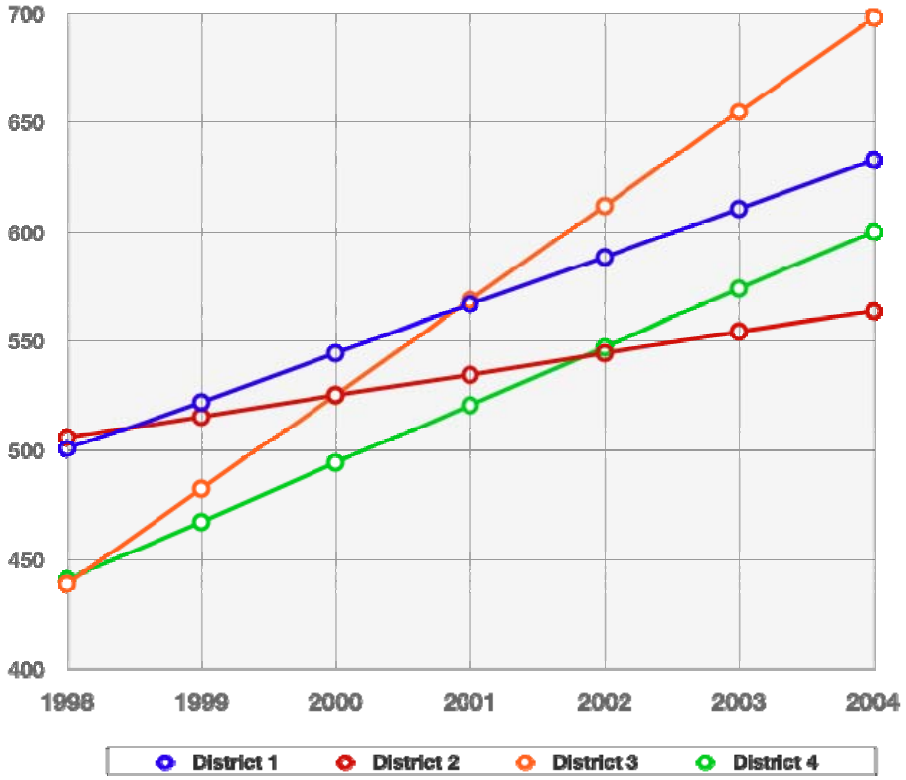
We now examine whether academic achievement improves for low-performing schools in general in some districts more than in others. Appendix A-11 presents the results of an HLM regression in which we have grouped II/USP and comparison elementary schools together in the nine largest school districts in terms of student enrollment in California. As can be observed, we have specified this model using a continuous time variable instead of dichotomous year indicators, in order to obtain an estimation of the average yearly API increase of low performing schools in different districts. This average serves as a good indicator of the school districts' *overall* success in improving student achievement in these low performing schools across the past six years.

The results suggest that student achievement grows at statistically significantly different rates in low performing schools across districts, even after controlling for important school-level student characteristics like free or reduced price lunch, ethnicity, parental education, and English language proficiency.

Exhibit 3.25 shows, as an example, the test score trajectories of treatment and control elementary schools in four school districts in California. As can be observed, academic achievement improves at significantly different rates across these districts. For instance, in school district 3, the low-performing schools' API increased, on average, by 43 points a year, while low performing schools in district 2 (the one that shows the smallest improvement over

time on this graph) increased academic achievement, on average, by 9.6 points a year. Again, this is after controlling for important observable school-level differences that tend to be correlated with academic achievement.

Exhibit 3.25: Average Test Score Trajectories of Elementary II/USP and Comparison Schools in Four School Districts in California[38]



Summary of Findings

The findings from this chapter can be summarized as follows:

- Rising performance is observed among the state's low performing schools, including both II/USP and comparison schools, during the period of II/USP implementation. This increase suggests a possible overall effect from the state/federal accountability movement.

- There is still some indication of a planning year "bump" as reported in the 2003 PSAA Evaluation Report, but the evidence for this effect is weaker. Cohort 3 schools do not show this difference, and in Cohort 1 the effect is not significant for all measures.

[38] This graph shows straight lines because it estimates an average yearly contribution of each school district to their low performing school's API. A graph of all 9 districts is shown in Appendix A-11.

- Some modest positive II/USP effects are also found in the implementation years, however, they are not consistent across cohorts or grade levels. While statistically significant, all these effects are very small and most education research would not consider them educationally significant.

- Any gains that are made by II/USP schools in relation to comparison schools dissipate when II/USP comes to an end. One possible reason for this is that schools may have implemented certain strategies during II/USP that they were unable to sustain when funding ended. We explore this possibility in greater depth in Chapter 4. Another possible reason is that some comparison schools may have entered other programs such as Program Improvement during these years. Participation in these programs may have enhanced performance of the comparison group in relation to the II/USP schools.

- These analyses control for HPSGP participation for Cohort 3. For II/USP schools in Cohort 3 also participating in HPSGP, outcome measures generally show a statistically significant positive effect.

- ELs have made statistically significant gains since 1999 compared to EOs in II/USP and comparison schools across almost all cohorts and school types. However, evidence of an II/USP effect for English learners varies according to school type. No II/USP effect is observed in elementary schools; with an inconsistent pattern regarding II/USP participation seen in middle schools. For high schools, a significant positive II/USP effect is observed for ELs. This seems worthy of note given the consistency of this pattern across all three cohorts.

- In both comparison and II/USP schools, our data indicated that the gap between special education students and non-special education students has increased. Due to substantially increased test participation rates for special education students over this time period, this general trend is not surprising. No consistent II/USP effect is observed for students in special education.

- The achievement gap between students eligible for free and reduced price lunch (FRLP) and all other students in II/USP and comparison schools appears to increase from 1998 to 2000, but then decreases from 2000 until 2004. The II/USP effect for FRLP students is varied across cohort and school type. However, given the positive II/USP effect observed for Cohort 2 elementary and middle schools, as well as Cohort 1 and 3 high schools, there is slight evidence of a positive II/USP effect for them. Given the high percentage of FRLP students in all of these schools, it is difficult to interpret the significance of any possible direct II/USP benefit for this population.

- Results of the analysis of the CDE growth targets tend to suggest that improving academic achievement in low-performing middle and high schools is harder than in elementary schools. This finding seems to indicate that in middle, and especially in high schools, it is not only difficult to achieve the required growth targets, but also especially hard to sustain those levels of growth.

- When the district influence on school performance was examined, we find that II/USP schools tend to perform fairly similarly to non-II/USP schools within individual districts. However, student achievement grows at statistically significantly different rates in low performing schools across districts, even after controlling for important student

characteristics. This suggests substantial potential for districts to affect the performance of low performing schools.

Chapter 4 – Factors Contributing to and Hindering Growth in II/USP Schools

Overview

In Chapter 3 we discussed the overall impact of II/USP on participating schools by examining trends in student outcomes among II/USP and similar non-II/USP schools. As we discussed, when comparing overall achievement patterns, little evidence of a substantial impact from II/USP participation was found. However, these analyses only allowed us to look at impact in the aggregate. From the first evaluation of II/USP, it appeared that II/USP was having an impact on *some* schools. We found substantial variation in the outcomes among II/USP schools, with some experiencing considerable and consistent growth in student achievement, and others experiencing little or negative growth. Thus, for this continuation study, we identified a subset of II/USP schools that had made considerable growth during and after II/USP participation (II/USP "growth" schools) in an attempt to learn from them. Identifying these schools, and interviewing a key administrator and one teacher at each school, allowed us to explore in greater depth specific factors that contributed to their improvement. In addition, by including two other subsets of schools—a subset of II/USP schools that made minimal growth in student outcomes, and a subset of schools that made high growth, but did not participate in II/USP, we were able to compare the challenges and activities among schools realizing varying levels of growth through II/USP, and to explore factors that contributed to improvement in the absence of II/USP.

In the first section of this chapter, we discuss the factors identified by principals and teachers as contributors to their growth and progress. In the following sections, we discuss what was reported to us regarding challenges that the schools faced in their improvement efforts, the district's role in supporting low-performing schools, and II/USP's contributions to growth in student achievement. In addition, we discuss one important factor described as substantially influencing the work of low-performing schools in California—the implementation of the No Child Left Behind Act (NCLB). When possible, we compare findings from the growth schools to schools that did not make substantial growth. In addition, when relevant, we draw comparisons between teachers' and principals' responses. We reserve detailed discussion of the state-monitored schools for Chapter 5.

Summary of Findings

From our telephone interviews, we found that the most essential factors for schoolwide growth in student achievement were reported to be a coherent instructional program tied to the standards, leadership, professional development (including instructional coaches and teacher collaboration), and systematic assessment and data-based decision-making. II/USP was reported to have a variable impact on the growth of participating schools. In cases where II/USP reportedly *did* have an impact, respondents reported that II/USP provided vision and focus, needed funds, and a sense of urgency and accountability. Factors that affected the impact of II/USP, however, included the level of buy-in of school staff to change; the clarity,

specificity, and relevance of strategies implemented; the level of focus in implementing the Action Plan; and the leadership's ability to create an urgent and focused climate for change. In addition, late distribution of funds, a low level of guidance provided for fund use, limited communication between schools, and a lack of follow-up support and monitoring were reported as hindering the successful improvement of schools within II/USP. Finally, we found that the implementation of NCLB has begun to draw attention away from the state accountability program, as schools and districts try to balance the competing goals and demands of the two systems.

Assessment of Success in Improving Student Achievement

As explained in Chapter 2, the primary criteria used to choose schools for our sample were based on the CDE's measure of growth within the state accountability system—the meeting, or not meeting, of API growth targets. In addition, we used absolute API growth as a secondary criterion. By and large, the "growth" schools identified for our sample agreed that they had made considerable progress in improving student outcomes. Of the 20 principals of growth schools (both II/USP and non-II/USP), only one principal and one teacher (both from the same school) were skeptical about their level of success. This principal and teacher (of an urban II/USP high school) felt they had made progress in some areas, but still had a long way to go. This school seemed to be aware of the progress they made, while also wanting to sustain the sense of urgency that seems important to continued improvement.

For those that agreed they were successful in improving student achievement, several mentioned other ways in which they have grown beyond test scores. One principal from an urban II/USP elementary school put it this way:

> *I think it's been a huge turnaround for the most part. The staff morale is completely different; the parents' perception of the school is different. And all the assessment data we look at, we see there is a big improvement. We share that information with the parents and staff on an ongoing basis, so the more successful we become…we continue to work on it [more], and we see more growth. It snowballs that way.*

Other measures of success noted by respondents included a positive climate, a greater sense of engagement and commitment by teachers, and a common vision among teachers.

When respondents at state-monitored schools or at schools that made low levels of growth were asked if they were making progress in improving student outcomes, responses varied, though most thought they had made some progress. Among the principals at the four schools that were not participating in the SAIT process, two believed they had made progress. One had started focusing on completing monthly assessments and targeting the lowest-performing students, and expected to see results reflected in test scores this year. Another principal at a Cohort 3 elementary school felt they had made progress, but because of extreme changes in student population, the progress they made had not been reflected in test scores.

Top Factors and Strategies Identified as Critical for Growth

In this section we discuss the most common factors and strategies for growth outlined by school staff at growth schools (both II/USP and non-II/USP), and the ways in which they addressed challenges during their improvement efforts. We draw on data from schools that made minimal growth when appropriate to highlight differences between the two groups.

During the interview, we asked each principal and teacher to identify the top three factors that contributed to their school's growth. In addition, we asked them to identify the one factor that was most critical. We should note that many respondents emphasized that there were not just a few factors that contributed to growth, and that many factors were interrelated. In the words of one principal at an II/USP elementary school, "It's a combination of things, there is no *one* thing you can do for school improvement. It's a total change effort." In addition, upon further discussion of a specific factor, respondents would often mention several other related factors that were closely tied to the one they identified. Nevertheless, we used the ranking strategies described above in an attempt to get them to be as specific as possible.

We summarize the responses for growth schools below. Responses were coded using categories and subcategories developed based on our conceptual framework for the study. Exhibit 4.1 includes responses from both principals and teachers in terms of all three factors provided and the most important factor.

Exhibit 4.1: "Growth" School Staff Responses to the Question: "What factors contributed to your school's growth?" by detailed Factors for Growth

		Ranking Domain as #1		Ranking Domain as One of Top 3	
Detailed Factors for Growth		**N**	**%**	**N**	**%**
School and Staff Capacity					
A.	Leadership	5	14%	9	9%
B.	Instructional coaches/support	4	11%	5	5%
C.	Professional development	4	11%	8	8%
D.	Professional community/teacher collaboration	4	11%	16	15%
E.	School culture/climate	2	5%	11	11%
F.	Experience, qualifications and characteristics of instructional staff	1	3%	6	6%
G.	Resources: funding, etc.	0	0%	0	0%
Curriculum/Instruction					
H.	Coherent curriculum/curriculum & instruction tied to standards*	7	19%	18	17%
I.	Vision and common goals for instruction	2	5%	4	4%
J.	Additional instructional time for students	0	0%	1	1%
K.	Whole school reform model	1	3%	3	3%
Systematic Assessment and Data-Based Decision-making					
L.	Use of data and monitoring of student learning	6	16%	12	12%
M.	Attention to student outcomes	0	0%	5	5%
District					
N.	District support	0	0%	1	1%
Other					
O.	Parent involvement	0	0%	1	1%
P.	Other	1	3%	4	4%
Q.	No response	1		10	
Total		**38**	**100**	**114**	**100**

Note: Percentages were calculated from number of actual responses.
*These were originally coded as two separate subcategories (curriculum and instruction tied to standards, and coherent curriculum/instruction), but due to a high level of overlap between these two categories, we have grouped them together for analysis purposes.

As can be seen above, the major categories with the largest percentages of responses included school and staff capacity, curriculum and instruction, and systematic assessment/data-based decision-making. We discuss each of these major categories in the upcoming sections. Within curriculum/instruction we focus on two subcategories that we group together as "coherence of instructional program." These include "Coherent curriculum/Curriculum & instruction tied to standards" and "Vision and common goals for instruction."

A note of caution when interpreting these data is that factors may have been mentioned by a respondent in the interview, but not included in their list of top three factors. For example, in cases where the respondent focused on the use of data to inform instruction as a top factor, they may have also discussed how critical teacher collaboration around these data were. Thus the respondent's assessment that professional community is important would not be reflected as a specific factor for growth.

Appendix B-3 shows responses broken down by respondent type (administrators and teachers) for growth schools. In general, teachers' and administrators' responses were similar, though we note some differences in the following sections. One notable difference is the ranking of leadership. Teachers more often ranked leadership as one of the top three factors for growth (15 percent of responses) than principals (6 percent of responses). This may be due to principals' reluctance to recognize themselves as a primary factor for their school's improvement.

In the discussion below, we will return to the data presented above and make further comparisons among groups of schools, as well as between teachers and principals where appropriate. In this chapter, we draw on data from interviews where the factor was discussed, not just where the factor was specified as one of the top three.

Using our conceptual framework, we group the factors into three major categories below: coherence of instructional program, school and staff capacity, and use of data. These are the factors that interviewers reported as the most commonly cited overarching themes. The fact that their numbers above appear to be relatively low reflects the interrelated nature of many of the factors for effective change, and suggests that the exact paths to successful improvement vary somewhat by site. These factors also align well with those noted in our prior PSAA Evaluation (2003), and we provide context by referring to findings from that previous study. In some cases the factors were specifically tied to II/USP, and we discuss them in this context. However, we also provide a summary specific to the perceived influence of II/USP later in this chapter.

School and Staff Capacity

In the 2003 PSAA Evaluation, we found that school capacity was an important mediating factor in the development of curricular and instructional coherence, and therefore to growth in student outcomes. Capacity has also been cited frequently in research on instructional improvement. Cohen and Ball (1999) spoke of the complexities of school capacity, noting that capacity is "a function of the interaction among [teachers, students, and educational materials]" and not a component of merely one thing such as teachers' skill or curriculum. In order to understand instructional capacity, it is important to view the organization and environment in which these factors exist. In this section we discuss various aspects of school and staff capacity that were discussed by respondents as factors for growth during II/USP

participation. Aspects of school and staff capacity were the most frequently listed factors for growth. Fifty-four percent[1] of respondents from growth schools reported an aspect of school or staff capacity as the most important factor; similarly, aspects of school or staff capacity made up 53 percent of the total factors listed among the top three for growth for all schools.

We will first consider findings regarding school leadership—both the leadership of the principal and aspects of distributed leadership. We will then review forms and types of professional learning, including instructional coaching and teacher collaboration, that were considered to be important to the schools' success in improving student achievement.

Leadership

Leadership has long been considered one of the most important contributors to effective schools.[2] Effective leadership at the school site is critical to generating and directing school improvement efforts. In our interviews with staff at growth schools, leadership was one of the three most prominent factors reported for growth. As mentioned earlier, teachers more often reported leadership as a critical factor for growth than administrators.

Principal Leadership

First we consider the role principals played at schools that were successful in improving student outcomes. Principals and teachers typically described the role of effective principals as encompassing a variety of dimensions and qualities that contributed to other factors for growth that we discuss, such as instructional coherence and professional development, and to an overall climate that fostered improvement. Several themes emerged around these dimensions. First, respondents described effective leaders as creating an environment that is focused and sets high standards. Two principals and two teachers, for example, reported a focus on high standards and expectations for teachers and students, and on monitoring schoolwide progress. Others reported that the principal "enforced" the school goal or vision by being an example to staff and students, providing direction, and reinforcing common goals (e.g., school focus on improving vocabulary) whenever possible.

Others discussed principals as creating positive learning environments that are supportive of teacher learning and leadership. For example, one principal reported focusing on creating an environment where "all stakeholders treat each other with respect and dignity." Others mentioned ways that professional learning was fostered through staff development opportunities, grade-level or departmental meetings, and other opportunities for adult learning (e.g., coaching, observing other teachers). One teacher mentioned the value of seeking teacher input when planning these activities. Finally, some discussed encouraging teacher collaboration by creating time in the schedule for teachers to meet and work together.

One final common theme that emerged around the principal's role was the difficult work of developing a staff that is highly qualified and buys in to the need for change and reform. Several principals mentioned that both removing recalcitrant teachers and recruiting and supporting effective, highly qualified teachers were challenging tasks within their domain.

Several examples of strong leadership were found among our growth schools. Below we provide specific examples of some of the actions and strategies these leaders took at their schools. These examples draw on information provided by both principals and teachers.

[1] Some numbers may differ slightly from Table 4.1 due to rounding.
[2] See (Terry, 1996; Muijs, Harris, Chapman, Stoll, and Russ, 2004)

The principal at El Madrone Elementary[3] came to the struggling school five years ago. The first thing she did after joining the school was to work on improving the school's physical environment, for example by cleaning up the school and painting the walls. She bought more books and materials, and brought out the Action Plan created with the II/USP External Evaluator several years prior to her arrival. She worked with teachers to address the Action Plan, including focusing classroom instruction on standards-based lessons. She supported professional development to build capacity among the teachers and worked directly with teachers on a regular basis. She encouraged a culture of change at the school, and holds her staff accountable for student progress. The principal and staff work together, using student data to monitor results and progress, and to identify areas for further work.

One principal at Churchill, an urban elementary school, started at her school just after the Action Plan had been written. She fostered a sense of leadership among her teachers, and made efforts to make the environment focused on change and professional learning. As a doctoral student she infused the school with new ideas and research-based strategies she learned in graduate school. The teacher at this school described the principal as a "mover and shaker" and observed, "Without the principal, you can't do all this. She came in; she knows we were there before her and will be there after she leaves. She's given a lot of the power and decision-making to us. Some teachers were more reluctant since it was new and different. Now they are in it, they are happy with it. She also brought in literature; she has access to [research] literature from [her] doctoral program. We made faces about it, but she said we are a community of learners, we need to do this to improve teaching."

Five district respondents (of 14) also emphasized the importance of leadership as a top contributor to growth. As one respondent stated when discussing a growth school in his district, "[the principal] walks the talk" and provides the school with the "leadership, enthusiasm, and commitment to move the school ahead." Two district respondents noted that simply having some stability in the school leadership was important to growth, as so many lower-performing schools experience issues of consistent administrator turnover.

Distributed Leadership

Though it is difficult for a school to substantially improve without a strong principal, having an effective principal at the school site is not all that is required for improvement. Elmore (2000) suggests that moving leadership from more general, management and operational notions of leadership to focus instead on "the guidance and direction of instructional improvement" requires the *distribution* of leadership among staff. Elmore suggests that as leadership becomes more distributed, the quality of leadership will evolve by enhancing the skills and knowledge of staff, creating a common culture of expectations, fostering productive relationships, and "holding individuals accountable for their contributions to the collective result."

Many of these qualities outlined by Elmore mirror those we heard regarding strong leadership at the school site. Recognizing the importance of distributed leadership, we asked respondents about the presence of other leaders at their school, and how distributed leadership (when present) contributed to their growth or improvement strategies. When asked

[3] For confidentiality concerns, all the school names listed are pseudonyms.

about the distribution of leadership at the school, 85 percent (17 of 20) of administrators at growth schools indicated there is some form of distributed leadership with staff beyond the principal and vice principal. Nine principals indicated that instructional coaches and other area specialists were considered part of the leadership team. Eight respondents noted that grade-level representatives or department chairs are important to the school leadership. Sixty-four percent of principals (9 of 14) from schools that did *not* make substantial growth and SAIT schools also indicated that they have some form of distributed leadership in the school.

Several growth school principals stated that such distribution of leadership creates teacher buy-in to the improvement process because teachers have a voice in the direction and approach those efforts take. At an urban high school, a site leadership team has been created that is represented by every program and department. The principal encouraged teachers to take up organizing and planning the professional development at the school, which resulted in a writing task force this year. A teacher from this school reported that the distribution of leadership among all staff enhanced collaborative practice and the "positive nature of the school."

Professional Development, Professional Community, and Instructional Coaches

Of the complex factors contributing to the individual and organizational aspects of instructional capacity, Cohen and Ball (1999) suggest that opportunities for professional and organizational learning are among the most crucial. We consider several forms of professional development here, ranging from workshops and trainings to the use of instructional coaches and teacher collaboration within grade levels or departments.

High quality professional development opportunities were reported to result in greater confidence among teachers in the use of standards and curricular materials, the analysis of student work and data, and the planning and differentiating of instruction. Overall, 58 percent of respondents at growth schools and 54 percent of district respondents cited professional development, teacher collaboration, and/or the use of instructional coaches as one of the top three factors for growth. In fact, professional development was ranked as the second most important feature of capacity (see Table 4.1). Several principals cited some challenges with professional development as well: not enough time to implement what is learned, substitute time needed to release teachers, and cost.

A variety of professional development opportunities (both in terms of form and content) were emphasized by responding principals as building teacher capacity or improving instruction at the growth schools. These included traditional forms of professional development such as weekly staff development meetings and trainings, and inservices provided by external providers such as consultants, district or county education office personnel, or through partnerships with local universities. In addition, respondents frequently noted the use of less traditional forms of professional development such as coaching, modeling of lessons, peer observations, collaboration with other teachers in the same department or grade level, and professional development coordinated or led by teachers. Professional development on using data and/or monitoring student instruction was the most common content area mentioned.

Typically, respondents reported that combinations of professional development activities were present at schools. For example, at one urban elementary school, they implemented a comprehensive professional development program that integrated and addressed teacher learning in a variety of ways. This program included three primary strategies: the hiring of

two instructional coaches who provided modeling and feedback to teachers (both struggling teachers and teachers who could be teacher leaders), observing teachers at other schools to view other instructional strategies and incorporate them into their own classrooms, and developing professional development around two focuses: writers' workshops and math problem solving.

We now focus specifically on two less traditional forms of professional development—instructional coaching and teacher collaboration.

Instructional Coaches and Facilitators

Principals at 69 percent of schools[4] cited the use of literacy and math coaches and other instructional facilitators at their schools as important components of their improvement strategies. Coaches were reported to increase staff capacity by providing subject matter expertise and hands-on, individualized professional development. They were reported to assist teachers with modeling and observing classroom lessons, managing and interpreting student data, trainings, and other professional development. They also assisted with strategies to improve instruction and planning for students needing additional support. They were reported in several cases to be critical to implementing improvement plans.

At one middle school, for example, where instructional coaching was described as one of the three most important factors for growth, both reading and math coaches were hired. They hope next year to have science and social studies coaches. The current coaches work with teachers in their classrooms by observing and modeling lessons, then following up with them to discuss what they did. The coaching work has happened on an almost daily basis for the last three years, and was reported to be so successful that the district decided to fund the position.

Some schools hired coaches themselves, but in others the district provided coaches. District-level coaches often worked across schools, which potentially fostered increased coherence within the district. Some respondents reported that II/USP funds were used to hire coaches, though principals reported that this type of instructional support personnel was difficult to sustain when funds ended.

Professional Community/ Teacher Collaboration

A strong professional community with a high level of teacher collaboration enables teachers to learn from each other, share ideas and strategies, and review student outcome data to improve instruction. Respondents at several schools reported that II/USP and the state monitoring process provided funds for collaboration.

Respondents reported a variety of forms of teacher collaboration. The frequency of collaboration reported ranged from two to three times per week, to monthly. Some reported that collaboration occurred during one-hour meetings, while others reported meeting up to several hours at a time. Collaboration meetings were held by grade level, class, or team structures. In general, however, respondents reported that *regular* and *frequent* opportunities to collaborate were important for developing a strong professional community, learning from each other, discussing data, and planning instruction. Collaboration among grade-level or departmental (in secondary schools) colleagues was most often reported as useful.

[4] Excluding SAIT schools

Teachers and principals reported the primary focus of teacher collaborative time as instruction-related and centered on a coherent set of goals. The activities included planning lessons together; reviewing and sharing teaching strategies; analyzing and reviewing data, such as performance based assessments; and identifying areas of concern in instruction and strategizing methods to address these areas. In addition, several teachers mentioned that teacher collaboration at their school was coupled with observations of their colleagues' classrooms or work with coaches. These observations and coaching provided discussion topics for meetings focused on implementing instructional techniques. Below is an example of one growth school's efforts to foster a strong professional community:

> *La Fortuna Elementary, in a suburban, northern California district reported collaboration as key to their growth. Teachers met three times a week by grade level during common planning time. During this time teachers met with coaches and discussed their English language development program. They examined assessment scores from benchmark assessments, and planned lessons accordingly. In addition, the school had weekly staff development days in which staff examined and discussed instructional goals. Students moved back and forth between classrooms and different levels of instruction. Therefore, every six to eight weeks during staff development time they reviewed benchmark assessments, made adjustments to instruction, added interventions, and regrouped as necessary. As the principal said, "this is collaboration all the way," since teachers shared students.*

Coherence of Instructional Program

Instructional coherence, an outgrowth of greater coordination and alignment of goals, activities, and resources at schools, has long been found to be a key component of effective school organizations. In the previous study, we found a strong relationship in our case study schools between the presence of a coherent instructional program and improvements in student achievement outcomes—that is, schools with more coherent programs also demonstrated greater and more consistent gains in student test scores. Indeed, a central goal of the II/USP planning process was to develop greater instructional coherence in schools.

In this follow-up study, 26 percent of responding teachers and administrators at growth schools cited a coherent instructional program as the most important factor contributing to their school's improvement. Here, we assess the contribution of coherence to growth by examining the factors listed by respondents along two interrelated dimensions of coherence: a common vision/focus for the school and coherence in curriculum and instruction.

Coherence Through the Establishment of a School Vision or Focus

An important component of instructional coherence is a schoolwide vision or set of common goals related to instruction (O'Day & Bitter, 2003). We define this vision or common focus as one that permeates the schools' activities, guiding professional development and fostering teacher collaboration across a clear set of goals. A clear theme that emerged from our interviews was the importance of this vision or common focus to school's improvement efforts. Strategies reported to foster such a vision included creating a mission statement for the school, developing a set of schoolwide goals, or focusing on one subject area to improve student achievement. Principals described this as having a "unity of purpose," "vision building," and devising a meaningful and strategic focus.

Respondents that reported having a school vision or focus spoke of its practical applications and the importance of this vision/focus being actively implemented, as opposed to a plan that is developed, "put on a shelf," and forgotten. Respondents at growth schools reported incorporating the vision/focus into school activities such as instructional planning, professional development, and data assessment, and described it as ubiquitous throughout the school. For example, at one urban Cohort 3 elementary school, the teacher described the vision they formed around high student expectations and instruction tailored to the needs of their student population. The school "filtered everything through [their] vision," including staff development and the articulation of instructional goals. This respondent noted the importance of having an "active" vision, not just a vision on paper. In other words, the vision was used to guide core activities in the school and evolved each year when the whole staff met to rethink the goals together.

Other schools developed a common focus on a specific area such as English language arts (ELA), vocabulary, English proficiency, or writing. Seven growth schools cited having a schoolwide focus on one academic area. In contrast, no respondents from schools that made low levels of growth reported having such a schoolwide focus. Respondents emphasized that creating a schoolwide focus is not a trivial task. Below is the description of one school's development of a focus, the practical implications of this strategy, and the instructional coherence it brought to the school.

> *As part of the II/USP planning process, Mirafuentes Elementary developed a specific academic focus for the school. They began by reviewing and analyzing SAT-9 scores. The school then held discussions, where teachers were given the opportunity to select a goal area. The principal ensured that there was formal buy-in from all stakeholders, which took several months to accomplish. A cross-section of stakeholders (administrators, teachers, and parents) held meetings and discussed the goal area. The advisory and site council were also consulted. By consensus, the school created the goal of building students' vocabulary. The school drew up a plan, activities, and benchmark measures that would result in raising vocabulary and student achievement. The principal was resolute in having the plan implemented. He took the plan "seriously" and kept it clear and simple to facilitate implementation. The school implemented activities aligned to the goal, developed assessments for vocabulary aligned to the standards, and added a block for teaching vocabulary apart from language arts. During monthly grade-level team meetings, teachers discussed vocabulary and related instructional strategies. Thus, at Mirafuentes Elementary the focus on building vocabulary permeated throughout the school, leading to an increased level of instructional coherence.*

In this case, the development of such a focus was attributed to the planning process in II/USP. Several other respondents that reported a schoolwide focus discussed the importance of II/USP to the development of this common focus. More information about this is discussed under the *II/USP Influence and Strategies* section of this chapter.

Coherence Through Standards-Based Curricula and Instruction

A second interrelated strategy reported for developing coherence was having a curriculum tied to the standards, or a common curriculum or instructional plan through the school. As we saw in the last study, having a common curriculum fostered consistency across classrooms in terms of what students were learning and how they were learning it. Through this

consistency, having schoolwide alignment of curriculum and instruction facilitated teacher collaboration. Three major strategies discussed by respondents for developing coherence in curriculum or instruction are discussed below: adopting a common curriculum, developing a pacing or master plan, and aligning instructional strategies across the school.

Common Standards-Based Curriculum. Since California's adoption of academic content standards in 1998, schools throughout the state have increased their focus on implementing a standards-based curriculum. In 2002, the state transitioned to a standards-based assessment system, creating more incentive for schools to align their instruction and curriculum to the standards. Ideally, aligning curriculum to the standards will foster curricular coherence by fostering consistency across classrooms and grades, facilitating teacher collaboration around the standards, and aligning professional development. Sixteen percent of respondents from growth schools felt curriculum and instruction tied to standards was the most important factor for growth. Seventeen percent of respondents at schools that did not make substantial growth also felt it was the most important factor in the progress they had made.

Adopting a common standards-based curriculum across the school was a common method used to align instruction with the standards. In our interviews with growth schools, the curriculum most often mentioned was Open Court. Other curricular packages growth schools reported adopting were Houghton-Mifflin and Harcourt Brace.[5] One principal noted that the district provided a reading coach and assessments to support the implementation of the curriculum.

At several schools, II/USP was cited as contributing to the alignment of their curriculum to the standards. In a few cases the External Evaluator was reported as emphasizing the need for the school's curriculum to align with the standards. The State Assistance and Intervention Teams that work with II/USP schools that have not made consistent and targeted progress also work on standards alignment. The principal at one SAIT elementary school, for example, reported that participation in II/USP clarified the importance of following the standards and learning methods to assess them. More information about the SAIT process will be discussed in Chapter 5.

Pacing or Master Plan. A pacing or master plan is another technique schools reported utilizing to foster curricular coherence. Pacing plans enabled teachers to hold discussions about the curriculum, plan lessons together, and share strategies to improve student comprehension and learning. Respondents at six of the growth schools reported using a pacing or master plan; no respondents from schools that made low levels of API growth reported using this technique. Growth schools implemented such a plan in varying ways—for example, by focusing on spiraling for grade levels, having a district liaison work with teachers to create a pacing guide, or developing a curricular "map." Below is a description of one large urban high school's process for creating a curricular map for its core classes and the impact they had on curricular and instructional coherence at the school.

> *The vice principal of Citrus Valley High School, in collaboration with a subset of teachers, developed curricular maps for the school's English and math courses. They used the standards to lay out an entire course, including the topics to be taught and the schedule for teaching them. The vice principal reported that the curricular maps became templates for the courses by breaking down the*

[5] Since several large urban districts have adopted Open Court districtwide, this breakdown may be a result of an overrepresentation of sampled schools in these districts.

standards and identifying areas to introduce supplemental materials. The curricular maps were reported to facilitate teacher collaboration, since teachers stayed on the same schedule and administered the same benchmark assessments regularly. Every two to three weeks teachers assessed their students using a 12-question test created by an outside publisher. Each teacher received an analyzed report of the assessment results by class. The vice principal and teachers would then meet to discuss the assessment results and the standards students were struggling with or learning. Teachers returned to the classrooms and made adjustments to their teaching and curriculum as needed. For Citrus Valley High School, the curricular maps not only created curricular and instructional coherence within courses, but they opened the doorway for teacher collaboration.

One district representative and the associated schools discussed the key role the district played in the development of instructional alignment. In this case, a district liaison worked with schools across the district to align their curriculum to the state standards by developing pacing guides and benchmark assessments.

Alignment of Instruction. Although mentioned less often, alignment of pedagogical techniques across classrooms was another strategy schools used to create instructional coherence. Three growth schools reported fostering the use of specific teaching strategies through the school, such as explicit direct instruction, thinking maps, or critical thinking techniques. In these cases, respondents reported the importance of professional development and training on these techniques to help reinforce the instructional strategies among teachers.

A schoolwide vision and curricular coherence are both important factors to improved school outcomes. As has been shown in the examples above, these factors rarely exist in a vacuum; rather, they are implemented in conjunction with other strategies that enhance schools' improvement efforts and increase student learning.

Systematic Assessment and Data-Based Decision-Making

In recent years, California public schools have engaged in numerous federal, state, district, and school-mandated assessments to monitor student progress. Data from these tests can be invaluable to schools by assessing the needs of students and pointing to areas for improvement (Neumann, 1996; Kannapel and Clements, 2005). Study respondents at 25 percent of schools[6] listed systematic monitoring and data-based decision-making as the most important factors for growth. We also asked respondents to rate on a scale of 1 to 10 the importance of using data for instructional planning to their growth, with a 10 meaning it was the most important factor in their growth. The mean rating by growth school respondents was an 8.5. Below we discuss assessment and data usage patterns among our sampled schools.

Many respondents emphasized the use of frequent, curriculum-based and standards-aligned assessments that could provide ongoing measures of progress. These most often included assessments associated with Open Court and Houghton-Mifflin, which can be administered on a six-week or quarterly basis. In addition, some schools created their own assessments that aligned with state standards, or used district-developed benchmark assessments. State assessments, including standardized tests such as the CST, CAT/6, and CELDT, were

[6] Excluding SAIT schools

reported to be useful at the beginning of the year in planning instruction, but were not frequent enough to be used for data-based instructional decision-making.

Nineteen respondents who discussed data usage each reported using data in more than one way. For example, one school used data to determine whether students were passing or failing, to identify the areas for improvement, and to form instructional strategies in ELA and mathematics. Teachers at another school used data to plan units of instruction, create specific strategies for needy students, and gauge student understanding. Foremost, teachers at these schools were reportedly involved with ongoing data reviews. In nearly 80 percent of growth schools, teachers reported discussing data regularly at grade-level meetings, in professional development sessions, and at other staff meetings. Teachers at 12 schools reported analyzing and interpreting data, while at nine schools they were said to receive data analysis or interpretation training, and collaborated with instructional coaches, peer coaches, resource specialists, principals, or assistant principals. To analyze data, teachers and administrators at several schools reported using data reports generated by district-provided software that consolidated standardized and benchmark test data. From their data, teachers were trained to identify the benchmarks and grade level standards that students failed to meet. For example, at one urban elementary school, teachers reviewed the district-produced assessment data for each student to identify and discuss every question that students responded to incorrectly. They analyzed the content of each question to find out what specific areas of instruction students missed, and sought to determine why students missed these questions.

Respondents at growth schools reported using data to identify student needs and to guide instruction. Based on these needs, teachers collaborated in grade-level meetings or with administrators to adjust and differentiate instruction or plan specific instructional strategies, units, areas of focus within subjects, and interventions for students. Respondents discussed various forms of instructional supports and interventions that were implemented as a result of these analyses, including math and English clubs, after-school programs, and tutors for students who needed extra support.

Teachers at many of these schools reported also meeting regularly to follow up on student progress after each periodic student assessment. They reported analyzing and interpreting data, modifying instruction, and implementing interventions based on individual and grade-level student needs, tying instruction to standards, reanalyzing data to determine student progress, and modifying instruction further. This cycle, enacted through teacher and school administrator collaboration, repeated throughout the academic year.

A few differences emerged between the data use practices of growth schools and schools that made low levels of growth (including two schools that were also SAIT schools). The main difference between reports from these groups of schools was the extent of data use to inform instruction. Growth schools reported using data frequently and extensively to guide instruction, while the other schools, more often than not, reported using it to generally identify students not performing well. For example, respondents at two low-growth schools reported using data to identify needy students and unmet standards. At one of these schools, teachers used data during grade-level meetings to identify and target students on the cusp of grade level proficiency. Respondents at several schools did indicate using data to target and guide instruction, though no further specifics or examples were cited. Respondents from growth schools more elaborately discussed use of data to inform collaboration, needed instructional adjustments, realignment with standards, intervention strategies, reanalysis of student progress, and instructional modifications.

Challenges to Improvement Efforts and Strategies to Address Them

In addition to discussing factors critical for school improvement, we also explored the barriers respondents faced in their improvement efforts. Principals and teachers discussed a wide range of barriers, but the challenges mentioned fell into four primary categories. The most common challenge cited by respondents (39 percent) was serving the needs of their student populations. These included challenges associated with student behavior, a transient student population, and a low level of student preparedness. The next most cited challenges were developing a highly qualified, skilled, and invested teaching staff (19 percent); increasing parental involvement (14 percent); and resources and funding (6 percent).

Addressing Student Needs

Respondents reported facing many challenges in working with their student populations. Specifically, 27 respondents discussed the challenges associated with serving diverse groups of students, including large populations of English Learners, students in poverty, or special education populations.

Substantial numbers of respondents also reported being challenged by a lack of student motivation, a transient student population, discipline/behavioral issues, and low levels of student preparation. The principal from one urban elementary school, for example, spoke about behavioral issues the school faces. As he reported, students like to "solve their problems with their fists." There is an enormous violence cycle the school must address. And at an urban high school, the vice principal noted that less than ten percent of the incoming ninth graders read at grade level, with about 70 percent reading three or more years behind grade level. The vice principal said, "I was a good teacher, but there's no way that I can bring in a kid reading at a second-grade level and by the time they are in tenth grade get them reading so they can pass the CAHSEE [California High School Exit Exam]."

Respondents reported implementing a wide range of strategies to address these numerous challenges. Most often reported was the use of intervention programs. These included mandatory after school tutoring, counseling, peer mediators, dropout prevention programs, character education classes, study skill classes mandatory for students who have three or more failing grades, and an extended learning program on Saturday. At one school a performance arts program was brought in to introduce "culturally responsive education." Two schools also reported implementing a series of incentives to reward good attendance. At one of these schools, the principal calls home when a student is absent. Intervention programs were also used to address student motivation and behavioral issues.

Others addressed the needs of their students by enhancing instruction and undertaking instructional improvement efforts. Respondents, for example, reported that their schools incorporated additional instruction time, focused on a subject area, and/or provided training to teachers. To address the needs of ELs at one school, they reported focusing on differentiated learning and English Language Development (ELD) instruction. They also reported providing professional development on EL instructional strategies, providing supplemental materials for ELs, and implementing ELD portfolios to track student progress.

Establishing a Skilled and Invested Staff

Respondents reported a variety of challenges faced in establishing a skilled and invested staff, ranging from high levels of teacher turnover to low teacher motivation. While some respondents mentioned staff stability as a top factor for their growth, teacher turnover was the most often cited challenge associated with establishing a skilled and invested staff.

Teacher Turnover

We asked all respondents about the level of teacher turnover at their school, and how the level of turnover affected their growth or their improvement efforts. The level of turnover varied between growth schools, schools that made low growth, and state-monitored schools. Of the 20 responding administrators at growth (II/USP and non-II/USP) schools, only 25 percent reported a high level of teacher turnover. The generally low level of teacher turnover was reported to be a facilitator to improvement efforts because it builds understanding among teachers and creates opportunities to develop the instructional program. At the six schools that made minimal growth, however, (including the two that are state-monitored), 67 percent of respondents reported a high level of turnover. As described by the principal of one elementary school, "A few years it was 12, last year it was 15 teachers who left. We have a total staff of 36." Approximately half of the state-monitored school respondents reported a high level of teacher turnover. One said teachers left because of their classification as an intervention school, and another reported not being able to find qualified teachers. This variation suggests that teacher turnover is more of a challenge in the low-growth and state-monitored schools, and can potentially be affected by the classification of a school as an intervention school.

Teacher Motivation, Morale, Qualifications, and Skills

In addition to teacher turnover, respondents cited low levels of teacher buy-in or motivation as a challenge. At one middle school, the principal introduced the use of instructional coaches. However, teachers felt threatened by the coaches. To obtain buy in to the strategy the principal invested time in faculty meetings and professional development to reiterate the benefits of instructional coaches. The principal had constant communication with teachers, and allowed them to speak freely. The principal walked into classrooms to reinforce the instructional coaches, and provided supplemental materials and positive feedback to teachers.

Finally, teacher union issues, low staff collaboration and low staff qualifications were discussed as further challenges associated with establishing a strong skilled and invested staff. At one elementary school, the principal noted difficulty removing ineffective teachers and documenting ineffective instruction. To address this issue, she identified teachers' strengths and weaknesses and used an evaluation system to notify teachers of areas of concern. The principal met with the teachers, observed practice, and provided assistance. Several principals reported that removing staff that did not buy in to the reforms was an important, but difficult, process.

Increasing Parent involvement

The importance of parental involvement to school success has often been cited in research.[7] Despite the importance of parental involvement to a child's success many respondents

[7] For example, see Desforges and Abouchaar, 2003.

reported struggling to involve parents in school activities. Respondents reported parent involvement as a challenge more often than as a factor in a school's growth. Only five percent of respondents listed parent involvement as one of the three most important factors for growth. However 35 percent of respondents listed parent involvement as one of the greatest challenges to the school's improvement efforts.

Principals and teachers discussed some of the challenges faced with parental involvement. Some respondents believed the parents of their students did not value education; others felt that parents who did not speak English were limited in their ability to assist their children with schoolwork.

One common strategy used to increase parental involvement was for schools to sponsor workshops for parents. Seven respondents reported using this strategy. This included parent education classes, which focused on parenting skills, or parent conferences showing parents classroom activities. One school participated in GearUp, a program that teaches parents how to help kids with reading. In addition to parenting classes, one school offered ESL and Even Start classes and a course for parents to complete their General Education Diploma (GED). One rural elementary school employed multiple approaches, including obtaining a grant for teachers to make home visits with parents, initiating a program in which parents serve as tutors after school, and planning weekly meetings with parents.

Finally, several respondents reported hosting events at the school to involve parents more in the school's activities. Some events schools hosted included family math, science, and reading nights; a pajama party, where parents and students went to school in pajamas and read together; and an arts festival. Schools also tried to involve parents by keeping them informed through parent newsletters and parent/teacher conferences, and having them participate in the school site council. In addition, three schools reported that II/USP helped articulate parental involvement and provided funds for a parent liaison and parent workshops.

Funding

Similar to parental involvement, funding was more often considered a challenge than a factor for growth. Only 3 percent of respondents listed funding and resources as one of the most important factors for growth.[8] On the other hand, 15 percent of the respondents listed funding and resources as one of the greatest challenges to growth.

Funding is critical for school improvement, but as one teacher said, "money doesn't buy great instruction." Additional discretionary funds either through II/USP or other sources, though, allowed schools to pay for instructional coaches and teacher collaboration time, after-school programs, supplies and materials for teachers, professional development, and upgrades to school facilities. However, decreases in funding, either due to the end of II/USP or other funding mechanisms, often meant reduction of supports and services. Two principals reported that receiving additional funds, such as those from II/USP, can be a challenge because it is one-time money with restrictions on its use. Another principal said that because he received funds from different programs he did not have the latitude to make adjustments to the school's budget without engaging in a tedious bureaucratic process in which he sends budget adjustments to the district, which is sent to the state, and then approval to make adjustments is received several months later.

[8] However, it should be noted that funding was often reported as a means to implement other strategies listed as top factors for growth.

The only strategies reported by respondents for addressing funding issues were to apply for grants or pass school bonds. Two principals reported applying for grants to supplement the school budget. However, only one of these principals had success in receiving grants. This principal worked with a consulting company that assisted him in writing and submitting grants. At this same school they were successful in passing a bond to build a multi-purpose building.

District Supports

In a review of recent research literature on the role of the district in school-level reform efforts, Marsh (2000) found that districts can play a "potentially critical role in improving teaching and learning." Districts acting as change agents tend to mobilize critical resources (human, social and physical), thus having a "better chance of enacting and sustaining state and local reform goals and policies."

We also found in the 2003 PSAA Evaluation that districts can influence the progress that schools make toward achievement goals and improving outcomes for students. This finding stemmed from both achievement analyses that demonstrated significant differences in achievement trajectories among several large urban districts in California, and from case study data that indicated that districts played an important role in both II/USP implementation and improvement strategies.

Our interviews in this study supported this finding further and indicated that there is variation in the level of support for school improvement efforts across districts. In some cases, districts implemented specific policies and supports for low-performing schools, in other cases districts implemented further accountability measures, and in other districts supports for low-performing schools were reported to be nearly void. In light of this variation, we consider what districts were reported to do that best supported schools in making systemic improvements.

While none of the principals noted district support as one of the top three factors for growth, they were asked about the district role in efforts to improve their school. On a scale of 0 to 10, with 10 as the most important factor for growth, administrators at growth schools gave the district support a mean score of 7.63. Nearly half of the principals of growth schools who responded to this question rated the district a 9 or 10. Respondents listed a broad range of policies and supports, including planning and pacing guides for instruction, funding for instructional coaches and additional textbooks, professional development for staff, coordinating assessments, data analyses and assisting with interpretation of results, and helping schools apply for grants.

In addition to these tangible supports, a few principals lauded the district for less tangible forms of support. In these cases the district helped to provide focus for schools or fostered a supportive environment. For example, one principal stated that their local district (which was part of a large urban district) gave them a "sense of where they are going;" another noted that the district's focus "helped in aligning the vision of the school."

The interviews with district respondents provided further context about the district role. Interviewed district staff listed a variety of strategies to enable improvement efforts in struggling schools. In some cases, district respondents reported systems that addressed *all* low performing schools through additional supports (such as coaches) or additional accountability requirements. One respondent from an urban district, for example, reported

that they established a program to identify schools within and beyond II/USP that were failing to improve and implemented various levels of sanctions at these schools, such as restructuring. Another reported encouraging and moving principals from higher performing schools to the lowest performing schools (that were facing sanctions under Program Improvement). Several respondents reported that while they had incentives in place to recruit teachers in high-need subject areas (such as mathematics and science), they were unable to establish programs to redistribute highly qualified staff to low performing schools due to union contracts. Other common strategies and supports discussed include the establishment of a common curriculum, additional assessment requirements, and professional development. These primary strategies mentioned by districts aligned well with the strategies mentioned by principals as useful for their improvement efforts.

Common Curricula and Assessments

Most of the districts in our sample (85 percent) have adopted some form of districtwide common curriculum at either the elementary or secondary level or both. A few districts noted that they included teachers in the process of selecting curricula. One district, for example, reported having teachers across grades come to learn about the various new texts, then pilot them in the classroom for about 12 to 16 weeks, and then come back together to analyze the strengths and weaknesses of each. They then voted on which package to choose for the district to adopt. Overall, the decision to move to a common curriculum across the district was noted to be a key factor for growth among their other low-performing schools. In addition to developing their own curriculum and corresponding assessments based on state standards, they also instituted a textbook adoption program in an effort to get all the schools in the district "on the same page."

In addition to districtwide curricula, 69 percent of district respondents noted that they have additional testing requirements in place for their schools, beyond state-mandated assessments. Two district respondents specifically listed the use of data or other assessment tools as a key factor for school improvement among low-performing schools in their districts. Among those who have added testing requirements, most are administered as trimester or quarterly assessments, primarily in English language arts and mathematics. Just over half of the district respondents (54 percent) noted that they provide technical assistance to their schools around the use and interpretation of assessment data. One southern urban district is now able to give information on "individual students by individual standards" so that principals and teachers can see exactly where to bolster instruction in particular content areas or for specific subgroups. In addition to being able to use the data more effectively, the district also has the capability to look at trends over time with individual students to see why there were changes.

Professional Learning

Many principals and teachers (40 percent)[9] mentioned professional development and instructional coaches as key district supports to the school. A principal of a school in a large, urban district described the professional development opportunities they have received as "very focused," "raising the bar with critical thinking and problem-solving with students," and "in-depth in terms of classroom practices." He found the opportunities to be of high quality and consistency.

[9] Excluding SAIT schools

District respondents also reported that professional development and instructional coaches are among the important supports they provide to schools. Almost all (92 percent) of district respondents listed some type of professional development or coaching activities that the district supports, though there was a broad range in regard to the kinds of professional development provided. Several districts noted the provision of mandatory trainings for curricular materials such as Open Court and Houghton Mifflin, and workshops on bilingual education or English language development strategies. In addition, two district respondents discussed the partnerships the district created with local universities to provide professional development in particular content areas (e.g., mathematics) or with local chapters of the California Subject Matter Projects to provide additional trainings for staff.

Additional Supports Needed

Principals were asked what additional supports they would like to see from the district. One principal from an urban middle school reported that he would like more guidance from the district on creating and conducting professional development, and another high school principal suggested additional opportunities for professional development. The principal of an elementary school in the central valley stated that he would like more guidance from the district on the instruction of English learners—a growing population in the central valley. They have asked for additional support from the district, but are told they are the "shining stars," and do not receive sufficient guidance on how to improve instruction for their EL students. Several principals listed funding from the district as problematic for their schools—either that they would like additional funding or would like more flexibility from the district on how to spend that money. Issues around budgets and funding were listed as a challenge by many district-level respondents as well.

II/USP Influence and Strategies

We have outlined several key contributors to growth in schools that consistently met their growth targets or made high growth in student achievement within II/USP. We next consider the extent to which II/USP was reported as contributing to the implementation of these factors and to the overall growth of the schools. The theory of action outlined in Chapter 1 hypothesizes that the planning and work with an External Evaluator, coupled with additional resources and the threat of sanctions, will result in the creation of coherence and increased capacity, and therefore improved student outcomes. However, our findings from the first evaluation and from this follow-up study do not show enhanced growth, on average, in II/USP participant schools.

In this section we start with II/USP schools showing substantial growth during this period and discuss the extent to which respondents believed that II/USP specifically contributed to this growth. Next, we turn to interviews from all sampled II/USP schools to examine why II/USP seemed to affect some schools more than others, and to look for aspects of the policy that may have impeded growth during this period.

Conducting this follow-up study several years after the first cohort completed participation in II/USP had both benefits and drawbacks in terms of exploring II/USP's effect. As a benefit, it allowed us to explore the effects of II/USP over a longer period of time—to what extent strategies implemented during II/USP were still in place and to what extent the program itself appeared to contribute to growth among schools.

A limitation associated with this timing, however, was that in many schools, staff turnover and the passage of time resulted in minimal knowledge of II/USP. In these cases, administrators, and more often teachers, were unable to distinguish between reforms associated with II/USP and those associated with other programs (e.g., Reading First, HPSGP). New administrators often did not know which program funded initiatives that started prior to their time at the school.[10] Of the 30 II/USP administrators interviewed, only 11 had a high awareness of II/USP and were able to discuss related reforms in great detail. Six principals had minimal or no knowledge of II/USP because they were not present for most of the reform, while the remaining administrators outlined only the basic reforms they undertook.

For teachers, an even larger portion had minimal to no knowledge of II/USP (10 out of 29 interviewed). Since these teachers had all been at the schools for at least three years, and in many cases far more, the minimal knowledge of these teachers indicated that II/USP as a program had not fully permeated the school. They more often remembered having an external person assisting their school at the beginning of the program, with less knowledge of the reforms associated with it during the implementation years. For example one teacher mentioned that she had heard her formal principal mentioning II/USP all the time, and was aware of the External Evaluator assisting her school in the first year, but had very little knowledge of the program beyond this. Only nine teachers interviewed had a high awareness of II/USP; the remaining had an understanding of the basic reforms associated with II/USP.

In this section we discuss three specific questions:

1. To what extent did II/USP contribute to progress in "growth" schools?

2. Why did II/USP facilitate growth more in some schools than in others?

3. What aspects of II/USP could be improved to better facilitate growth?

To what extent did II/USP contribute to progress in "growth" schools?

To answer this question, we must first consider the dilemma associated with the minimal knowledge of II/USP among respondents. Among growth schools, three principals and six teachers had minimal to no knowledge of the strategies employed using II/USP funds. Though they were able to discuss successful reforms implemented during the time of II/USP, we unfortunately cannot make a direct connection between II/USP and these reforms. When these respondents were asked what contribution II/USP had to their reforms, they often responded that they were certain the funding contributed in some way, but discussed less the influence of the planning year or an External Evaluator.

To further assess the contribution of II/USP to growth, we examined interview responses from principals and teachers from our growth schools who had at least a basic knowledge of II/USP (13 principals and 9 teachers). Respondents were asked about II/USP's contribution to growth and what component contributed the most. In general, these respondents reported that II/USP played a significant role in their improvement efforts. We asked respondents to rate the importance of II/USP to their school's growth on a scale from 0 to 10, with 10 being the most important factor for growth. Six of 13 principals from growth schools with knowledge of II/USP rated II/USP a 9 or a 10 on this scale. Only two of the nine teachers who had at

[10] Eleven principals reported serving as principal of their current school for two or fewer years.

least basic knowledge of II/USP rated it a 9 or 10. Respondents reported II/USP contributing to growth by providing a vision and focus, funds, and a sense of urgency and accountability. Below we discuss these factors.

At several schools respondents reported that II/USP provided vision and focus. Eight principals and five teachers from growth schools reported that the planning process and/or work with the External Evaluator were components of II/USP that contributed strongly to their growth. Typically, the planning effort was reported to develop a focus or vision for the school. Primary strategies were focused around the factors for growth discussed in the previous section: implementing curricula and assessments tied to the standards, increasing teacher collaboration, and developing common goals. For example, one principal reported that developing a shared vision helped to get everyone on the same page and build a positive school culture. In the words of one teacher, whose school used II/USP to adopt a reading program, implement a focus on standards, and increase expectations for students, "The planning was great. The planning was what brought us up, created goals and objectives for the years to come. That was the most important thing."

The additional funds provided by II/USP were reported by many schools to be important. Eleven principals and six teachers of growth schools mentioned that the funding associated with II/USP was a component that contributed strongly to their school's success in improving student outcomes. The funding allowed them to implement strategies that they would not have otherwise had the opportunity to do, or allowed them to implement them in a more intensive and accelerated way. The most common strategies implemented with II/USP funds mentioned by school respondents included the following:

- **Professional development.** Respondents reported that II/USP funds were used to provide professional development around content standards, new curriculum adoptions, or academic areas of focus determined through the planning process. In addition, at some schools teachers were given opportunities to visit other schools or release time to focus on analyzing data or planning instruction.
- **Additional instructional staff/coaches**. These positions included reading coaches, mathematics coaches, technology staff, or additional staff to reduce class sizes.
- **Instructional materials.** These included textbooks, literacy and mathematics materials, and computers.

In addition, respondents reported that their schools were able to implement whole-school reform models, increase communication with parents by hiring a parent liaison, and provide teacher collaboration time.

In some cases, II/USP was reported to provide a sense of urgency and accountability. At several schools, identification as a low-performing school resulted in an increased awareness of the need to improve student outcomes. In a Cohort 1 elementary school, for example, the principal reported that II/USP pushed the school to reduce complacency among school staff and to raise the bar in terms of expectations for student learning. In the words of another principal, "Having gone through that process of having someone come in and look at the school from a different perspective opened up the eyes of people here." In two cases, the Action Plan was used to provide this sense of urgency and accountability. For example, at one Cohort 1 school, the principal, who came in during the first year of implementation, said, "I came in and brought the [Action Plan] in, told them to get it off the shelf… We are being held accountable to the state to implement this plan."

Why did II/USP facilitate growth more in some schools than in others?

To answer this second question, we examine the strategies and implementation processes associated with II/USP at *all* sampled schools (including those that made low growth in student achievement and state-monitored schools). While we found that II/USP reportedly did contribute to growth in a small subset of schools, other schools felt less of an impact from II/USP.

In terms of strategies, principals at low-growth and state-monitored schools reported using II/USP funds to implement similar strategies as the growth schools. These strategies primarily focused on instruction. For example, one elementary school principal who felt they had made significant progress by implementing monthly assessments and targeting their lowest-performing students thought that II/USP was very important to the progress they had made. He mentioned using the money to pay for additional teachers to reduce class sizes and to implement intervention classes.

However, fewer principals among schools that made low growth and state-monitored schools were able to articulate exactly what the II/USP funds had allowed them to do (approximately half of the principals versus 12 out of 16 principals at growth schools). For example, at one Cohort 2 middle school, the principal arrived two years ago, as II/USP was coming to an end. He reported that nothing had been accomplished or implemented in the previous three years using II/USP funds. This principal described II/USP as "pouring money on the problem" and as "educational welfare," and said it did not focus enough on identifying issues and pushing to address those issues. Another principal, who had started at the school several years ago, shortly after the Action Plan had been written, stated, "I inherited an Action Plan that was not well thought out. There appeared to have been no guidance whatsoever, the plan was not clear, it was discontinuous... There seemed to be nothing planned around instruction."

One possible explanation for these reported differences in reports is that there were somewhat more *new* principals at the low-growth and state-monitored schools than at the growth schools. While 69 percent of our administrator respondents at growth II/USP schools had served at their schools for 3 or more years, only 54 percent of administrator respondents at the low-growth or state-monitored schools had.

In examining data from all sampled schools, there appear to be several themes that emerge regarding where II/USP is most successful. The themes are as follows:

- *Buy-in of school staff to the need for change*. Respondents at three growth schools that rated II/USP highly in terms of its importance to their reforms mentioned a readiness to change among staff. This readiness reportedly resulted in more buy-in to the strategies associated with II/USP. For example, the principal at one school said, "One thing is that people knew for the most part that they weren't doing well, but did not know why or where." The External Evaluators allowed them to identify these needs, and helped teachers to work towards goals that helped them to improve instruction.

- *Strategies are clear, specified, and align with school needs.* The majority of growth schools implemented clear and specific strategies using the II/USP funds. These strategies often included those outlined at the beginning of the chapter as critical factors for growth, and were aligned with the needs of the schools. In some cases, this meant implementing a standards-based curriculum that had not yet been in place.

In other cases, when such a curriculum was already in place, it meant developing a system for assessing students and using data to ensure that they are implementing the curriculum in the best way possible. This alignment with needs helped to create buy-in around the goals and focus on the strategies.

- *Initial focus placed on, and ongoing revisiting of, Action Plan.* At two growth schools, principals explicitly talked about using the Action Plan as a continuous guide for improvement. The administrator at one of these schools, who started at her school after the Action Plan had been written, referred to it as "our bible" and discussed how the school became unified and organized around the plan. They systematically went through the plan and "tackled" what they could, which included building a mission and vision statement that they posted throughout the school, updating the school's facilities, purchasing instructional materials, and implementing additional professional development programs. At another school, all staff reportedly met every year to revisit the goals laid out in their Action Plan. This ensured that they were focused and that their goals continued to evolve to meet their needs.

- *Leadership provides urgent and focused climate for change.* In these cases, the principals emphasized the accountability associated with II/USP and the need for improvement. Several principals specifically mentioned "using" the state accountability system as a reason for necessary change. One principal said, "II/USP gave me a tool of urgency: I could say 'we have to improve because the state of California says we have to.'" Others had framed this accountability more positively. In the words of one principal, "I told everybody that if we are willing to take money, then we have to make changes that are coming about. We will take the money and do what we need to do with it to make this a better school." Alternatively, at one school that did not experience growth during II/USP, the respondent mentioned that there was a lack of monitoring and accountability: "There wasn't a lot of buy-in from the teachers and parents, there was no monitoring….there were no consequences."

CSR Schools

From our data, we wanted to see if schools participating in II/USP through the Comprehensive School Reform (CSR) program had any different experiences or facilitators to growth than the non-CSR II/USP schools. Of the 40 schools in our sample, only 5 participated in the CSR component of the II/USP program. We were able to interview four of the five CSR schools in our sample. Given this small number of CSR schools, it is difficult to make any firm conclusions from our data. Respondents reported varying experiences within the CSR program. At one school, one model was implemented during the first year of participation in II/USP. However, since the curriculum associated with the model did not align with the California standards, the model was dropped in the second year. At one urban middle school, which is currently a SAIT school, the teacher felt the External Evaluator contributed substantially to the school. However, the principal felt differently. He believed participation in SAIT had benefited the school more than CSR. Respondents at the other two schools, however, felt that participation in CSR improved the school. These two schools, from the same school district, participated in the same reform model. Respondents at both schools felt the program strengthened instruction, promoted teacher collaboration, and provided useful professional training. In fact, one teacher felt that the model had been central to the school's success.

What aspects of II/USP could be improved to better facilitate growth?

To answer this third question, we consider the strengths and weaknesses of the policy as reported by respondents at the school and district levels. Many respondents spoke positively about the overall concept and focus behind II/USP. They believed giving the schools additional funding was helpful, and thought the planning concept was a good one. Six out of 14 district staff interviewed specifically mentioned funding and what the funding enabled schools to implement/purchase as a strength of II/USP. Five district respondents noted that the focus on improvement associated with II/USP was helpful. In the words of a district administrator, the process "forced schools to look at themselves."

Nonetheless, there were aspects of II/USP that respondents reported could have been improved to facilitate their school improvement efforts. Several of these aspects were also identified in the 2003 PSAA Evaluation Report, and include more timely and appropriate distribution of funds, additional guidance for planning and use of II/USP funds, greater communication with other participating schools, greater consistency and skills among External Evaluators, increased monitoring during implementation, and more positive incentives. We discuss each of these briefly below. Another important constraint that was mentioned by respondents was the temporal nature of the funds. This is discussed in greater detail in the *Sustainability of II/USP Reforms* section.

Timely Arrival of Funds

Three district respondents reported that late distribution of funds significantly hindered the implementation of II/USP within their districts, particularly among year-round schools. Schools are required to spend their II/USP funds within the year they are allocated, minus a small amount of carryover funds (15 percent). The three districts that mentioned the late distribution of funds discussed the limitations associated with these spending requirements. In one large urban district, the administrator reported that notification letters for funding were not received until October or November. Without money allocated during the summer, schools cannot implement professional development or hire new staff in a timely way. This was reported to be particularly problematic for year-round schools. This district provided funds up front to participating schools in July for several years, but were not able to do this for all years of II/USP. Another district administrator at a large urban district echoed this response. However, in her district, they were unable to provide up-front funds to schools in any year, and schools did not receive funds until sometimes as late as May. She reported that many schools in her district had to return funds to the state government because they were unable to spend them by the end of the year. These restrictions also affected how the funds were spent, since the uncertainty associated with the funds reportedly prevented school administrators from focusing the spending on a comprehensive, coherent set of strategies.

Guidance and Flexibility for Fund Use

Many respondents (ten principals and three district administrators) mentioned the use of funds as an area of improvement for II/USP. First, they reported that they would have liked to have the flexibility to use these funds for purposes other than the External Evaluators. Although several respondents discussed the benefit of the outside perspective the External Evaluators brought to their schools (see below), others thought they could have spent their money in more effective ways. In the words of one principal, "A lot of money was spent on the outside evaluation, and more should have gone to the students."

Secondly, respondents reported that providing guidance about how best to use the funds for sustainable, effective strategies would have been useful. They thought that having better guidelines for the use of the money would have prevented the implementation of strategies that were not effective, and would have helped schools determine how best to use the money for the biggest impact. Some thought there should have been restrictions on the money. For example, one principal, said, "They should have had some formula [for spending the money] …. Automatically, 20 percent off the top should have been for professional development." This principal criticized the External Evaluator's plan, asserting that once they left, they had nothing to show for their work. When asked about any aspects of II/USP that he would change, another said, "I would have made it less of a free-for-all. More specific structures to outline where things would go." Others did not go so far as to suggest placing requirements on the use of funds, but suggested that additional guidance in terms of what should go in the Action Plan and what strategies to implement would have been useful.

Improved Communication

Five principals noted the need to have greater communication among key stakeholders within II/USP, including better communication between the state, districts, and schools regarding the provisions of the policy, and better communication among schools participating in II/USP. In the words of one principal, "Schools should have had better access to individuals who had a full understanding of II/USP and what you could do and couldn't do."

In addition, there appeared to be minimal communication among schools participating in II/USP, particularly between those that were making strong growth and those that were not. One principal mentioned that a good addition to II/USP would be to set up a system of communication for schools to learn from each other about what the provisions of the policy are, and what strategies have worked. Another stated that she would "rather see schools come together, principals meeting and sharing strategies" than to bring outside people (e.g., External Evaluators) into the school.

Greater Consistency and Skills Among External Evaluators

At the time of this study, three to five years had passed since respondents had worked with an External Evaluator. Therefore, although the External Evaluator provision was a key aspect of the II/USP policy, not very many respondents provided details regarding the External Evaluators' work. Greater detail about the planning year and the External Evaluators' role can be found in the 2003 PSAA Evaluation Report.

Despite a generally low level of knowledge about the External Evaluators, many respondents still mentioned the mixed quality of the External Evaluators that worked with their schools. On the one hand, approximately half of the respondents with at least basic knowledge of II/USP (12 out of the 23 teachers and principals) specifically mentioned that the work with the External Evaluator was valuable.[11] For example, at one school the principal and teacher both spoke highly of their External Evaluator team, with the teacher reporting that it was the "crucial element" for them. She discussed how the External Evaluator worked closely with the teachers as a team to identify their areas of need and to develop a plan. They also helped them to work through trust issues among the staff. The principal agreed with this statement.

[11] Two of these respondents discussed a *second* External Evaluator the school hired as effective. They did not have a positive experience with the first one hired.

On the other hand, four principals and four teachers with at least basic knowledge of II/USP, as well as two district administrators, mentioned the External Evaluators when discussing ways in which the policy could be improved.[12] The primary themes associated with these responses were that the External Evaluator did not work intensively enough with their school, that the External Evaluator did not understand the needs of their school, and that the benefit gained from the work of the External Evaluator was not enough given the cost. These issues were particularly salient for teachers. For example, one teacher said, "[The state] needs to look at [the External Evaluators] much more closely. In some cases there are people on the list to be consultants who are not at all able to carry out that function." This teacher believed that many External Evaluators were selected based on connections to the schools, rather than on their track record with improving schools.

One teacher talked about the lack of information available on the effectiveness of the External Evaluators. This, along with the mixed impressions discussed above, suggest a greater need for accountability in regards to the external consultants' work, as well as the need to provide more information on the skills and experience of external consultants in programs like II/USP.

Greater Follow-Up Support and Monitoring

Several respondents (six principals, four teachers, and one district administrator) discussed the minimal level of follow up support and monitoring associated with II/USP. Two main issues discussed included the low level of support from the External Evaluator beyond the initial Action Plan development phase, and the lack of monitoring by the state.

In terms of follow-up support, one principal thought it would be a good idea for the External Evaluator to return toward the end of the implementation process to provide advice on the next steps. Others thought more frequent support would have been helpful. For example, one principal talked about how they did not get a "whole lot of follow-up from anyone" and "were left to ourselves to implement the plan." One teacher who had a positive experience with their External Evaluator (who also assisted the school during the first year of implementation) said, "We have this wonderful opportunity for two years to work with this group, and they helped us a lot. But after two years they are gone. There isn't even any support for them to come once a year and give us a new shot in the arm."

Several principals would have liked to see more monitoring and accountability during the implementation of the program. In the words of one principal who started midstream in the II/USP process at a school that is now state-monitored, "In II/USP the monitoring was not there…the staff didn't have much buy-in." He later said he would have liked to see "someone over the principal to make the principal and the teachers accountable."

More Positive Incentives

Finally, though not mentioned by a large portion of respondents, the need for more positive incentives was strongly asserted by a few. These respondents stated that the threat of sanctions was not the best way to implement such accountability, and that the state should consider more positive supports and incentives to this end. One principal of a school that is now participating in the SAIT process felt that programs like II/USP just remind staff of a

[12] An additional two principals and teacher that had minimal to no knowledge of II/USP spoke negatively of the External Evaluator position (though not specifically of the External Evaluator at their school).

school's inadequacy. He believed not enough support was given to teachers who enter very difficult situations, and that the perception of inadequacy only exacerbates the situation.

Sustainability of Growth and II/USP Reforms

The analyses of student outcome data discussed in the previous chapter revealed that any gains realized by II/USP schools relative to the comparison schools dissipated either during or after II/USP participation. These results raised questions about the sustainability of effects that may be observed from II/USP after the funding ends. We addressed this question in our interviews by asking respondents at growth schools whether they expected to continue meeting their growth targets, and why. Since the growth schools in our sample were selected on the basis of having made consistent and/or high growth for several years, our sample has proven they can sustain their growth one to two years after II/USP implementation. This should be taken into consideration in interpreting their responses.

The vast majority (16 out of 17) of principals at growth schools expected to continue meeting their API targets (see the following section for a discussion of AYP targets). Within this group, the primary reason provided for expecting continued ability to meet growth targets was increased capacity among their teachers. Specifically, they cited increased teacher knowledge of what and how to teach, increased knowledge of how to look at and use data to inform instruction, increased buy-in to new instructional programs and focus, and increased expectations among teachers. There was hesitation on behalf of one principal, who felt that they may have reached a plateau in their growth and may need to focus more on the achievement gap and subgroups to improve further.

The optimistic reports above suggest that some strategies implemented through II/USP could have lasting effects. For example, strategies that seemed to alter the overall operation and focus of the school might result in longer-term changes even without funding. These strategies include training teachers to focus on instruction tied to the standards, implementing systems for the use of data to inform instruction, and developing a climate of high expectations. In some cases, implementation of sustainable strategies was intentional. As one principal stated, "My whole thing was to build capacity, long after we don't have money and after we are done with the plan…We're doing things that are *systematically* changing our school."

Despite the notable optimism among these respondents with regard to their growth targets, there was some indication that key II/USP reforms had either ended, or would need to end, due to the temporal nature of the funds. Most schools in our sample had only recently finished II/USP funding, and many still had some carryover funds. One principal described it this way:

> *"Sustaining [the instructional coaches] is our challenge right now… I've thought about that and am looking into grants. The problem I see with the whole accountability system is that when you are not doing well, you get money, which is great so you can do things you could not do otherwise. But once you start to succeed, it's almost like you are punished since your money is taken away. It's ironic."*

Other schools also reported struggling to sustain reforms. At one, they had to release a bilingual reading coach and a technology person once their funds ended. And at another, where they primarily used the money for release time for teachers to examine data from benchmark assessments and plan instruction, they were not sure where they would find

money to sustain these efforts in the upcoming years. This principal said, when asked about what aspects of II/USP could be different to make it more effective, "I would like that schools that are making progress with the use of this funding be considered for sustained or reapplication of funding." Finally, at a Cohort 2 elementary school, the principal explained that the after-school program which was extremely successful has been cut in half since they lost the II/USP funds. They now have approximately 100 students on the waiting list.

These findings raise questions about what schools can do to sustain their work after a program like II/USP comes to an end. One potential vehicle for sustainability that was mentioned in several cases was the district. For example, one principal reported that their focus and the skills that had developed through II/USP had been enhanced by the focus and goals of the district. The district's implementation of Open Court and a common math program had helped to support reforms that II/USP had begun.

Tension and Overlap with Federal Accountability Provisions

We have already discussed many important factors that have contributed to and hindered the implementation and resultant growth of II/USP schools. However, we found that an important additional influencing factor is the implementation of the federal No Child Left Behind Act, introduced in 2001. This section addresses the coexistence of PSAA and NCLB. The tension and overlap between these two systems are critical to the likely future impact of programs like II/USP.

We asked all principals and district administrators how they were balancing state provisions with the federal NCLB accountability requirements. This section summarized what we learned from these discussions, and specifically addresses three issues: perceptions of API and AYP targets, the emphasis schools have placed on these targets, and the effects NCLB implementation has had on strategies implemented by schools. We begin by describing some of the programmatic differences between NCLB and PSAA.

Differences Between NCLB and PSAA

Outcome Measures and Targets

A central component of any educational accountability system is an outcome measure that enables the monitoring and comparison of achievement across schools. NCLB and PSAA utilize two different outcome measures; while NCLB uses "Adequate Yearly Progress", PSAA uses the Academic Performance Index (API), as described in Chapter 1.

The API ranges from 200 to 1000 and ranks schools based on their overall performance on several assessments, including the California Standards Test (CST), the California High School Exit Exam (CAHSEE), and the California Achievement Test (CAT/6). Under the PSAA, a statewide performance target of 800 has been set. The state has also defined schoolwide yearly growth targets and a comparable improvement (CI) target for numerically significant ethnic/racial and socioeconomically disadvantaged subgroups. The API schoolwide target is set at 5 percent of the difference of the school's API Base and 800, while the subgroup or CI target is set at 80 percent of the schoolwide target.

NCLB uses Adequate Yearly Progress (AYP). AYP encompasses several measures and targets, including participation rates, "percent proficient" on assessments including the CST and CAHSEE, API performance, and graduation rates for high schools. Specific targets are

set within each of these categories. All schools, districts, and numerically significant subpopulations are expected to meet "percent proficient" targets (also known as annual measurable objectives, or AMOs) for English language arts and mathematics. For example, in 2005, for an elementary school to meet AYP, they must have had a 95 percent participation rate on statewide assessments, 24.4 percent of students proficient on ELA assessments, 26.5 percent proficient on mathematics assessments, and at least 590 on the API (or one point of growth). To meet AYP, all numerically significant subgroups, including English Learners and students with disabilities, must also meet these expectations. The expected percentage of students at or above proficiency has increased, and will continue to increase, on a schedule laid out by the state and guided by the federal government.

In both accountability systems, specific targets are set based on their respective outcome measures to assess whether schools are meeting state and national standards. However, several important differences exist between the API and AYP targets. First, the API targets are based on *growth,* while AYP targets are primarily based on a fixed target in any given year—in other words, the percentage of students above a certain level of achievement. In addition, the API includes an incentive structure for schools to raise the achievement of their lowest-performing students, since in the API calculation for growth, additional weight is given to increasing the lowest performing students' scores. The AYP targets encourage schools to move students from "basic" or below to "proficient" in order to increase the percentage of students in or above this targeted band. Finally, AYP places more emphasis on the progress of student subgroups by incorporating English Learners and special education populations into the list of subgroups, and requiring each subgroup to meet the same performance standards as all other students. In summary, the API is more oriented towards schoolwide progress, while AYP is more targeted specifically at individual students below proficient.

It is important to note that it is possible for a school to meet one set of targets without meeting the other, in any given year. For example, a school could increase test scores enough to meet their API growth targets schoolwide and among significant ethnic/racial and socioeconomic subgroups, but not meet the percent proficient target for their EL students and therefore fail to make AYP. Alternatively, a school could meet all percent proficient targets (as well as the additional targets such as participation rate) to make AYP, yet not *increase* their scores enough in one year to meet their API growth targets.

One additional important difference between the two systems is that NCLB also holds districts accountable for meeting AYP targets. Districts must have 95 percent participation rates, as well as specified percentages of students meeting proficiency standards. The PSAA did not incorporate a district-level accountability system.

Consequences of Failure

In the PSAA, only schools participating in II/USP and HPSGP are subject to state interventions and sanctions if they do not make the targets for a minimum level of growth. Under II/USP, if a school does not make significant growth in their API in both of the implementation years, the school becomes state-monitored and is assigned a School Assistance and Intervention Team,[13] and receives additional funding through their district (see Chapter 5). A school remains state-monitored until it meets its API targets for two

[13] Schools can also receive this sanction if they made some growth during the implementation years (but did not meet their targets) and then made no or negative growth in a subsequent year.

consecutive years. Currently state policy does not clearly state what consequences schools that do not show growth through the SAIT process will face.

Under NCLB, any Title 1 school is subject to sanctions if it fails to make AYP for two consecutive years. At this point the school enters "Program Improvement" (PI). A school can remain in PI status for up to five years if it continuously fails to make AYP. Districts are responsible for providing technical assistance and corrective actions during the PI process. There are a broad range of sanction and intervention options that gradually increase in severity each year that a school does not make AYP. These interventions include notification to parents that the school is in PI status, giving them the option to change schools; replacement of school staff; and ultimately, school restructuring. A series of other instructional modifications and support enhancements are required along the way. A school does not exit PI until making AYP for two consecutive years.

Another important distinction between PSAA and NCLB is that under PSAA, the state (by way of the SAIT process) is responsible for intervening in underperforming schools. With NCLB, the district is the primary entity responsible for taking corrective action with underperforming schools.

Under NCLB, districts (or Local Education Agencies) can also be identified for "Program Improvement" by failing to make districtwide AYP for two years in a row.[14] PI districts are expected to create a plan for improvement and to implement that plan in the upcoming year. They face additional corrective actions in the third year.[15] Districts do not exit PI status until they have made AYP for two consecutive years.

Perceptions of API and AYP Targets

Many interview respondents reported that AYP targets are more challenging than the API. When asked specifically whether they thought their schools would meet their API and AYP targets in upcoming years, principals were more often confident about meeting their API targets than their AYP targets. Of the schools where these targets were discussed, 83 percent of principals expressed a high degree of confidence that they would meet their API growth target in the upcoming years, compared to only nine principals who confidently expressed they would meet their AYP targets.

Respondents cited several reasons why they were confident about meeting their API growth targets. Many of the schools interviewed had been consistent in meeting the API target, and principals were confident about sustained performance. In addition, familiarity with the API reportedly allowed teachers to know 'where and how' to teach to increase the API. Clear achievement standards under the API facilitated principals' setting growth targets and expectations for learning and instruction to meet that target.

Due to what they considered to be AYP's unrealistic proficiency targets, additional targets for participation rates, and additional focus on English learners and students with disabilities, respondents reported less confidence in meeting *AYP* targets:

[14] In addition, districts must fail to meet an API threshold for the socioeconomically disadvantaged subgroup to be identified for PI.

[15] Corrective actions could include the replacement of district staff, appointing a state trustee in place of the superintendent, among others.

- *Unrealistic proficiency expectations*: Six principals noted that both in the short term and long term, meeting the increasing AYP targets was an unrealistic expectation. In 2006 schools will be expected to have 24.4 percent and 26.5 percent of all students proficient in ELA and mathematics, respectively. The percentages will increase in upcoming years. In contrast, as a school's API score increases, their yearly API growth targets decrease.[16] One vice principal of an urban high school noted the difficulties they face meeting the increasing "percent proficient" targets. She stated that although performance expectations rise, the performance level of students entering high school does not change substantially from year to year. Another principal mentioned the targets were 'statistically impossible' to meet in the long term.

- *Multiple targets*: Under AYP, in addition to increasing the percentage of students above a level of achievement, schools must also meet set participation rates and high school graduation rates. The current criteria require that schools and districts test at least 95 percent of students schoolwide and for numerically significant subgroups. Two schools specifically mentioned that meeting both the AMO *and* participation rate threshold was a significant challenge.

- *Additional focus on ELs and students with disabilities*: As noted earlier, under NCLB, a school must meet the specified targets for all numerically significant subgroups, including ELs and students with disabilities. Until now, API has not included these two subgroups within their subgroup targets. Nine principals expressed concern about meeting NCLB's requirement to test these two groups, and others were discouraged, claiming AYP targets are unreasonable for both ELs and students with disabilities. One principal suggested that the tests be made available in the primary language of the EL.

Although districts were not asked directly about their opinions on the rigor of the state and federal accountability targets, there were a few salient responses. Four of the 14 district respondents referred to NCLB requirements as less realistic than the API growth targets. Major concerns arose specifically in relation to the AYP targets for ELs and students with disabilities in low-performing schools. In addition, four respondents preferred the "growth" model associated with the API since it gives recognition to the improvement of the lowest-performing schools. One district administrator reported that "there are a lot of schools who have made substantial growth on the API, but aren't making the AYP yet, particularly those schools with high EL populations." Six of the 14 district respondents also cited "conflicting pressure" between the API and AYP. Two of these respondents said that NCLB was creating greater pressure because of the Program Improvement process. Two others mentioned that it is "problematic" to have two different targets because it creates confusion, particularly when a school is identified as "succeeding" according to one set of targets but facing sanctions according to another set.

System Emphasis

We asked school and district respondents how they were balancing state accountability requirements with NCLB. We found substantial variation regarding school and district responses in relation to their focus on the two systems.

[16] As noted previously, the API growth target is set at 5 percent of the difference of the school's API Base and 800—California's target for all schools. As a school's API Base increases, the difference between the API Base and 800 decreases, and thus 5 percent of that difference also decreases.

A substantial portion (6 out of 14) of district respondents reported focusing more on meeting AYP targets. These six reported recommending to their schools that they focus on making AYP (instead of meeting API growth targets). At one district, many schools had already regularly met their API targets and had shifted their focus towards meeting the AYP targets. In another case, the district reasoned that by meeting AYP targets and closing achievement gaps, schools' API scores would increase. Two districts mentioned their desire to avoid the sanctions associated with Program Improvement as a motivator for focusing more on AYP.

On the other hand, five districts reported setting achievement goals independently of accountability targets and focusing on providing schools with a basic set of tools (professional development, supplemental services, benchmark assessments, pacing guides). These districts were optimistic that accountability targets for both NCLB and PSAA would be met. Only one district claimed to be focusing more on API, citing a "respect for API and its intent" and a "focus on growth" as reasons.

Of the 34 sampled schools (including SAIT schools), respondents at 11 schools explicitly mentioned that they were more focused on the federal accountability system and the AYP. In contrast, only one school administrator explicitly mentioned that they were more focused on state accountability and the API. Reasons provided for focusing on AYP were similar to those stated by district respondents, including perceived stricter sanctions associated with failing to make AYP and the greater challenges associated with making AYP. At one school, for example, the principal believed that by maintaining focus on NCLB, the state requirements would be met. Several principals mentioned that federal requirements associated with the credentialing of teachers have pushed them to place a greater focus on NCLB. Two of these schools were having staffing shortages because of the difficulty in finding fully qualified teachers.

Some schools were relatively unconcerned about the multiple targets set by the NCLB and PSAA accountability systems, feeling that they could address both targets if fundamental issues such as standards were addressed. These schools reported focusing on comprehensive content coverage, addressing the state standards, setting goals, and focusing on testing skills. However, respondents at several schools were clearly discouraged about managing the requirements of both accountability systems. One principal described this as getting "pulled from both ends" and felt it was a struggle to find a "clear strategy" to address this conflict. Another felt he was "trapped between two accountability systems."

NCLB's Effect on Strategies

Given the focus that some schools and districts have reported placing on AYP, it is important to examine the extent to which schools and districts reported making instructional decisions that relate differently to these programs. With our district respondents, we have already seen that six gave direct instructions to their schools to focus on AYP. Another recommendation given to schools was to focus on the California Standards Tests (CSTs) and not on the norm-referenced CAT/6 (which is included in API but not in AYP). Another district motivated by its desire to exit PI status reported assisting principals in identifying basic students and targeting three to four of them to reach proficient status.

Among the school-level interviews, there were a few salient strategies designed with considerations specific to NCLB. Three principals described targeting strategies that identify students who are below proficient and focus resources on these students. For example, the principal of a rural elementary school reported knowing that the school "only needs 5

students per class to be proficient for AYP" and works with these lower-performing students for one hour, three days a week. A principal of an urban elementary school reported training teachers on how to use data to identify the performance levels of students to target basic-level students who are closest to the next performance band. A case that exemplifies how schools are making conscientious decisions based on coexisting accountability systems is one elementary school in the Central Valley, where instruction is shaped largely in response to both AYP and API targets. Here, half of the year is dedicated to focusing on the API: "resource staff works with below-basic and far-below-basic students." The second half of the year, the school identifies "students on the cusp" and targets those students for AYP by providing them with after-school or other supplemental interventions.

Summary and Conclusions

In this section we summarize our overall findings and conclusions from our examination of II/USP and factors for growth. Below, we briefly summarize our primary conclusions along three main areas: overall factors that contributed to or hindered growth among II/USP schools; II/USP's contribution to school growth; and the overlap of federal and state accountability provisions. In Chapter 6, we provide overall recommendations and policy implications for ongoing state accountability initiatives based on these findings.

Factors Contributing to and Hindering Growth in II/USP Schools

1. *School and Staff Capacity:* Strong leadership and professional learning opportunities, including the use of instructional coaches and opportunities for teacher collaboration, were noted as key factors among our growth schools. In particular, distributed leadership was noted to promote buy-in to reforms among teachers. The use of instructional coaches provided one-on-one mentoring opportunities and support, and high levels of teacher collaboration (i.e., a strong professional community) enabled greater instructional coherence and use of data to inform instruction.

2. *Coherent Instructional Program:* As was discussed in the 2003 PSAA Evaluation Report, we found that the development of a coherent instructional program contributed to growth. Two inter-related strategies were discussed by respondents: the establishment of an active vision and shared goals; and the implementation of common, standards-based curricula and instruction. The increased coherence was reported to enable greater teacher collaboration around instruction and to focus improvement efforts at the school.

3. *Systematic Assessment and Data-based Decision Making:* Assessing students on a regular basis using curriculum-related or school/district developed benchmark assessments was reported as a key contributor to growth. Growth schools spoke explicitly of how reviewing the data (typically within grade levels), identifying areas of need, and adjusting instruction accordingly provided a focused and informed learning environment. The district was often reported to play a key role in establishing systems of support for data analysis.

4. The most common challenges to growth reported by respondents were meeting the needs of their student population, establishing a skilled and invested staff, increasing parent involvement, and funding. Schools reported a variety of strategies they were implementing, with varying success, to address these challenges.

5. District policies and supports were sometimes reported to have a strong positive influence on the school's improvement efforts. Key policies and supports established by districts included districtwide common curricula, districtwide benchmark assessments, professional development opportunities, and technical assistance (e.g., for the interpretation of data). The implementation of these supports and policies varied considerably among districts with some schools reporting very minimal support from their district offices.

II/USP's Influence on Growth

1. Many respondents reported a low level of knowledge about specific II/USP strategies employed at their schools. Among principals, this limited knowledge was typically due to turnover in administration. This situation limited our ability to assess II/USP's direct contribution to the factors listed for growth. The limited knowledge also raises questions about the long-term impact of II/USP among participating schools with high levels of turnover.

2. A substantial subset of growth school respondents who *did* have a strong knowledge of II/USP spoke highly of the contribution II/USP had to their reform efforts. Most typically, respondents noted the funding and the focus that II/USP provided. A few attributed their strategies for growth to the planning year process. Strategies employed using II/USP funding primarily included professional development, additional instructional staff/coaches, and instructional materials.

3. Several factors appeared to contribute to the ability of *some* schools, and not others, to improve through II/USP. These included the level of buy-in among school staff to the need for change; the clarity, specificity, and alignment of strategies employed; the level of focus placed on and ongoing review of the Action Plan; and the level of urgency and focus fostered by school leadership.

4. Additional factors were reported to hinder the effectiveness of II/USP to successfully influence growth in schools. These included the late dispersal of funds, the limited level of guidance and flexibility for fund use, the minimal opportunities for schools to communicate with others in their situation; mixed experiences with External Evaluators; the limited level of follow up and monitoring; and the focus placed on negative labels and sanctions.

5. There was evidence that some schools were not able to sustain key strategies employed during II/USP once this funding ended. While some reported striving to implement capacity building strategies with long-term effects, others focused on shorter term needs such as instructional support staff.

Tension and Overlap with Federal Accountability Provisions

1. Respondents generally reported more positive perceptions of the API growth targets than of the AYP. Specifically, respondents spoke favorably of the API growth model as a means to recognize progress made among the lowest performing schools. AYP was reported by many to be unreasonable in its expectations for meeting percent proficient targets among all students, and among subgroups.

2. Despite more favorable perceptions of the API, respondents at both the district and school level reported focusing their efforts more on meeting the AYP. Several reported that by meeting the more challenging targets associated with AYP, they would likely meet their API targets as well. Others noted that more severe sanctions were associated with not making AYP. Some school staff reported taking specific actions towards meeting AYP by targeting "basic" students to move them into the "proficient" band. Others, however, reported that their focus on meeting standards and providing effective instruction would enable them to make progress towards both sets of targets.

3. Many school and district respondents reported that the overlapping and in some ways conflicting requirements of the two programs created a greater sense of confusion and overwhelmed school staff trying to focus on specific targets for growth.

Based on these findings we conclude that growth, where observed in II/USP schools, was mediated by a variety of factors, including district policies and school capacity. However, in a small subset of these cases, II/USP was reported to enable the successful implementation of some of the improvement strategies cited by respondents. However, several policy-related factors hindered the effectiveness of II/USP in all schools. Those that overcame these hindrances appeared ready for change and had leadership that promoted a sense of urgency in their efforts to implement the Action Plan and other focused strategies for improvement.

Chapter 5. School Assistance and Intervention Teams

Overview

One of the purposes of this continuation study was to study new aspects of II/USP, and in particular the School Assistance and Intervention Team (SAIT) process. In this chapter we discuss our primary findings regarding the implementation and effectiveness of the SAIT process. In the first section of this chapter, we provide a background summary of the SAIT structure and state monitoring process, focusing on the revised process that was implemented in the 2003-04 school year. The chapter's second section outlines telephone interview findings regarding the implementation and perceived effectiveness of this process. Here we also outline the perceived barriers to school improvement and strengths and weaknesses of the SAIT process. A summary of findings and conclusions are outlined in the final section of the chapter. Specific recommendations for the SAIT process are provided in Chapter 6.

Summary of Findings

Overall, we found that SAIT members and the staff at SAIT schools complied with the state-outlined process for monitoring. However, we also saw that the role of SAIT members varied among schools and organizations; in some cases SAITs conducted additional coaching and training in addition to their expected monitoring role. In addition, we found that even though SAITs were following the prescribed program with fidelity, there were mixed perceptions regarding the effectiveness of SAITs in the overall improvement of schools. While it appears that the SAITs implemented their prescribed school reform measures (the nine *essential program components*, or EPCs), there were mixed reports about the quality of assistance provided by the teams.

It was noted that while the interventions carried out by the SAITs may be *necessary* for schools to make progress, they may not be *sufficient* for creating meaningful and sustainable school reform. An *Academic Program Survey* (APS) is utilized by state-monitored schools and SAITs to measure the presence or absence of the EPCs. However, respondents reported that the SAIT process does not accurately assess the quality of implementation of the components, nor certain factors outside the scope of this survey such as leadership quality or school culture and climate, which tremendously impact a student's ability to learn.

The mixed opinions about SAIT effectiveness in facilitating school reform resulted in the examination of the role of SAITs in school improvement efforts beyond monitoring. There are valuable lessons regarding the implementation of monitoring programs and overall barriers to improvement in underperforming schools derived from the examination of SAIT in this light. These findings and lessons are outlined in the sections below.

The SAIT Process

Context

The SAIT process was first implemented in the 2002-03 school year. Schools that did not make positive growth on the schoolwide Academic Performance Index (API) in both of the two funded implementation years were required to contract with a SAIT.[1] These teams are expected to work closely with a school and its district over a three-year period to identify school needs and implement corrective actions to improve student achievement.[2] To exit the program, schools need to make significant growth for two consecutive years.

The initial SAIT process was revised in the 2003-04 school year to focus specifically on developing effective instructional programs in English language arts (ELA) and mathematics. To this end, the state identified nine essential program components that it deemed most critical to student success. The Academic Program Survey was created for schools and SAITs to assess the presence or absence of these components through interviewing school personnel and examining school data. All of the schools in our sample participated in this revised process.

SAIT Characteristics

The SAIT requirements call for teams to be composed of educators with experience in school reform and state curricula and standards. The teams include retired educators and other individuals from private companies, county offices of education, and nonprofit organizations.

To implement the SAIT process, the California Department of Education instituted a succession of approval processes and generated lists of approved SAIT organizations and SAIT leads. In 2002-03, the CDE approved 26 SAIT provider organizations. In July 2003, CDE staff developed and released a Request for Applications to solicit additional SAITs from county offices of education, accredited colleges and universities, and non-governmental educational organizations. They received 49 applications in which 21 county offices and 24 educational organizations were approved as SAIT providers. For the 2004-2005 school year, a total of 45 teams were approved.

To become an approved SAIT provider, organizations are required to submit a proposal and document that they have successfully worked with low-performing schools to plan and implement reform, and provided intensive support for schools to implement standards-aligned programs.[3] SAITs are chosen for their knowledge and skills in the following areas:[4]

[1] Additional schools have since been identified as SAIT schools. Schools that made significant growth during II/USP implementation but did not meet growth targets were identified for the SAIT process if they made *negative* growth in any year thereafter.

[2] Districts/Schools receive a combination of federal Title 1 funds and state general funds when participating in the SAIT process. Districts/Schools receive $75,000 (elementary and middle schools) or $100,000 (high schools) to hire a SAIT. Districts are expected to provide matching funds equivalent to half of this amount. Districts/schools then receive $125 per student for three years to implement the corrective actions. Districts are expected to match these funds 100%.

[3] See *Application Request for Providers of School Assistance and Intervention Teams* at http://www.cde.ca.gov/ta/lp/iu/arsaitltr.asp

[4] These are listed verbatim from the *Application Request for Providers of School Assistance and Intervention Teams* (see previous footnote).

1. State board-adopted academic content standards and frameworks;

2. The teaching of standards-based reading, writing, language arts, and mathematics for students by grade span;

3. Universal access materials and other strategies to help ELs acquire full academic proficiency in English and meet standards;

4. Student Testing and Reporting (STAR) Assessments, as well as curriculum embedded assessments, standardized, criterion-referenced, and other forms of assessment and their use to guide school planning;

5. Accelerated interventions for underperforming students and schools, including the state board-adopted reading intervention programs;

6. Professional development that addresses standards-based instruction focused on state board-adopted or aligned instructional materials that are in use at the school; and

7. Ability to provide the intensive support necessary for the school to successfully implement recommendations made by the SAIT.

All teams had to provide assurance that a representative number of team members had instructional or administrative experience in the grade span with which they will be working. After being approved by the state, these providers could advertise their services to schools in their region, county, or district.

SAIT Intervention Process

SAITs first conduct an audit of the school's instructional program by following a set of specific, well-defined steps. The initial step of this process is training school leaders on the use of the APS to evaluate to what extent a school is addressing the nine EPCs that were deemed critical to student success in reading/language arts and mathematics by the CDE. These areas (see Appendix B-4) differ slightly for kindergarten through eighth grade and for ninth through twelfth grade.

A school then conducts a self-assessment using the APS and reviews student achievement data. After completion, the SAIT visits the school to verify the accuracy of the survey responses through interviews with teachers and administrators and through a document review. Through this process, the SAIT assesses whether each EPC is minimally, partially, substantially, or fully in place. Finally, the team develops benchmarks for the three-year state monitoring period and compiles a Report of Findings and Corrective Actions for the CDE.

After this report is finalized, the SAIT assumes a role of support and monitoring. The process requires that the team provide "intensive support and expertise to implement the school reform initiatives in the plan" and is expected to monitor implementation of the corrective actions at least three times a year. The local governing board and the CDE have provided additional guidance on the support and monitoring provisions through SAIT training sessions and written documents (e.g., the RFP for SAITs). SAITs are expected to work with school and district staff to implement the nine EPCs, for example, by helping them to obtain standards-aligned textbooks, schedule professional development for teachers and the principal, and ensure effective implementation of teacher collaboration sessions. In terms of

monitoring, SAITs are expected to review data and progress towards the corrective action goals and benchmarks, and prepare a report of progress, including next steps. The report must be presented to the local governing board.

Findings from Telephone Interviews

As described in Chapter 2, we conducted interviews with the SAIT leaders along with the principal and one teacher at each of the ten state-monitored schools[5] included in our telephone interview sample. This allowed us to understand the SAIT process both from the perspective of SAIT representatives as well as from the school communities. Impressions regarding program implementation and the effectiveness of SAIT teams from the perspective of SAIT members, as well as principal and teacher respondents, are summarized below.

Implementation of the SAIT Process

Role of the SAIT

Overall, there was consistency among the reports of both the SAIT leaders and school staff regarding the scope of SAIT activities. When providers were asked to describe their roles, all respondents spoke of following state protocols with fidelity. One provider described her team's role as "following state protocols by the book" and many described the prescriptive nature of their roles. However, most providers (7 out of 9) described additional measures such as spending additional time at the schools to ensure the implementation of corrective actions, or providing additional coaching and/or consultative services to their schools. In at least two cases, these additional activities were conducted under separate, additional, contracts.

This variation in roles was also echoed by school staff. While all principals and teachers described SAITs that followed set protocols, some staff described very active teams, while others described minimum intervention. Overall, 4 of 15 respondents described their SAITs as performing duties above what was required, 9 described them as executing outlined requirements, and 2 described them as not performing the minimum required. For example, responses ranged from "Our SAIT person has been at almost every leadership meeting and has met weekly with the school and even a few times in the evening" to only "they monitor for compliance in the nine essential components." One teacher we interviewed was not at all aware of the intervention team at his school.

In sum, while in most cases the SAITs reportedly adhered to the basic requirements of the state monitoring process, some went above and beyond their articulated roles, and in at least two instances respondents felt they did less than what was required. This variation in roles indicates that participants in the SAIT process may be receiving substantially different amounts of assistance and support. Some of this variation may stem from the quality of services provided by the SAIT, and some may stem from variations in the agreements made between schools and SAITs when beginning the process. One SAIT leader explained, for example, that they told the schools ahead of time that they would provide more than the minimal amount of support and guidance through the process, and therefore fulfilled this agreement during the process.

[5] Nine SAIT members were interviewed, because one team provided services for two of the sampled schools.

Audit Process

According to both SAIT leaders and school staff, during the audit process the SAITs met with administrators to discuss the EPCs, interviewed teachers (in some cases only a subset of teachers, in other cases all teachers at the school), and reviewed assessment data and documentation of the nine essential program components (such as purchase orders for textbooks or school schedules).

During this audit process, the SAITs also reviewed the schools' assessment of the nine essential program components on the APS. In general, SAIT leaders reported that the schools had accurately assessed their implementation of the components. One, for example, called the school's self-assessment "brutally honest." In only a few cases, the SAIT leader reported having to adjust down the school's self-assessment of the APS ratings. In all cases, the APS revealed that most of the essential program components were *not* fully in place at the start of the SAIT process.

According to the interviewed SAIT leaders, the implementation of the audit process was generally smooth, and in close compliance with the process outlined by the state. One provider even reported that carrying on at one school after the originally assigned SAIT decided to withdraw resulted in relatively little disorder.

Respondent principals and teachers generally agreed that the audit process was implemented without incident. According to one principal, "[A] SAIT doesn't ask extraordinary things of schools, it just asks that they have these basic systems in place." However, there was variation in the degree to which SAITs involved teachers in the audit process. Some teams worked only with administrators, which caused some teachers to feel somewhat disconnected from the reform effort.

District Role

Districts are expected to play a key role in various aspects of the SAIT process. Specifically, they are expected to select the SAIT that will work with each school, distribute the SAIT funds to the schools, and work collaboratively with both the school and SAIT to implement the nine essential program components and improve instruction. The district also sets up a School/District Liaison Team (SDLT) that meets initially to discuss findings and corrective actions, and then periodically afterwards to discuss a school's progress and ongoing needs. Despite this high level of expected involvement, the actual involvement of the district in the SAIT process was reported to be variable by SAIT respondents.

While most (5 of the 9) SAIT respondents reported being very pleased with the district's level of involvement in their school reform efforts ("The district really puts its money where its mouth is—it supports all SAIT schools"), two of the nine were less pleased ("They seem to stand in the way more often than they help out"). The remaining two SAIT interviewees described the district role as neutral. Those who were pleased reported that district staff were responsive to their requests for support, took an active interest in supporting the school during the process, and communicated often with the SAIT. Those who expressed concern about the district's involvement discussed a lack of responsiveness from the district liaison, lack of cooperation in making changes to ease the implementation of SAIT requirements (e.g., busing schedules and assessment schedules), and delays in providing requested school-level supports (e.g., instructional coaches).

There was also variation in reports specifically regarding the role of the School/District Liaison Team. Two providers interviewed indicated that their districts (one large and one small) did not have such teams. In addition, those with teams reported a wide variation in the frequency of SDLT meetings, ranging from twice per year to monthly. The primary function of these teams was communication. Five respondents reported providing updates on school's progress to the district liaison during these meetings; three reported using these opportunities to gather answers to questions and updates on requests that the schools have made (e.g., for textbooks).

Monitoring and Support

After the audit process was complete, SAIT responsibilities fell within the category of monitoring and support. The level of monitoring and support within the SAIT process (as well as within other programs for low performing schools) is important to consider, given our finding from the 2003 PSAA Evaluation that the level of continued monitoring and support within the II/USP implementation years was insufficient to ensure effective implementation of strategies developed in the planning year. SAITs are required to submit a monitoring report three times per year, which involves an update on the implementation of the nine essential program components.

While some of the providers focused solely on the monitoring role (visiting at least three times per year), others (7 of the 9) reportedly provided additional supports in terms of coaching, professional development, and assistance with problem solving. For example, one provider indicated that they went beyond requirements specified by the state for the monitoring process and spoke of holding monthly meetings at the school in which all stakeholders reported on continuing activities. Another reported sitting in on department meetings, as they felt it would give them better insight into the functioning of the school. Others reported providing additional services, sometimes through additional contracts, such as trainings for teachers, coaching of administrators, and consulting services to develop master schedules. No respondents mentioned conducting classroom observations as part of their support and monitoring, since this was not allowed.[6] Over half of the SAIT respondents stated that their ability to follow up and monitor the successful implementation of the essential program components and to assess progress made by schools was limited by this restriction.

Interviews with principals at state-monitored schools generally confirmed reports by SAIT respondents about their involvement in schools after the initial audit process. However, teachers reported being much less connected with the process, and were sometimes unaware of the support and monitoring roles of the SAITs. For example, one SAIT member spoke of conducting regular visits to a particular school, while a teacher at the same school was unaware of any follow-up. Similarly, another team member reported regular SAIT visits while a department chair at this school expressed a lack of follow-up. A likely reason for this disparity is that SAITs reported most often meeting with the principal and other administrators during monitoring visits, rather than with teachers.

[6] One respondent did mention going on a district walkthrough to look at "artifacts" such as scoring rubrics and student work. However, this respondent emphasized that they were *not* observing instruction.

Effectiveness of the SAIT process

Since schools did not start implementing the revised SAIT process until the 2003-04 school year, it is difficult to draw concrete conclusions about the effect the SAIT process has had on student achievement. However, it is worth noting that of the 30 schools that entered the SAIT process in the 2003-04 school year, 70 percent met both their schoolwide API and comparable improvement targets for that year. Of the remaining schools, 13 percent met one of the growth targets, while 17 percent did not meet either growth target. While a substantial portion of schools undergoing the SAIT process met their growth targets in the first year, it is not clear to what extent the process contributed to this growth, as the intervention had only been in place for a part of the school year and in some cases for only a few months.

Perceived Contribution of SAIT Process to School Improvement

We did gather data, however, on participants' expectations for school growth as a result of the SAIT process, perceptions of the effectiveness of the EPCs, and barriers schools faced in improvement under SAIT. We should note that out of the 10 state-monitored schools in our sample, three had entered the process in the 2003-04 school year, and seven had entered in the 2004-05 school year. Thus, most schools in our sample had only completed one year of the SAIT process at the time of the interview (in spring 2005).

Both SAIT and school staff were asked to assess the overall level of progress at their schools since beginning the SAIT intervention. Generally, most SAIT leaders (7 of the 9) reported some level of progress as a result of the intervention. In addition, 63% of the SAIT respondents speculated that the schools they were monitoring would meet their growth targets for the year. The providers who felt their schools had made significant progress cited instructional coaching, strong leadership, teacher involvement in districtwide collaboration, a sense of urgency among stakeholders, better collaboration among teachers, and increased focus on instruction as reasons.

SAIT leaders were also asked to assess progress made on the EPCs. Out of the three SAIT schools that had participated for two years, one had reportedly put all nine essential program components in place. The others, along with all schools that entered the intervention in 2004-05, reportedly had some, but not all, EPCs fully in place. Although not addressed through the EPCs, a third of the respondents reported improvements in school climate and culture since intervention began.[7]

A difficulty cited by SAIT members in assessing their schools' overall progress was their inability to observe classrooms.[8] One provider said, "I think I've seen progress, but it's hard to know for sure without visiting classrooms." Another said that "what goes on in the classroom is critical," and therefore important to observe in providing support and monitoring progress.

In contrast to the effectiveness perceived by SAIT members, teachers and principals were split on the overall effectiveness of the SAIT process. Some principals and teachers (4 of the 14 responses) felt that the SAIT process was generally effective and produced positive

[7] Evidence cited included stronger organization, an increase in positive attitudes among staff, and more respect and collegiality demonstrated towards SAIT providers.

[8] Currently, SAIT members are not allowed to observe classrooms during the SAIT process. A plan is in place for a Level 2 SAIT process, for schools that have all EPCs in place, but have still not improved. Classroom observations will be allowed for this Level 2 process.

changes such as increased collaboration, overall encouragement, and organization to the school. An equal number classified the process as ineffective, with negative consequences such as intrusiveness, lack of substantial support, less time for electives and activities not related to ELA and mathematics, and the inability of SAITs to address salient issues. The remaining six responses by teachers and principals were mixed or neutral about overall effectiveness.

Perceived Effectiveness of the EPCs

According to SAIT providers, the list of EPCs contained many elements important to establishing effective mathematics and ELA instruction. As one provider said, "[The EPCs were] extremely helpful. They were very useful for schools that have not kept up with basic state curriculum efforts." Another said, "EPCs are the infrastructure to get a school moving [in the right direction]." One noted that they helped to focus the process on academic issues.

While all of the SAIT leaders interviewed agreed that the EPCs were central to their work, however, some felt the EPCs were treated more as a checklist rather than a full consideration of the quality of each component. They said that even if a school fully implemented all of the components, it might still not be successful. "Full compliance doesn't necessarily lead to a deep understanding of the program," said one team member, while another stated that "[EPCs] are necessary but not sufficient."

In addition, respondents reported a few specific concerns about the list of essential program components. First, six out of nine respondents commented that the APS should specifically address the instruction of students with special needs and/or English Learners. One provider called this omission a "tremendous shortfall" of the process. Although these subpopulations were reportedly addressed in the training of SAIT providers, SAIT respondents still felt they should be made integral to the list of EPCs. Secondly, five out of the nine SAIT respondents reported that school climate and culture issues, such as student and teacher absence and cultural sensitivity, should be included in the EPCs as well. According to one provider, "That is critical in terms of creating an environment for school reform." Finally, several providers and school/district respondents mentioned that the intervention courses required for high school students who are substantially behind grade level in reading and mathematics limited their enrollment in core course requirements and electives.

Salience of Sanctions

The basic underlying rationale of the II/USP and SAIT processes is to give schools the necessary resources, incentives, and tools to improve. As part of this, sanctions were introduced as a disincentive against failure. In the 2003 PSAA Evaluation, we found that the sanctions (which later became the SAIT process) associated with II/USP lacked salience among school staff. We found a similar limited saliency of sanctions for the SAIT process as well.

While some SAIT member respondents said the threat of sanctions motivated schools to cooperate with them, others reported that the sanctions had little, or a negative effect. For sanctions to be effective, participants in the process must be motivated to avoid them. However, SAIT member respondents noted that multiple threats of sanctions seemed to have decreased motivation. One SAIT member stated, "Further sanctions may give [schools] an excuse to 'throw in the towel'… My fear is they will lose the will to live." Echoing this, one teacher said that her school had become numb to the threat of sanctions. She said, referring to

Program Improvement, II/USP, and the SAIT process, "It was like getting three different shots. By the third you don't feel it."

The threat of sanctions also caused some unrest among teachers and administrators. At one elementary school that had recently been combined with a middle school there was reportedly a lot of "finger pointing" among grade levels about who was responsible for their becoming state monitored. At another, there was a fear of teacher turnover because staff would prefer to work in less restrictive environments.

At the time of this study, the sanctions for schools that do not improve within the SAIT process had not yet been finalized by the state. Thus, respondents did not know what sanctions might lie ahead. This further decreased the salience. Some respondents, for example, reported that the uncertainty caused school staff to feel like the outcome was out of their control. One teacher said, "There's a lot of fear and uncertainty. We don't know what next year will bring." In general, principals did not think it likely that severe sanctions, such as the state taking over their school, would result. One SAIT member said, "They've never heard of the state taking over a school."

Perceived Barriers to Improvement

While many SAIT respondents reported progress made by schools through the SAIT process, when asked what they saw as the biggest barriers to improvement, many cited factors outside the scope of the process. They included the following:

- *Ineffective Leadership:* Four of the nine SAIT respondents mentioned poor leadership as a barrier to school improvement. While the scope and quality of leadership is not fully addressed through the EPCs, providers cited a lack of urgency among administrators to implement reform efforts.

- *Lack of Investment/Motivation of School Personnel:* Another perceived barrier to improvement (discussed by three SAIT respondents) centered on the lack of investment or motivation by school personnel as indicated by high teacher absences, lack of communication between teachers, a culture of low expectations, and a lack of focus on academic achievement.

- *The Impact of Other Accountability Programs:* While SAIT providers did not say that NCLB has constrained their work, the general consensus was that schools and districts overwhelmed with sanctions were often pessimistic about the likelihood of success on any of the reform efforts. As discussed in Chapter 4, the requirements of the various accountability programs sometimes conflict, leaving schools and SAITs unsure as to how to best respond.

Perceived Strengths and Weaknesses of SAIT

Both SAIT members and school staff were asked what they believed were the major strengths and weaknesses of the process. These were sometimes directly related to the state monitoring process, while in other cases they were external.

- *Perceived Strengths of SAIT Process:* According to SAIT team members, the primary strength of the process was the increased support it provided to schools, including

additional money, technical support, and resources (e.g., textbooks). A few mentioned that the collaboration that develops between the SAIT team, the school, and the district was important. Finally, one provider said that the strength was the focus on underperforming schools, particularly in areas without a lot of community advocacy.

Teachers and principals noted similar strengths. School staff felt that money spent on staff development, intervention programs, and direct assistance to leadership were benefits of the program. They noted increased data tracking and teacher collaboration that could create long-term change. Principals also reported that the SAIT process increased their authority and credibility when implementing reform measures.

- *Perceived Weaknesses of the SAIT Process:* SAIT members reported that the process was lacking primarily in scope and accountability. Regarding scope, weaknesses included the failure to address the needs of student subgroups in need of special attention, the inability to observe classrooms, and the failure of the process to address school climate and culture since a difficult school climate can hinder the focus on instruction. Concerning accountability, weaknesses included lack of clarity regarding sanctions for schools that do not show improvement, absence of sanctions for poor teams or External Evaluators, and a lack of 'teeth,' as according to one member "there is no language in the law that says schools have to do anything."

  Teachers and principals highlighted the difficulty associated with implementing the SAIT provisions. In the short run, school staff reported that the process was too intrusive, entailed time-intensive paperwork, and was stressful and distracting. They also identified the process as focusing too much on the core subjects at the expense of electives and that it ultimately centered on threatening sanctions rather than providing services or support.

Summary and Conclusions

This section summarizes overall findings from our investigation of the SAIT process. Policy implications and recommendations for the SAIT process are outlined in Chapter 6. Below, we briefly summarize our primary conclusions:

1. All interviewees generally agreed that the state monitoring process was useful for instituting the basic infrastructure necessary for effective ELA and mathematics instructional programs. However, while the EPCs were necessary for struggling schools, they were often deemed insufficient for meaningful and sustainable change, as they did not address the quality of the intervention nor key barriers such as poor leadership.

2. SAIT providers noted several specific problems with the EPCs. First, the needs of traditionally underperforming cohorts (specifically ELs and special education students) were considered a critical omission from the EPCs and the SAIT process in general. In addition, some SAIT providers reported that the exclusion of a school climate/culture component was problematic in schools with a "toxic" school culture and low levels of motivation among students and teachers. Other providers and district/school respondents discussed the complications associated with the intervention requirements for secondary schools, reporting that students were often unable to complete core course requirements and enroll in electives.

3. While all providers in our sample were reported to adhere to legislation guidelines, the level of service provided by SAITs appeared to vary broadly. While a few provided extra services pro bono, others reported charging for added support, and some only conducted the minimal monitoring required by law.

4. Several SAIT providers reported that their ability to assess and monitor progress within the SAIT process was hindered by their lack of access to classrooms. They reported that their inability to observe instruction limited their ability to assess key issues that may hinder growth.

5. Despite clear expectations for a large district role within the SAIT legislation, participation of districts in the process was reported as variable. While some districts reportedly provided a high level of support and met regularly with the SAIT, others were reported to create barriers to change at the school.

6. The consequences of failing to make growth while state monitored were not clear to respondents. These unclear sanctions increased confusion and anxiety surrounding the reform efforts and in some cases reportedly decreased motivation.

At this early point, we are unable to assess the actual effectiveness of the SAIT process on improving student outcomes. While a substantial percentage of SAIT schools met their growth targets in the first year of participation (2003-04), we have minimal evidence at this point to confirm a link between these outcomes and the SAIT process.

According to respondents' reports, however, the SAIT process has helped schools make progress in establishing a basic infrastructure that fosters effective ELA and mathematics instruction. On the other hand, variable levels of support and some clear omissions of important components (i.e., the instruction of key subpopulations) were reported as hindering the effectiveness of the process. In addition, lack of consistency in the role of the SAIT and the district raise questions about the level of guidance and accountability provided by the state for these roles.

Chapter 6. Conclusions and Recommendations

Introduction

In the previous chapters we outlined primary findings in three areas: achievement outcomes for II/USP and comparison schools (Chapter 3), strategies and factors that have contributed to or hindered growth in student achievement in II/USP schools (Chapter 4), and the implementation and effectiveness of the School Assistance and Intervention Team (SAIT) process (Chapter 5).

In this chapter, we draw on our prior discussions to present central findings relevant to II/USP and consider their implications for future state policy. Acknowledging that II/USP has been replaced by the High Priority Schools Grant Program (HPSGP), we focus our recommendations on state-level accountability in general and on the SAIT process (which is expected to be used with HPSGP schools). We also address the difficulties associated with the simultaneous implementation of the accountability and assistance provisions of the federal No Child Left Behind (NCLB) Act.

We divide our discussion into three parts. In the first, we focus on our central cross-cutting conclusions concerning II/USP and factors influencing the progress of low-performing schools in California. In the second section, we provide recommendations and discuss policy implications at the state and local levels related to these overarching conclusions. In the final section we turn to the SAIT process, providing overall conclusions and recommendations for future investigation and implementation.

Overall Conclusions Regarding II/USP

In this section, we discuss our overall findings and conclusions regarding state-level accountability policy across four main areas: the overall impact of II/USP on student achievement, factors contributing to or hindering growth in student achievement within II/USP schools, the role of the district in school improvement efforts, and the impact of NCLB implementation on the focus and implementation of the state accountability program. In the next section we discuss recommendations that stem from these major findings.

Conclusion 1: **Overall, the impact of II/USP participation on student achievement has been negligible. Any small advantage experienced by II/USP schools relative to comparison schools during program participation dissipated before or soon after program completion.**

In Chapter 3, we discussed the overall student achievement trends in both II/USP and comparison schools as measured by the school-level API and by student-level SAT-9, CAT/6, and CST scores. We examined achievement trajectories to identify any differences between II/USP schools and similar schools that did not participate in II/USP.

As discussed in the 2003 PSAA Evaluation Report, one important trend is that statewide **low-performing schools (both II/USP and non-II/USP) have made large increases in STAR scores since the implementation of the PSAA**. Elementary schools made the largest gains, while middle and high schools made more moderate gains.[1] In addition, when examining student-level standardized scores, low-performing schools' scores have moved closer to the state average over the past six years.

However, **we find only modest differences in student outcomes between II/USP and comparison schools. In addition, these modest gains, where observed during the period of program participation, dissipate after program completion.** In some cohorts, in some grade levels, II/USP schools experienced slightly larger achievement gains in comparison to similar low-performing schools in the state during II/USP planning and implementation years. More often, however, there was no observed effect, and trends were inconsistent across cohorts and grade levels. Perhaps the most prevalent trend is that while we see some statistically significant but relatively small positive effects during II/USP planning and implementation, we find significant but relatively small negative effects of II/USP in the years after the program ends.[2] For Cohort 1 elementary schools, this negative effect occurs even earlier, during the implementation years.

In addition, when positive gains in relation to comparison schools were observed, they were on the order of 0.02 standard deviations, which is quite small when compared to the overall gains made by low-performing schools across the years. This modest level of improvement, even if sustained, can be considered fairly insignificant from an educational standpoint (Cohen, 1969), and when considering the amount of funds and effort invested in II/USP.

These findings call into question whether II/USP is an effective strategy for improving the state's low-performing schools. This does not mean, however, that nothing can be gained from these efforts. A major focus of this study was to learn lessons for future state policy, and these lessons are reflected in our recommendations below. We have learned a considerable amount through the implementation of II/USP about what works well and not as well within a statewide accountability system.

In the 2003 PSAA Evaluation Report, we provided several possible explanations for the minimal-to-nonexistent overall effect of II/USP during the early planning and implementation years. These explanations are still reasonable given the current policy context, and should still be considered when interpreting these findings. First, we proposed that the II/USP could have had both a *direct* effect on participating schools and an *indirect* effect on schools that did not participate. This indirect effect could have resulted from the attention to student outcomes that resulted from the implementation of the PSAA overall, and II/USP in particular. For example, some districts responded to PSAA by developing policies for low-performing schools that addressed more than just II/USP schools.

A second possible explanation can be attributed to the wide variation in growth among II/USP schools. Given that some made considerably more growth than others, we found that

[1] For example, Cohort 1 II/USP and comparison elementary schools gained approximately 190 API points between 1998 and 2004, while Cohort 1 II/USP and comparison middle schools gained an average of approximately 107 points since 1998. Cohort 1 II/USP and comparison high schools increased by an average of approximately 79 API points in this same time period.

[2] After the first two years of implementation the program ends at schools. However, many schools did receive an additional year of funding.

growth was heavily influenced by factors outside of the program, such as district policies and school capacity.

Now that additional time has passed, another explanation that we believe the state must strongly consider is that the accountability theory of action model shown in Chapter 1 of this report may have broken down as implemented through II/USP, especially given current contextual factors. For example, the ability to garner attention and create a focus around a clearly identified set of measurable objectives for II/USP schools has reportedly been at least partially compromised by the somewhat conflicting goals of NCLB. Also, the enhanced motivation to improve that was expected to result from identification, labeling, and consequences reportedly did not occur in all schools. Some schools reported that they saw II/USP as a wake-up call, while others found it de-motivating to be labeled as a failing school—especially when they felt they were being held to unrealistic expectations (more prevalent under NCLB, but sometimes cited as a factor under II/USP), receiving inadequate guidance, and faced with unclear sanctions. Regarding motivation, some also noted the seeming irony of receiving supplemental funds while failing, then subsequently having the funds removed upon success.

The theory of action for II/USP assumes that increased capacity through external assistance and the receipt of supplemental resources is sufficient to result in sustained change. Respondents, however, sometimes expressed concern in regard to the base resources on hand. Examples of these concerns involved the availability of highly qualified staff, the mixed quality and usefulness of the II/USP External Evaluators, the temporary nature of II/USP funds (short term and removed upon success), and the late distribution of these funds. There may be a natural tendency for school staff to attempt to transfer out of schools labeled as failing, especially if it appears that the elements perceived to be needed for success are lacking. State data do not make it easy to track this kind of movement, but we believe it may create a substantial hindrance to long-term success in underperforming schools.

Given the focus in this study on the *longer-term* effects of II/USP, particular attention must be paid to the decline in student achievement growth in II/USP schools, in relation to the non-II/USP comparison group, that is found in later years (post-implementation). We propose two possible explanations for this. First, the change in relative growth may result from increased growth among comparison schools, as policies such as NCLB or HPSGP[3] place pressure on other low-performing schools in California. Secondly, the change in relative growth could be a result of a decline in II/USP schools' growth after the program ends. We discuss both explanations below.

NCLB has placed increased pressure on *all* low-performing schools. Schools that did not participate in II/USP are currently being pushed to improve student outcomes and meet high expectations for student performance mandated under NCLB. The pressure felt by these schools, particularly those identified for Program Improvement, could be leading to greater attention to student outcomes and a resultant increase in scores in comparison to the II/USP schools. In the 2003 PSAA Evaluation Report we proposed possible reasons for a "planning year bump" among II/USP schools, including the increased attention to student outcomes and instruction resulting from the labeling of their school as "underperforming," public scrutiny, and the planning year process. These same factors may now be contributing to growth in non-II/USP schools identified for improvement or undergoing early sanctions associated with NCLB.

[3] However, only 22 percent of the comparison schools participated in HPSGP.

Additional concerns must be raised regarding the sustainability of interventions like II/USP. Such programs must be examined closely in considering the benefits and drawbacks of a program where schools receive an infusion of funds for a limited number of years. Our interview data suggest that despite optimistic expectations of long-term growth, many schools had to reduce or eliminate programs implemented using II/USP funds. In particular, schools that hired additional staff such as instructional coaches had to either eliminate the position or reduce the position to part time. A decline in achievement growth could be an anticipated result of such program reductions.

Despite the limited effect that II/USP has had on schools, we see *some* preliminary evidence of a positive effect of HPSGP participation within our achievement analyses (among schools that participated in both programs). Though these results should be interpreted with caution since the full effect of HPSGP has not been explored in this study, there may be evidence that some of the changes incorporated into this successor program may have improved the effectiveness of this model for school improvement. These changes include a greater focus on the *lowest*-performing schools (i.e., schools in decile 1), increased funding (double the amount received by II/USP schools), and the expectation of greater district involvement.

Conclusion 2: **Despite the lack of an overall program effect, there is evidence that II/USP participation contributed to growth in <u>some</u> schools. Respondents in these schools identified specific factors and strategies they believe led to their improvement.**

The bad news in the state's efforts to turn around low-performing schools through II/USP is described above. The good news is that some previously struggling schools were able to make substantial progress during the time of II/USP implementation, and some attributed this growth to II/USP participation. In the 2003 PSAA Evaluation, we found substantial variation in growth among participating schools. Thus, in conceptualizing this continuation study, it seemed that the most informative next step would be to identify schools that did make high growth, attempt to identify what worked for them, and consider what could be replicated in future state efforts. Consequently, we identified subsets of schools that had made relatively high or low growth during II/USP. Interviewing staff at these schools allowed us to learn about the factors and strategies that contributed to their growth, as well as the potential impact of II/USP.

Several essential factors for growth in student achievement were identified by schools that made consistent and/or high growth in student achievement during II/USP. These included the following:

- *Capacity:* Leadership, professional community, and professional learning supports, such as coaches, were reported as important. Often, leadership was distributed beyond just the principal, to include coaches, department heads, and teacher leaders.

- *Instructional Coherence:* Various aspects of instructional coherence, including a common curriculum, curriculum tied to the standards, or an overall instructional vision or focus for the school, were among the factors cited as contributing to growth.

- *Systematic Assessment and Data-Based Decision-Making:* Systematic and regular assessment of students was reported to provide necessary and important information to inform instruction. Attention to, and monitoring of, student learning using regular benchmark data were discussed as important ways to identify needs of students and to inform instruction.

Several aspects of II/USP reportedly influenced the growth of schools within the program and the ability of schools to implement these factors for growth. Some growth school respondents spoke highly of the contribution of II/USP, citing the funding and the focus provided through planning as most important. However, many others did not experience high growth, nor had positive experiences with the program.

Growth attributed to II/USP appeared to vary based on several characteristics, including the buy-in of school staff to the need for change, the clarity and specificity of strategies and the alignment of those strategies with the needs of the school, the focus placed in the long term on implementing and revisiting the Action Plan, and the extent to which leadership created an urgent and focused climate for change. These qualities appeared to be less present in schools that did not meet with success in improving student outcomes during II/USP.

Additional factors reported to hinder the effectiveness of II/USP included late distribution of funds, limited guidance on how to best use the funds, limited communication among II/USP schools and between the schools and the state, mixed experiences with External Evaluators, limited follow-up support and monitoring, and the focus on negative labeling and sanctions.

Of particular note is that while the subset of growth schools identified for this study generally felt that they were successful in improving student learning, it was clear from our interviews that **few schools were communicating with each other to share these effective improvement strategies**. Several respondents mentioned that they would have liked to have more opportunities for communication with other schools participating in II/USP to learn more about the program and what was working in other schools. Given the mixed quality and experience of the External Evaluators, one respondent emphasized that it would have been more useful to learn from other schools in their situation than to learn from an external entity.

Conclusion 3: **Local districts were found to influence the achievement trends in low-performing schools. While district supports were reported as key to some schools' improvement, these supports did not appear to be present in all schools.**

As in the 2003 PSAA evaluation, we again observed a significant district effect on the achievement trajectories of low-performing schools in the state. That is, while II/USP and comparison schools performed in similar ways *within* districts, the pool of low-performing schools varied considerably in performance *across* districts (controlling for characteristics of the student population and other factors). These findings, coupled with our interview data, suggest that districts serve as intermediaries between state-level policy and school-level implementation and that districts have the ability to affect the overall growth and performance of their low-performing schools.

From our interview data, we have seen districts can institute policies and supports designed to improve the work of all their low-performing schools, irrespective of program participation. These include technical assistance and professional development, particularly around systematic assessment and data use, as well as the targeting of resources to low-performing schools. In addition, they can provide a focus for schools' improvement efforts. However, such supports and focus did not appear to be present in all districts or realized by all schools. Respondents from many schools in our sample reported additional areas of support they would like from their districts.

Conclusion 4: **While possibly intensifying the focus on accountability overall, the presence of differing state and federal accountability systems has diffused the attention schools are paying to PSAA.**

The implementation of NCLB has created an additional layer of achievement targets and expectations for schools in California. Schools are expected to meet both API and AYP targets, and face sanctions from both the state *and* federal governments if they fail to do so. NCLB, by focusing efforts on all schools, has possibly raised the focus on accountability among schools *not* participating in II/USP.[4]

Some respondents did not report the additional layer of accountability as a major distraction, saying that their plans to improve student achievement overall should meet the needs of both programs. Many others, however, at both the school and district levels reported that having two separate accountability systems in place has "confused" and/or "overwhelmed" school and district staff. Keeping track of these two sets of targets and thinking strategically about how to meet them were considered difficult. In addition, respondents referred to the conflicts between the two policies. While the API provides incentives for schools to focus on their lowest-performing students, the AYP provides incentives for schools to move students who are just below proficient into the proficient band. AYP also places particular focus on EL and special education students, while this emphasis has not yet been added to the API.

Respondents in this study generally reported that the AYP targets associated with NCLB were more challenging to meet than the API targets. In part, this may be due to the fact that they are all relatively low-performing schools (within the bottom five deciles). Thus, while they may make growth in any year, they may still struggle to meet the absolute proficiency targets that are set for all schools in California. In addition, respondents spoke specifically of the difficulty they face meeting the subpopulation targets associated with NCLB. Of particular concern were the targets for English learner and special education populations.

Similarly, respondents reported that the focus on *growth* in the state system was more appropriate to the improvement goals of low-performing schools than the focus on meeting a set proficiency target. The API growth targets allow for recognition of growth and progress among the lowest-performing schools that may still be far from meeting proficiency standards.

Finally, respondents reported an increased level of attention toward the AYP targets. In part, this attention reportedly stemmed from the fact that these targets are considered more challenging to meet. Some thought that by making AYP, they would likely meet their API targets as well. However, respondents also said that they considered the sanctions associated with AYP to be more severe, and that this was a reason for greater focus on this measure. In addition, some districts were said to have told their schools to focus more on making AYP.

Recommendations for Ongoing Accountability Efforts

As II/USP has essentially come to an end, the recommendations presented in this report are not specifically directed at this program, but rather more generally at the state's ongoing efforts to improve low-performing schools. This study, as well as our prior evaluation of II/USP, has shown no sustainable aggregate II/USP effect on student outcomes. This seems to

[4] See our discussion of this possibility within the first conclusion discussed.

suggest that more comprehensive and dramatic changes in state policy will be needed to yield long-term sustainable improvement in the state's underperforming schools.

The improvement of low-performing schools has long been a difficult and intractable process. The realization of substantial progress in these schools on an annual and continuous basis is exceedingly challenging. In addressing this challenge, we must first consider the overall context for this effort. Accounting for variations in the cost of education, California schools are among the lowest funded in the nation. At the same time, we have adopted some of the highest academic standards. The dissonance between investment and desired outcomes may be one contributing factor to the expectation that more and more California schools will be designated as Program Improvement schools when NCLB standards are further raised in future years (e.g., in 2007-08). As described by Mike Kirst (2005), Stanford professor of education policy and former California State School Board President,

> *[California's] outcome oriented accountability system has never been aligned with the school finance system that relies on specified inputs, processes, and categorical programs that control about one-third of local budgets…The finance system is exceedingly complex … and provides neither equity or adequacy. It is not based on the necessary funds to enable all pupils to meet the state's high academic standards.*

No schools are more subject to the tensions resulting from this state-level dissonance between educational resources and expected outcomes than those with high percentages of students in poverty. The correlation between high poverty and II/USP participation is clear, with an average state poverty rate of 50 percent, compared to an 80 percent poverty rate for II/USP schools. Poverty has clearly been established as the strongest correlate with student performance, placing high poverty schools at a substantial disadvantage in relation to other schools—especially when the goal is specified as equally high expectations for all.

In light of this, an important question for the state to consider is the base resources available to these most challenged schools in relation to other schools in the state. Using current state data, it is not easy to track resources at the school level. However, one statistic of concern from a recent report to the state legislature in regard to the implementation of Proposition 227 is that the state's highest EL schools (i.e., those with total enrollment that is 61 percent or more EL) have a lower percentage of fully credentialed staff (86.9 percent) than the state average (92.5 percent) (Parrish et al., 2005). This is of particular concern given the dual education challenges facing ELs (acquiring English while simultaneously meeting high academic standards) and the fact that EL concentration is highly correlated with poverty. If the state's most challenged schools are not assigned some of its best teachers and most proven administrators, and incentives are not provided for these staff to remain in these difficult environments, it seems unrealistic to expect short-term, relatively small, monetary infusions like II/USP to make a difference.

Going back to some of the most basic principles of public education funding, as well as vast subsequent research, it has been clearly established that some classifications of students cost more to educate than others.[5] Differences in student need have also been acknowledged in adequacy court cases throughout the country and through current education funding systems in virtually all states (Chambers et al., 2004). This principle is also reflected in federal categorical aid programs for students in special education and poverty. California awards

[5] For example, see Berne and Stiefel, 1994; Peternick, Smerdon, Fowler, and Monk, 1997; Clune, 1994; and Reschovsky and Iimazeki, 1997.

categorical funding to school districts based on their number of English learners, students in special education, and students in poverty. However, the exact extent to which these resources trickle down to schools is unknown and current state data make this difficult to track.

In short, it would seem that a vital precondition for long-term improvement in the state's lowest-performing schools is that they have a base of resources that exceeds those allocated to schools facing fewer challenges. If the lowest-performing schools in the state have higher percentages of uncertified teachers, less experienced principals, and lower levels of overall funding in comparison to higher performing schools, any state plan for sustained improvement in the lowest cohorts of schools seems destined to fail.

At the same time, high expectations for all districts, schools, and students clearly set the state on the right path. Some schools and districts seem to respond very positively to this substantial and relatively recent redirection in attention to the state's most challenged schools and school districts. A primary recommendation that we will further elaborate is that we must identify and acknowledge these districts and schools that have shown substantial success despite such formidable hurdles as high levels of student poverty—at least to the extent that we emphasize those that continue to fail. These districts and schools should be acknowledged because they have clearly earned this credit, and even more importantly because if anyone in the state knows how to turn around low-performing, high poverty schools, it is the leadership and instructional staff in schools that have realized this substantial accomplishment.

The other good news, further substantiated by the current study, which has examined and attempted to contrast relative success and failure in high poverty, low achieving schools, is that the general ingredients for success in these schools are fairly well known. As cited in Chapter 2, there is substantial literature in regard to school success that repeatedly points to such factors as ongoing monitoring of student progress, strong instructional leadership, shared goals, and a positive and academically focused school climate.

These factors are also very well aligned with what we found in the current study. The real question is exactly what these elements mean in practice. In other words, what are some of the specific strategies schools use under these general headings, and how can they be replicated over a broad range of low-performing schools? The recommendations below try to focus on the kinds of systemic changes and specific strategies that we believe will be needed for accountability interventions like II/USP to succeed in the future.

Recommendation 1: **The state should consider the resources needed for sustained academic success in low-performing schools and ensure that they are present and sustained in applicant schools and their districts.**

The state should specify ongoing resource standards to address the needs of the state's most challenging, highest-poverty schools, and ensure they are allocated effectively by districts to schools, as a precondition for programs like II/USP. Given that schools reported being unable to sustain key strategies implemented through II/USP, sufficient resources to *sustain* improvement efforts must be an important component of any program that provides focused resources for change. Since the district was found to be a key intermediary between state-level policy and school-level implementation, the state must ensure that districts have the resources to provide the necessary assistance and support to their schools, and that they allocate them to low-performing schools as needed. For example, they might be required to

ensure a teaching force in these schools that at least equals, or exceeds, the district average on such characteristics as experience and degree level.

At this point in time, several foundations are about to contract for an adequacy study for the state, supported by the Governor and leadership in the California Department of Education and bi-partisan leadership in the California State Legislature. This study will systematically consider the funding needed for public education given the state's specified academic goals and the composition of pupil needs in districts across the state including students living in poverty, English learners, and students with disabilities. This type of study should prove useful in determining these ongoing resource standards.

Recommendation 2: **The full power and potential of districts and other intermediate agencies should be brought to bear on the problem of shoring up low-performing schools. Districts, and other relevant agencies, should be expected to play a key role in providing support, vision, and guidance for school-level improvement, and should be held accountable by the state when this does not occur.**

Given the high number of underperforming schools in II/USP and in California overall, the state will need all the help it can get in providing the intensive support and monitoring required for schools undergoing improvement efforts. This study demonstrates (and common sense clearly suggests) that districts can make a substantial impact on improving low-performing schools when they focus their attention, energy, and resources on this effort. Districts are clearly an important vehicle to ensure proper and effective implementation of programs such as II/USP, and to ensure the growth of the lowest-performing schools in their jurisdiction.

Within such a system, districts should be held accountable for providing additional focused assistance for the schools most in need, and to provide long-term and sustained support to maintain progress made through programs like II/USP.

The following should be considered key areas in which districts should play a primary role:

- *Fostering coherent instruction within schools*. Districts can encourage coherence by ensuring that schools have opportunities for teachers to collaborate and by monitoring the development and implementation of a vision and set of coherent goals for each school.

- *Encouraging the use of data to inform instruction.* Districts can provide schools with benchmark assessments that provide frequent and useful data to identify students' needs and inform instructional strategies and plans. Districts can also provide the means to analyze these data through systems that provide user-friendly and valid outputs of data on student outcomes, disaggregated in useful and effective ways.

- *Providing sufficient resources for an effective learning environment.* Districts can provide training and incentive structures to ensure that the lowest-performing schools have strong, highly-qualified staff and skilled and effective leaders.

- *Providing instructional supports and technical assistance to schools.* Districts can provide necessary instructional supports, including instructional coaches and curriculum specialists, to provide individualized support to schools undertaking improvement efforts. In doing so, the state should consider the resources available

through counties and intermediaries. These agencies can provide supplemental supports and technical assistance that some districts, particularly small districts, are unable to provide.

As mentioned above, the district can also play a key role by ensuring that funds are allocated in a way that supports their lowest-performing schools. One urban district in the study sample, for example, allocates dollars to schools on a per pupil basis, rather than in full time equivalents (FTEs). Therefore, schools with less experienced and therefore less expensive staff have additional resources available to spend on instructional resources, or on the hiring of more experienced staff. Such student-based funding systems within districts generally feature weights ensuring additional district resources for schools enrolling high percentages of students with special needs, such as students in poverty, ELs, and special education students. This is just one example of ways in which districts can allocate money to better align resources with the needs of their schools, based on the characteristics of the students they enroll.

To hold districts accountable for allocating these resources and providing this assistance, the state should explore ways to provide a system of incentives, guidance, support, and sanctions to districts that fail to make progress with their lowest-performing schools. Consequences such as removing latitude in governance, should be explored, however, since the effectiveness of district-level sanctions is yet to be determined, the state should regularly examine data to assess the effectiveness of any such system put in place.

It also seems worth considering whether the state should remove itself entirely from the business of directly sanctioning schools, and instead leave these activities to districts, holding districts accountable for results. This is likely much more manageable from a state perspective, is likely to be substantially more efficient, and seems to place responsibility clearly within the public education hierarchy that has been created by the state. If the state attempts to do everything itself (e.g., assume direct responsibility for all the state's underperforming schools), it seems very likely to be overextended and destined to fail.

Recommendation 3: **The state should acknowledge schools that improve within programs like II/USP or are successful with high percentages of students in poverty.**

As described above, the vast majority of low-performing schools in the state and across the nation are those with high percentages of students in poverty. The odds against success in these schools are formidable, and yet some schools are able to defy these odds to make substantial academic gains and to sustain this success. The findings from this study suggest that garnering attention by focusing only on schools that are failing has the potential to de-motivate schools. While there may be no way to avoid this kind of focus and the resulting attention on failing schools through a comprehensive accountability system, the fact that they *lose* supplemental resources when they succeed through a program like II/USP may not provide the best incentive for sustained improvement.

To ensure proper acknowledgment of success, the state should ensure that continued sufficient resources are available to schools that make growth. At the least, high poverty schools that are beating the odds should receive comparable *public acknowledgement* to those who are failing. This acknowledgment can serve both as a continuing motivator for growth, as well as a means to identify models for other high poverty schools that are not succeeding across the state. Successful, high poverty schools can provide concrete examples of what can be done to successfully improve student achievement.

Recommendation 4: **The state should work with districts to develop vehicles whereby schools in need of improvement can be linked with and learn from schools that have been successful in improving outcomes with comparable populations of students.**

In light of the limited communication reported and evident among schools in our sample, we recommend that the state and districts work in tandem to better facilitate this communication by establishing structures for schools to learn from each other. Districts should be assisted in learning how to pair struggling and successful schools, in cases where this is possible within a given district. The state should further assist in facilitating pairings of schools across districts as needed. The state should also work with districts to determine useful techniques for enhancing knowledge transfer across the two types of schools. The state should track the performance of schools that have been paired with others in an attempt to assess when these pairings have been successful and to make adjustments as needed based on results.

Although it is likely that costs will be associated with this kind of system (e.g., stipends to staff from successful schools offering assistance to similar, struggling schools), this may prove a more effective use of state funds than such vehicles as External Evaluators. It seems reasonable to expect that the state's greatest talent pool in regard to turning around low-performing schools resides in the administrators and staff who have demonstrated that this can be done.

The first step in developing such a structure is to establish clear criteria for success and to review data in light of these criteria to identify schools. The state should allocate personnel to review data yearly to this end. We suggest using straightforward criteria similar to those used in this study to enable the identification of consistent-growth and high-growth schools. While we focused on meeting growth targets and high levels of API growth to measure the success schools had in improving student outcomes, we encourage the state to also consider using the AYP, since schools are also now expected to meet these targets established by the federal government. Criteria should be kept simple to enable an easy classification of schools by growth each year, and to enable effective pairings of schools.

Another method to ensure communication between schools about best practices is to disseminate information that is compiled on strategies used by schools that have enabled growth. For example, disseminating the findings in this study will enable detailed information on factors for growth to reach schools that are undertaking improvement efforts. Other efforts have already been taken to do this, and should be continued. For example, ensuring a high level of participation at conferences such as the *On the Right Track* symposium sponsored by WestEd and the CDE could enable schools entering programs like HPSGP to learn from others. In addition, the National Center for Educational Accountability (NCEA) with several affiliates (including Just for Kids) has published a Web site where schools can identify other California schools that have met with success and learn about them.[6]

Recommendation 5: **Similarly, the state should identify *districts* that have successfully improved student outcomes in their low-performing schools, and facilitate opportunities for other districts to learn from them.**

Given that districts can play a key role in improvement efforts at their schools, and influence the implementation of state-level programs like II/USP, we recommend a similar effort as above, but targeted at districts. In this case, the state would again set simple and

[6] see: http://www.just4kids.org/bestpractice/evidence_finder.cfm?sub=tools

straightforward criteria for high-growth districts that can be reviewed on a yearly basis. Using these data, we recommend that the state provide opportunities, either through regular conferences and meetings, or through a more intensive partnering program, to facilitate learning across districts. Attention should be paid to the characteristics of districts to ensure that districts are matched with others that have similar student populations.

Recommendation 6: **The state should foster the use of benchmark data as a feedback mechanism for informing instruction and identifying students for extra support, as described in this report.**

A key factor that fostered growth in student outcomes identified through this study was the systematic analysis and use of data to inform instruction. Respondents emphasized the importance of frequent (e.g., monthly or every 6 weeks) benchmark assessments tied to the school's curriculum and tied to the state standards. Related ongoing communication and collaboration around data among grade-level and cross-grade-level colleagues were vital to this process. This was commonly mentioned as a key factor for turning around schools in our study, as well as in the general literature regarding effective practices. Not only does this seem a vital component for school growth, it can be much more easily replicated and made available to schools, along with requisite training, than less tangible factors such as strong leadership.

We recommend that systems be fostered and disseminated statewide that encourage and enable schools to utilize data on a regular and ongoing basis to inform instruction. Although training and ongoing support will be needed for the successful implementation of such systems, this is the kind of intervention that might be fostered through the types of teaming arrangements between districts and schools described above.

As it is clear that successful systems are already in place in some districts and schools realizing considerable gains in student achievement, the state role may be reduced to further identifying such systems and encouraging their further development and dissemination. These systems already in place were typically based on assessments associated with districtwide curricula tied to the state standards, and were supported by software packages that enabled the analysis and dissemination of assessment results in a user-friendly format.

One critical aspect to consider in establishing these systems is the training necessary for teachers and administrators. In order to understand the data, and to know how to utilize them in a way that supports instruction, teachers and administrators will need to receive specific training on the interpretation and use of results. Districts and counties should be expected to include such training as an important component of their professional development programs. In addition, districts would need to provide support by way of personnel who could be available to answer questions about the data. Counties could be the source of these personnel, as well as the types of teaming arrangements described above.

Recommendation 7: **The state should work with districts, and other agencies, to incorporate long-term guidance and monitoring into assistance programs for underperforming schools.**

Given the reportedly mixed experience with External Evaluators and the concerns raised by respondents about the limited support and monitoring received during the implementation years, we recommend that the state accountability system (e.g., HPSGP) incorporate sustained support and monitoring, while holding external providers accountable for assisting

and realizing results within their schools. This recommendation is bolstered by the finding in our achievement analyses that, for some cohorts, a decline in growth relative to comparison schools occurred after the planning year.

Specific areas in which sustained guidance and monitoring can be provided is in the use of funds, development of a coherent instructional program, and the development of strategies to sustain growth when funding ends. This guidance and support could come from either an external entity (that works with schools beyond just a planning year), or district offices.

Recommendation 8: **Analogous to what was reported for the school level, the *state* should use data on an ongoing basis to identify the extent to which state-level programs make an impact, and use these data to inform and alter policy and programs in support of low-performing schools as needed.**

As we set expectations for schools and districts to regularly use data as a basis for shaping policy and practice, we suggest the same process for the state. Above we recommended (see Recommendation 4) that the state allocate personnel to review data on school-level progress on a yearly basis. Here we suggest that these data be used not only to foster communication among schools with varying levels of growth, but also to examine state-level policies in regard to holding districts accountable for their schools' performance and in support of struggling schools to see what is working well and what needs to be revised. External evaluation studies such as this provide a means to gain formative and summative information on programs. However, given the high-stakes environment and urgency to improve student outcomes, the state itself should establish more mechanisms to review policies regularly, assess what components of its policies are on the right track, and adjust policies on an ongoing basis as needed.

Throughout such a review process, however, the state should pay attention to maintaining consistency across years. Too much change in reaction to short-term outcomes could create confusion about policy provisions and implementation expectations. Thus, before changes are made they must be thoughtfully considered and fully substantiated by data.

Recommendation 9: **The state should consider methods to better align the state and federal accountability systems. The state should take into consideration that stakeholders in low-performing schools generally report that while they consider the API to be a better outcome measure, they feel pressure to address AYP targets.**

Given the conflict and confusion associated with two overlapping accountability systems, we recommend that the state focus further on their alignment. However, we recognize that this is not an easy task and should be done carefully to preserve the most effective aspects of the state system.

For example, our data suggest that a growth model for accountability is important. For low-performing schools that are far from reaching the absolute proficiency targets set through NCLB, API growth targets provide incentives for schools to make continuous improvement. These targets account for the fact that these schools are starting from a very low base of performance.

Potential actions the state can take to this end include examining better ways to align the expectations and associated sanctions of the state model with the federal model. For example, the state could focus on having the *same* schools identified under *both* programs for similar

sanctions. Having some schools recognized for success within one model, but designated as failing in another, can result in confusion and unclear expectations. Additionally, the state can continue to press the federal government for greater incorporation of the state API into the AYP measure. Since there appears to be greater buy-in to a growth model, and to the API measure in general, this may be an important aspect of the state model to attempt to preserve.

SAIT Conclusions and Recommendations

In addition to these cross-cutting recommendations that address California's statewide accountability system, below we provide specific policy implications and recommendations for the ongoing implementation of the SAIT process. A number of these recommendations are based on the assumption that SAIT will remain in its current form. Based on our discussion above, however, we recommend that the state strongly consider some major changes in the overall design of the state accountability system, which would have implications for the SAIT process. As noted above, we recommend that the district and other intermediate agencies be more fully incorporated into the state system, and be held fully accountable for the outcomes of its schools. Under such a system, SAITs would not exist in their current form, where the level of district involvement is variable (despite specified expectations of substantial district involvement). We believe this overextends the state beyond what it can reasonably expect to effectively accomplish and bypasses other intermediary agencies that ultimately must be fully incorporated into any successful long-term reform and improvement process for the state's overall K-12 education system.

Making substantial, sustained progress in the state's lowest-performing schools is a very challenging undertaking, and each layer of the state's education system must be made responsible, and held accountable, in regard to their respective roles. With this idea in mind, something like the current SAIT process would be directed at districts that over time are failing to realize improvement in a high percentage of their low-performing schools. At the same time, the state might provide materials, training, and procedures for districts to follow in regard to SAIT-like processes at their schools.

Primary SAIT Recommendation: **The state should engage in ongoing assessment of the most successful and effective methods for realizing school improvement within the state monitoring process.**

It is still too early to know whether the current SAIT process, focused on the nine essential program components, will be effective in the aggregate in improving schools that failed through the II/USP process. Regardless, SAITs are the critical "next step" after II/USP in attempting to turn around these continuously low-performing schools. We received mixed assessments regarding the perceived helpfulness of SAITs from school respondents. In addition, it is reasonable to expect that whether SAITs are largely successful in the aggregate or not, that some SAITs will be more successful than others and that some strategies will be more effective than others. In light of the importance of this work, we recommend an ongoing and systematic data collection and assessment of the effectiveness of individual SAIT providers.

Specifically, the state should, on a yearly basis, assess the extent to which SAIT schools are improving (by looking at API growth) to see which of the EPCs were implemented during that year and what other types of supports the school had. This ongoing assessment should

provide a means to further focus the EPCs, to provide guidelines for their implementation, and to determine what aspects of the process are less effective.

Another vital component of this ongoing assessment of the SAIT process would focus on individual teams. These teams face a huge challenge in attempting to turn around underperforming schools. Their efforts are (at a minimum) based on an initial audit and three monitoring visits throughout the year.[7] This is relatively little intervention in relation to a large task. Many SAIT providers may not be up to this challenge, while others may have striking records of success. We believe that ongoing assessment of which SAIT providers are successful, under what conditions, and at what types of sites is critical to ensure that effective SAIT providers be given more work, and that those who show no impact either be retrained or replaced. It will also be important to learn from success. In addition to being given more work, SAIT providers who demonstrate success should be asked to provide input in regard to future SAIT policy and provide training to other SAITs.

To get a full assessment of the efficacy of individual providers, we also believe it will be important to talk with individual SAIT schools. In schools that have improved through SAIT, to what extent and in what ways do they think the SAIT helped them? From the perspective of schools not making progress, why did the SAIT process not yield better results? Also, assuming that the district will be more fully incorporated in the SAIT process, similar questions might be asked in regard to the district contribution. In what ways did the district help them, and what more might be done?

This ongoing data-based assessment should inform the state's decision-making process in regard to the SAIT system. Improvements should be made in real time when the data indicate a need to do so. We caution, however, against too much change based on limited short-term data, which could lead to confusion and ambiguity among participants in the process. This is an even stronger argument in favor of well-considered change, strongly based on data collection, analysis, and the resulting evidence of what is working and what is not.

Additional issues to consider in relation to SAIT: Below, we outline some specific key areas identified through this study that are potentially problematic, and should be further addressed through future research:

- **The state should consider adding components that address the instruction of English learner and special education populations.** SAIT providers noted that the needs of these groups were a critical omission from the EPCs and the SAIT process in general.

- **The state should consider including the enhancement of school climate and culture as a component within the state monitoring process.** Some SAIT providers reported that the exclusion of a school climate/culture component was problematic in schools with a "toxic" school culture and low levels of motivation among students and teachers.

- **Requirements for intervention program participation should be closely examined for adverse impact on enrollment in core courses at the secondary level.** Some SAIT, district, and school respondents discussed the complications associated with these intervention requirements, reporting that students were sometimes unable to complete core courses and enroll in electives.

[7] However, we found that many SAITs provided additional supports during the school year.

- **The state should reconsider allowing classroom observations for SAIT members.** Several SAIT providers reported that their ability to assess and monitor progress within the SAIT process was hindered by their lack of access to classrooms.

- **The list of EPCs should be bolstered with specific strategies and guidelines that enhance the quality of implementation of the EPCs, and that can be used in conjunction with the components to lead to greater success.**

- **The state should clarify the next steps to address schools that fail to improve through the state monitoring process.** These plans should be clearly communicated to participating schools. Unclear consequences were reported to increase confusion and anxiety surrounding the reform efforts and were reported to decrease motivation.

References

Argyris, C. and Schon, D.A. (1978). *Organizational learning: A theory of action perspective.* Reading, MA: Addison Wesley.

Bassok, Davis & Raymond (2004). *Not Condemned to Fail: An analysis of California's Low Performing Schools.* Stanford, CA: Hoover Institution Press.

Berne, R. and Stiefel, L. (1994). Measuring equity at the school level: the finance perspective. *Educational Evaluation and Policy Analysis*, 16 (4), 405-421.

Carnoy and Loeb (2002). Does External Accountability Affect Student Outcomes A Cross-State Analysis. *Educational Evaluation and Policy Analysis*, 24(4): 305-332.

Chambers, J.G., Parrish, T.B., Levin, J.D., Smith, J.R., Guthrie, J.W., Seder, R.C., & Taylor, L. (2004, March). *The New York Adequacy Study: Determining the Cost of Providing All Children in New York an Adequate Education – Final Report.* Palo Alto, CA: American Institutes for Research.

Cohen, J. (1969). *Statistical Power Analysis for Behavioral Sciences*. New York, NY: Academic Press.

Cohen, D.K. and Ball, D.L. (1999). *Instruction, Capacity, and Improvement.* Philadelphia, PA: Consortium for Policy Research in Education.

Clune, William (1993, June). *The Shift from Equity to Adequacy in School Finance*. Also published in Vol. 8 Educational Policy No. 376, 1994.

Darling-Hammond, L. (1996). Restructuring Schools for High Performance. In S. Fuhrman & J. A. O'Day (Eds.), *Rewards and Reform: Creating Educational Incentives that Work.* San Francisco, Jossey-Bass Publishers (144-192).

DeBray, E., Parson, G., & Woodworth, K. (2001). Patterns of response in four high schools under state accountability policies in Vermont and New York. In S. Fuhrman (Ed.), *From the capitol to the classroom: Standards-based reform in the states* (pp. 170-192). Chicago: University of Chicago Press.

Desforges, C. & Abouchaar, A. (2003). *The impact of parental involvement, parental support and family education on pupil achievement and adjustment: A literature review*. London: Department for Education and Skills.

Elmore, R. F. (2000, Winter). *Building a new structure for school leadership.* Washington, DC: The Albert Shanker Institute.

Elmore, R. F. (2001, April). *Psychiatrists and light bulbs: Educational accountability and the problem of capacity.* Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.

Finnigan, K.S. & O'Day, J.A. (2003) *External Support to Schools on Probation: Getting a Leg Up?* (In Press) Research report jointly published by the Consortium for Policy Research in Education and the Consortium on Chicago School Research, University of Pennsylvanian, Philadelphia, PA.

Grissmer, D.W., Flanagan, A., Kawata, J., & Williamson, S. (2000). *Improving student achievement: What NAEP state test scores tell us*. Santa Monica, CA: RAND Corporation.

Gunzenhauser, (2003). *High Stakes Testing and the Default Philosophy of Education.* Theory Into Practice, Vol. 42, No. 1, pp. 51-58.

Gwynne, J. & Easton, J. Q. (2001, April). *Probation, organizational capacity, and student achievement in Chicago elementary schools.* Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.

Harr, J.J., Pérez, M., McLaughlin, D., & Blankenship, C. (2005). *Participation of and accommodations for students with disabilities: Report on task 1.1: Participation policies* (NAEP State Analysis Project). Palo Alto, CA: American Institutes for Research.

Hoy, Wayne K. and Hannum, John W. (1997). Middle School Climate: An Empirical Assessment of Organizational Health and Student Achievement*. Educational Administration Quarterly*, 33(3), pp. 290-311.

Kannapel and Clements, (2005). *Inside the Black Box of High-Performing High-Poverty Schools*. A report from the Prichard Committee for academic Excellence, Ford Foundation: Lexington, KY.

Kirst, Michael (2005). Recent Evolution of California Accountability Policy. Paper prepared for State Legislative Seminar, Sacramento, CA, January 18, 2005.

Marsh, Julie A. (2000). *Connecting districts to the policy dialogue: A review of literature on the relationship of districts with states, schools and communities.* University of Washington: Center for the Study of Teaching and Policy.

Mintrop (2004). *High-Stakes Accountability, State Oversight, and Educational Equity*. **Teachers College Record,** Vol. 106, Issue 11, page 2128.

Muijs, D., Harris, A., Chapman, C., Stoll, L., & Russ, J. (2004). Improving schools in socioeconomically disadvantaged areas - A review of research evidence. *School Effectiveness & School Improvement, 15*(2), 149-175.

Neumann, Fred M. and Associates, (1996). *Authentic Achievement: Restructuring Schools for Intellectual Quality*. San Francisco, CA: Jossey-Bass.

Noddings (2004). *High Stakes Testing – Why?* Theory and Research in Education, SAGE Publications, Vol. 2, No. 3, pp. 263-269.

O'Day, J.A. (2002) Complexity, Accountability, and School Improvement, *Harvard Educational Review*, Vol. 72, No. 3, pp. 293-329.

O'Day, J.A. & Bitter, C. (2003). *Evaluation of the Immediate Intervention/Underperforming Schools Program and the High Achieving/Improving Schools Program of the California Public School Accountability Act of 1999*. Palo, Alto, Ca: American Institutes for Research.

Parrish, T.B. et al. (2005). *How Are English Learners Faring in California Post-Proposition 227? – Preliminary Draft of Final Report for AB56.* Final report submitted to the California Department of Education. Palo Alto, CA: American Institutes for Research.

Peternick, L., Smerdon, B.A., Fowler, W., & Monk, D. (1997). Using Cost and Need Adjustments to Improve the Measurement of School Finance Equity. *Developments in School Finance 1997,* 151-168.

Raymond and Hanushek (2004). Developments in School Finance: 2003. *Shopping for Evidence Against School Accountability,* (117-129). Stanford, CA.

Reschovsky, A. and Iimazeki, J. (1997). The Development of School Finance Formulas to Guarantee the Provision of Adequate Education to Low-Income Students. *Developments in School Finance 1997*. Washington, DC: National Center for Education Statistics.

Russell, Higgins and Raczek, (2003). *Accountability, California Style: Counting or Accounting?* Teachers College Record, 106(11).

Terry, Paul. (1996) The Principal and Instructional Leadership. Paper presented at Annual Conference of the National Council of Professors of Education Administration, 7 August.

Zigarelli, Michael. (1996). An Empirical Test of Conclusions from Effective Schools Research. *Journal of Educational Research,* 90(2), pp.103-109.

Evaluation Study of the Immediate Intervention/Underperforming Schools Program of the Public Schools Accountability Act of 1999

Appendices

Submitted to:

Kim Wells, Contract Monitor
Evaluation Unit
Policy and Evaluation Division
California Department of Education
1430 "N" Street
Sacramento, CA 95814

September 30, 2005

Appendix A:
Supplementary Methods and Tables for the Achievement Analysis

Appendix A-1:
Methodology

Methodology

In our analyses of API scores, we have data for multiple time points for each school; and in our analysis of student-level scores, we have data for multiple students and multiple time points for each school. To take the multi-level nature of the data into account, we employed hierarchical linear modeling methods. These methods make it possible to distinguish the effects of measured student-level and/or school-level factors, as well as the effect of time, and they also address the fact that students within the same schools are likely to have characteristics in common that we are unable to measure (such as common community characteristics). The equations for the model for the API are shown in Exhibit A1, and those for the student-level scores are shown in Exhibit A2.

The primary hypotheses of interest concern differences in achievement trajectories between II/USP and comparison schools over the period from 1998 through 2004. To test these time-specific hypotheses for Cohort 1, we created an indicator variable (IIUSP1) to reflect participation in II/USP Cohort 1 (coded 1 for participants and 0 for comparison schools), and six variables to reflect the calendar year (YEAR99, which is coded 0 in 1998 and 1 in 1999, 2000, 2001, 2002, 2003, and 2004; YEAR00, which is coded 0 in 1998 and 1999, and 1 in 2000, 2001, 2002, 2003, and 2004; YEAR01, which is coded 0 in 1998, 1999, and 2000, and 1 in 2001 2002, 2003, and 2004; YEAR02, which is coded 0 in 1998, 1999, 2000, and 2001 and 1 in 2002, 2003, and 2004; YEAR03, which is coded 0 in 1998, 1999, 2000, 2001, and 2002 and 1 in 2003, and 2004; YEAR04, which is coded 0 in 1998, 1999, 2000, 2001, 2002, and 2003 and 1 in 2004). The hypothesized effects we seek involve interactions of IIUSP1 and YEAR99, YEAR00, YEAR01, YEAR02, YEAR03, and YEAR04.[1] If participating and comparison schools were similar prior to participation, we would expect the interaction of YEAR99 and IIUSP1 to be zero, whereas if participation has a positive effect on subsequent growth, we would expect the interaction of YEAR00 and IIUSP1 to be positive, as well as the interaction of YEAR01 and IIUSP1, as well as YEAR02 and IIUSP1, and so on.

To determine whether the trajectories for CSR schools differ from those for regular (Action Plan) II/USP schools, we created an indicator variable (CSR), coded 1 for CSR schools, and 0 for other schools, and we included the interactions of the CSR with each of the year variables (YEAR99 through YEAR04). To test the time-specific hypotheses for Cohort 2, and 3, we created a parallel set of indicator variables. We also included an HPSGP indicator for Cohort 3 schools.

[1] The equations that appear in Exhibits A1 and A2 display the models in conventional two-level hierarchical linear model form. The "time-level" model represents the effects of year on achievement, as well as the effects of school characteristics. The "school level" model represents the effects of II/USP participation status on the level-one slopes for year. These school-level effects of II/USP represent the interactions between II/USP and year discussed in the text.

Exhibit A1: Multilevel Model for API Scores

a) Time-level model

$$y_{tj} = \beta_{0j} + \beta_{1j}Year99_{tj} + \beta_{2j}Year00_{tj} + \beta_{3j}Year01_{tj} + \beta_{4j}Year02_{tj} + \beta_{5j}Year03_{tj} + \beta_{6j}Year04_{tj} + \beta_{7j}Pct\_Black_{tj} + \beta_{8j}Pct\_Asian_{tj} + \beta_{9j}Pct\_Hisp_{tj} + \beta_{10j}Pct\_Meals_{tj} + \beta_{11j}Pct\_ELL_{tj} + \beta_{12j}Mobility_{tj} + \beta_{13j}Avg\_Pared_{tj} + \beta_{14j}Pct\_Full_{tj} + \varepsilon_{tj}$$

where:

$YEAR99_{tj}$ is coded 0 for 1998; and 1 for 1999, 2000, and 2001;

$YEAR00_{tj}$ is coded 0 for 1998 and 1999; and 1 for 2000 and 2001;

$YEAR01_{tj}$ is coded 0 for 1998, 1999, and 2000; and 1 for 2001;

$YEAR02_{tj}$ is coded 0 for 1998, 1999, 2000, 2001; and 1 for 2002;

$YEAR03_{tj}$ is coded 0 for 1998, 1999, 2000, 2001, 2002; and 1 for 2003; and

$YEAR04_{tj}$ is coded 0 for 1998, 1999, 2000, 2001, 2002, 2003; and 1 for 2004.

other terms are defined as in Appendix A-4.

b) School-level model for Cohort 1

$$\beta_{0j} = \gamma_{00} + \gamma_{01}IIUSP1_j + \gamma_{02}CSR1_j + \nu_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}IIUSP1_j + \gamma_{12}CSR1_j$$
$$\beta_{2j} = \gamma_{20} + \gamma_{21}IIUSP1_j + \gamma_{22}CSR1_j$$
$$\beta_{3j} = \gamma_{30} + \gamma_{31}IIUSP1_j + \gamma_{32}CSR1_j$$
$$\beta_{4j} = \gamma_{40} + \gamma_{41}IIUSP1_j + \gamma_{42}CSR1_j$$
$$\beta_{5j} = \gamma_{30} + \gamma_{51}IIUSP1_j + \gamma_{52}CSR1_j$$
$$\beta_{6j} = \gamma_{40} + \gamma_{61}IIUSP1_j + \gamma_{62}CSR1_j$$

where:

β_{0j} is the intercept for school *j* in the time level-model;

β_{1j} is the slope for YEAR99 for school *j*;

β_{2j} is the slope for YEAR00 for school *j*;

β_{3j} is the slope for YEAR01 for school *j*;

β_{4j} is the slope for YEAR02 for school *j*;

β_{5j} is the slope for YEAR03 for school *j*;

β_{6j} is the slope for YEAR04 for school *j*;

$IIUSP1_j$ is a 0/1 variable indicating whether school *j* is a member of IIUSP Cohort 1;

$CSR1_j$ is a 0/1 variable indicating whether school *j* is a Cohort 1 CSR school; and

v_{0j} is a random error term representing unmeasured factors related to the intercept of the growth curve for school j.

The model for Cohort 2 is similar. In the model for Cohort 3, an HPSGP indicator variable also appears.

Exhibit A2: Multilevel model for standardized SAT-9 and CAT/6 scale scores

a) Student/time-level model[2]

$$y_{itj} = \beta_{0j} + \beta_{1j}Year99_{itj} + \beta_{2j}Year00_{itj} + \beta_{3j}Year01_{itj} + \beta_{4j}Year02_{itj} + \beta_{5j}Year03_{itj} + \beta_{6j}Year04_{itj} + \beta_{7}Girl_{itj} + \beta_{8}EL_{itj} + \beta_{9}FLunch_{itj} + \beta_{10}Asian_{itj} + \beta_{11}Black_{itj} + \beta_{12}Hispanic_{itj} + \beta_{13}Others_{itj} + \beta_{14}HighPared_{itj} + \beta_{15}Grade3_{itj} + \beta_{16}Grade4_{itj} + \beta_{17}Grade5_{itj} + \varepsilon_{itj}$$

where:

$YEAR99_{itj}$, $YEAR00_{itj}$, $YEAR01_{itj}$, $YEAR02_{itj}$, $YEAR03_{itj}$ and $YEAR04_{itj}$ are coded as in Exhibit A1;

$Grade3_{itj}$, $Grade4_{itj}$, *and* $Grade5_{itj}$ are 0/1 variables indicating the student's grade level (with similar variables included in models for middle and high schools); and other terms are defined as in Appendix A-4.[3]

[2] For this analysis, we combined the student and time levels. This in effect assumes that there are no unmeasured differences across years, once the trend captured by the six dummy variables (YEAR99, YEAR00, YEAR01, YEAR02, YEAR03, and YEAR04) are accounted for. The model in effect also assumes that there are no stable unmeasured differences among cohorts (i.e., groups of students who enter in the same year).

[3] The student/time-level model also includes the missing value dummy variable for parent education described in Appendix A-4.

b) School-level model for Cohort 1

$$\beta_{0j} = \gamma_{00} + \gamma_{01} IIUSP1_j + \gamma_{02} CSR1_j + \gamma_{03} Pct\_Meals_j + \nu_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} IIUSP1_j + \gamma_{12} CSR1_j$$
$$\beta_{2j} = \gamma_{20} + \gamma_{21} IIUSP1_j + \gamma_{22} CSR1_j$$
$$\beta_{3j} = \gamma_{30} + \gamma_{31} IIUSP1_j + \gamma_{32} CSR1_j$$
$$\beta_{4j} = \gamma_{40} + \gamma_{41} IIUSP1_j + \gamma_{42} CSR1_j$$
$$\beta_{5j} = \gamma_{50} + \gamma_{51} IIUSP1_j + \gamma_{52} CSR1_j$$
$$\beta_{6j} = \gamma_{60} + \gamma_{61} IIUSP1_j + \gamma_{62} CSR1_j$$

where:

β_{0j} is the intercept for school *j* in the student/time level-model;

β_{1j} is the slope for YEAR99 for school *j*;

β_{2j} is the slope for YEAR00 for school *j*;

β_{3j} is the slope for YEAR01 for school *j*;

β_{4j} is the slope for YEAR02 for school *j*;

β_{5j} is the slope for YEAR03 for school *j*;

β_{6j} is the slope for YEAR04 for school *j*;

$IIUSP1_j$ is a 0/1 variable indicating whether school *j* is a member of IIUSP Cohort 1;

$CSR1_j$ is a 0/1 variable indicating whether school *j* is a Cohort 1 CSR school;

Pct_Meals_j is a variable indicating the percent of students eligible for free or reduced lunch in school *j*; and

ν_{0j} is a random error term representing unmeasured factors related to the intercept of the growth curve for school j

The model for Cohort 2 is similar. In the model for Cohort 3, an HPSGP indicator variable also appears.

Appendix A-2:
Synthetic API Calculation

Synthetic API Calculation

School-level API scores are the primary emphasis of the II/USP program, and thus they are a central outcome measure. To use API scores in the analysis, however, several challenges must be overcome. In particular, because the focus of the evaluation is on *change over time* in school performance, it is important for performance to be measured on a consistent scale over the six years under study (1998-2004). While official API scores are available for spring 1999, 2000, 2001, 2002, 2003, and 2004 they are not available for 1998. In addition, the method used to calculate official API scores changed across the three years for which scores are available, due to the incorporation of additional assessments.[4]

To overcome these challenges, we created a set of *synthetic* API scores for each school, using the school's official 1999 base API as the starting point. We refer to these scores as "synthetic" to reflect the fact that, although they are derived from each school's official API scores for the years under study, we have transformed them slightly to make them more comparable over time, and we have computed scores for 1998, for which official API scores are not available.

To calculate the synthetic 2000 API, we added the school's official API growth from 1999 to 2000 to the school's base score.[5] (See box below.) To calculate a synthetic 1998 API, we used student-level *SAT-9* scores to compute a synthetic API for both 1998 and 1999, employing the rules the CDE used to calculate the official 1999 base year scores, but including all students tested. We then computed the change from 1998 to 1999 using this synthetic score for both years and subtracted it from the 1999 base.

Derivation of Synthetic API Scores

1998 synthetic API = 1999 base API *minus* 1998-99 API growth

1999 synthetic API = 1999 base API

2000 synthetic API = 1999 base API *plus* 1999-00 API growth

2001 synthetic API = 2000 synthetic API *plus* 2000-01 API growth

2002 synthetic API = 2001 synthetic API *plus* 2001-02 API growth

2003 synthetic API = 2002 synthetic API *plus* 2002-03 API growth

2004 synthetic API = 2003 synthetic API *plus* 2003-04 API growth

[4] The rules used to exclude students based on mobility also changed between 1999 and 2000.

[5] Each year after 1999, two different API scores are available for each school – a base score, which is used as the basis for calculating growth over the coming year, and a growth score, which is used as the end-point in calculating growth over the previous year. The two scores are required because the rules used by the State to define API scores changed somewhat each year The base score incorporates the changes in the definition of the API that have been adopted since the previous year, while the growth score is based on the previous year's definition. A school's growth from 1999 to 2000 is computed by subtracting the 1999 base score from the school's 2000 growth score.

Appendix A-3:
Number of Schools with Missing Synthetic API data

Number of II/USP, and Comparison Schools[6] with Missing Synthetic API Data for One or More Years Using Data from 1998 to 2004

	Elementary	Middle	High	Missing	Total
		Cohort 1 II/USP schools			
Schools with complete API scores	245	66	43	0	354
Schools missing API scores	52	12	8	4	76
Total	297	78	51	4	430
		Cohort 1 comparison schools			
Schools with complete API scores	639	150	91	0	880
Schools missing API scores	72	19	29	0	120
Total	711	169	120	0	1000
		Cohort 2 II/USP schools			
Schools with complete API scores	233	86	60	0	379
Schools missing API scores	24	16	11	0	51
Total	257	102	71	0	430
		Cohort 2 comparison schools			
Schools with complete API scores	190	109	122	0	421
Schools missing API scores	39	17	28	0	84
Total	229	126	150	0	505
		Cohort 3 II/USP schools			
Schools with complete API scores	244	44	56	0	344
Schools missing API scores	49	7	19	11	86
Total	293	51	75	11	430
		Cohort 3 comparison schools			
Schools with complete API scores	399	148	133	0	680
Schools missing API scores	85	29	20	19	153
Total	484	177	153	19	833

[6] By 2004, 398 Cohort 1 comparison schools and 81 Cohort 2 comparison schools had been dropped from the group because they received II/USP in subsequent years.

Appendix A-4:
List of Control Variables

Student-level and school-level control variables included in models

Variable names	Values
Student-level variables	
GIRL	1=female, 0=male
ASIAN	1=Asian, 0=not Asian
BLACK	1=African/African American, 0=not African/African American
HISPANIC	1=Hispanic, 0=not Hispanic
OTHERS	1=American Indian or Alaska Native, Filipino/Filipino American, Pacific Islander and other, and 0= not American Indian or Alaska Native, Filipino/Filipino American, Pacific Islander or other
EL	1=English Learner, 0=Other students
R_FEP	1=Re-designated Fluent English Proficient (R-FEP), 0=Other students
ELMISN	1=if English fluency variable missing, 0=not missing
FEP	1=Fluent English Proficient (FEP), and 4=English only
FLUNCH	1=eligible for free or reduced price lunch, and 0=not eligible
HIGHPARED	1=Not a high school graduate, 2=High school graduate, 3=Some college, 4=College graduate, 5=Graduate school/post graduate training (missing cases imputed using the school mean)
PARED_MISSING	1=if ParEd missing, 0=not missing
SPECED	1=students received special education
School-level variables	
PCT_ASIAN	Percent Asian students (0 to 100)
PCT_BLACK	Percent African American students (0 to 100)
PCT_HISP	Percent Hispanic students (0 to 100)
PCT_ELL	Percent English language learners (0 to 100)
AVG_PARED	Average education level of students' parents (1 to 5)
MOBILITY	Percent of students first attending this school in current year (0 to 100)
PCT_FULL	Percent of teachers with full credential (0 to 100)
PCT_MEALS	Percent of students eligible for free or reduced price lunch (0 to 100)

Appendix A-5:
II/USP Status

Status of Elementary Schools that Participated in II/USP (n = 853)

II/USP Cohort	Exited		Watched		Monitored		Closed	
	N	%	N	%	N	%	N	%
Cohort 1	236	79.7	16	5.4	37	12.5	7	2.4
Cohort 2	165	64.7	37	14.5	47	18.4	6	2.4
Cohort 3	74	25.1	212	71.9	1	0.3	8	2.7
Total	475	56.1	265	31.3	85	10.0	21	2.5

Status of Middle Schools that Participated in II/USP (n = 232)

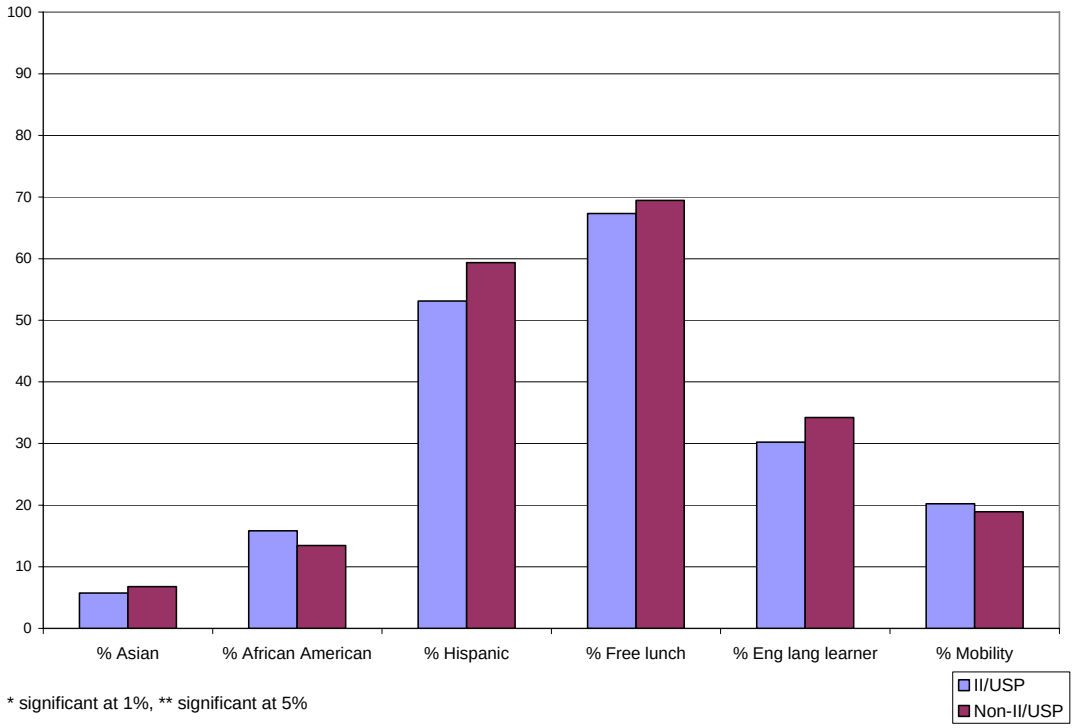
II/USP Cohort	Exited		Watched		Monitored		Closed	
	N	%	N	%	N	%	N	%
Cohort 1	53	68.8	11	14.3	11	14.3	2	2.6
Cohort 2	47	46.1	37	36.3	16	15.7	2	2.0
Cohort 3	14	27.5	35	68.6	1	2.0	1	2.0
Total	114	49.6	83	36.1	28	12.2	5	2.2

Status of High Schools that Participated in II/USP (n = 203)

II/USP Cohort	Exited		Watched		Monitored		Closed	
	N	%	N	%	N	%	N	%
Cohort 1	28	52.8	6	11.3	18	34.0	1	1.9
Cohort 2	18	25.7	36	51.4	12	17.1	4	5.7
Cohort 3	14	19.7	53	74.6	3	4.2	1	1.4
Total	60	30.9	95	49.0	33	17.0	6	3.1

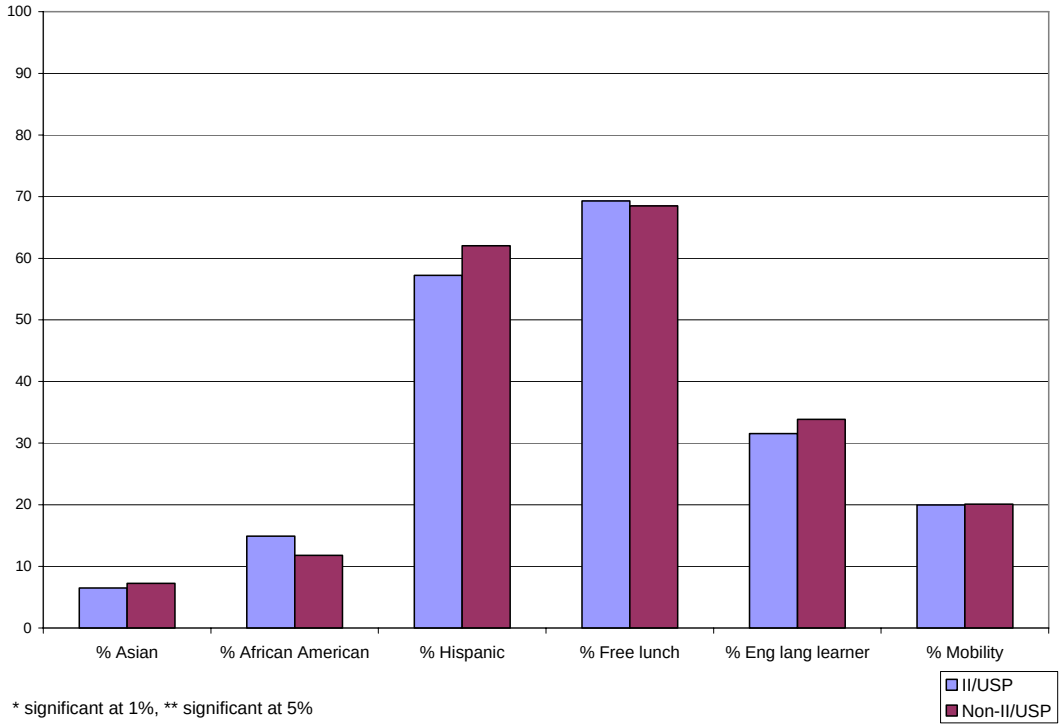
Appendix A-6:
Demographic Characteristics

Demographic Characteristics of Cohort 1 II/USP and Comparison Middle Schools



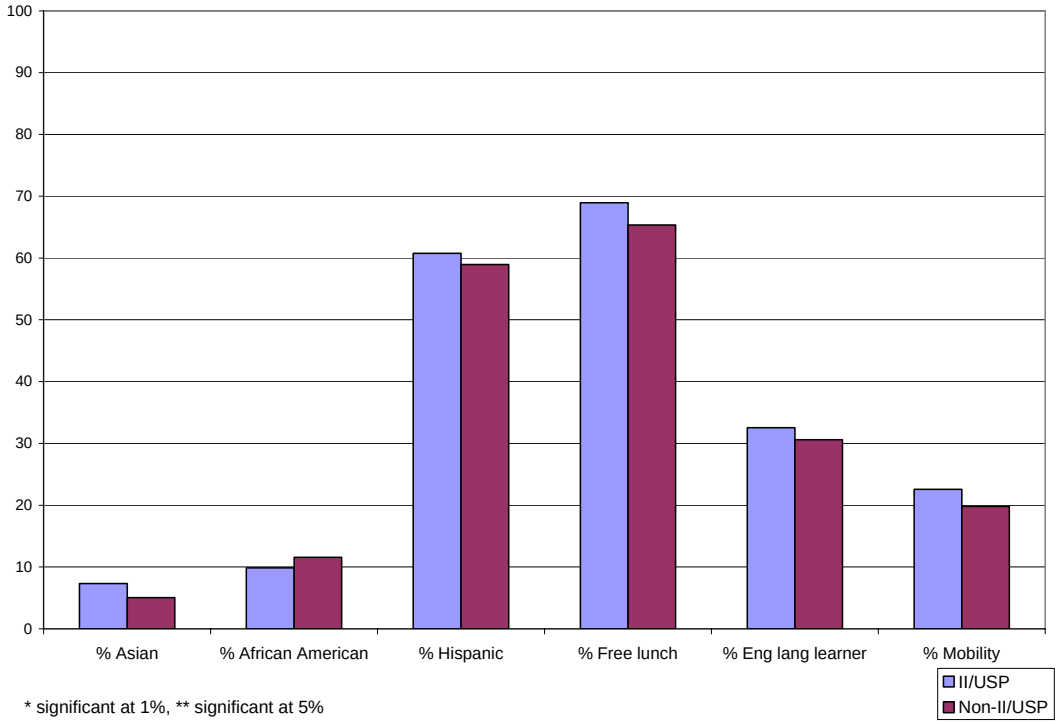
* significant at 1%, ** significant at 5%

Demographic Characteristics of Cohort 2 II/USP and Comparison Middle Schools



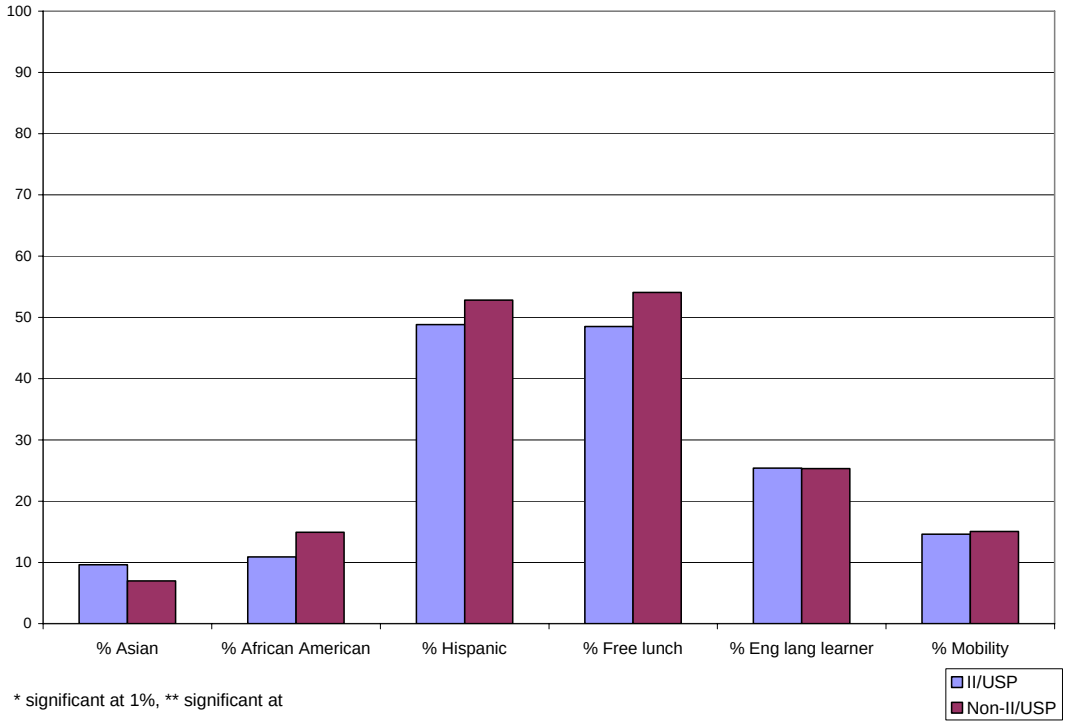
* significant at 1%, ** significant at 5%

Demographic Characteristics of Cohort 3 II/USP and Comparison Middle Schools



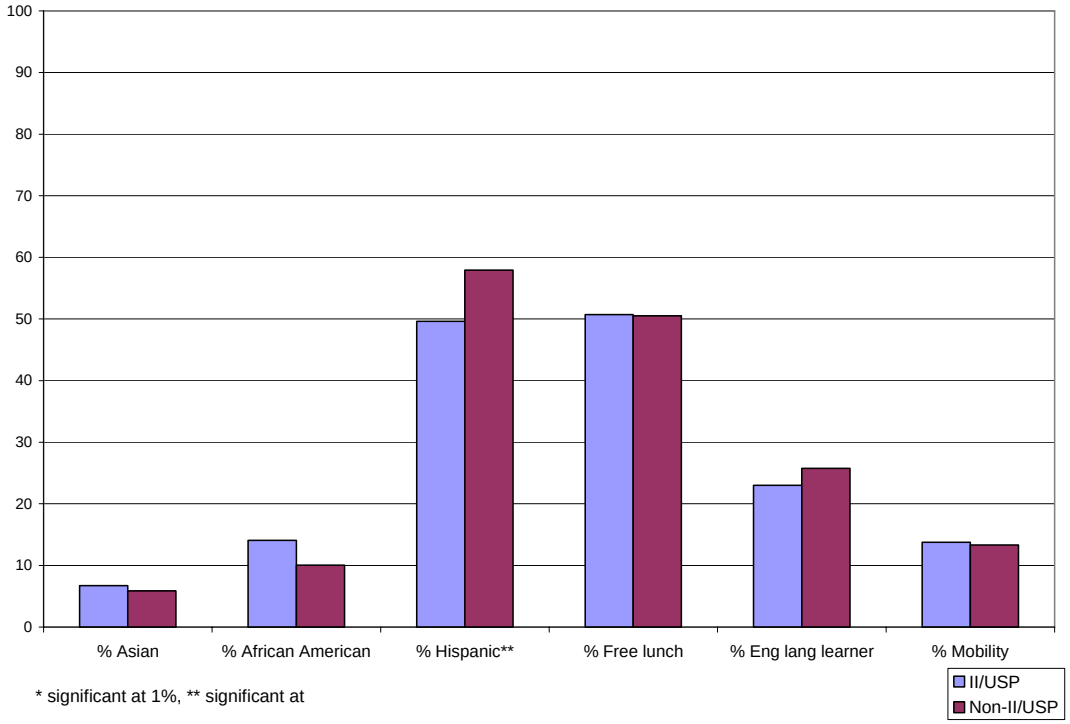
* significant at 1%, ** significant at 5%

Demographic Characteristics of Cohort 1 II/USP and Comparison High Schools



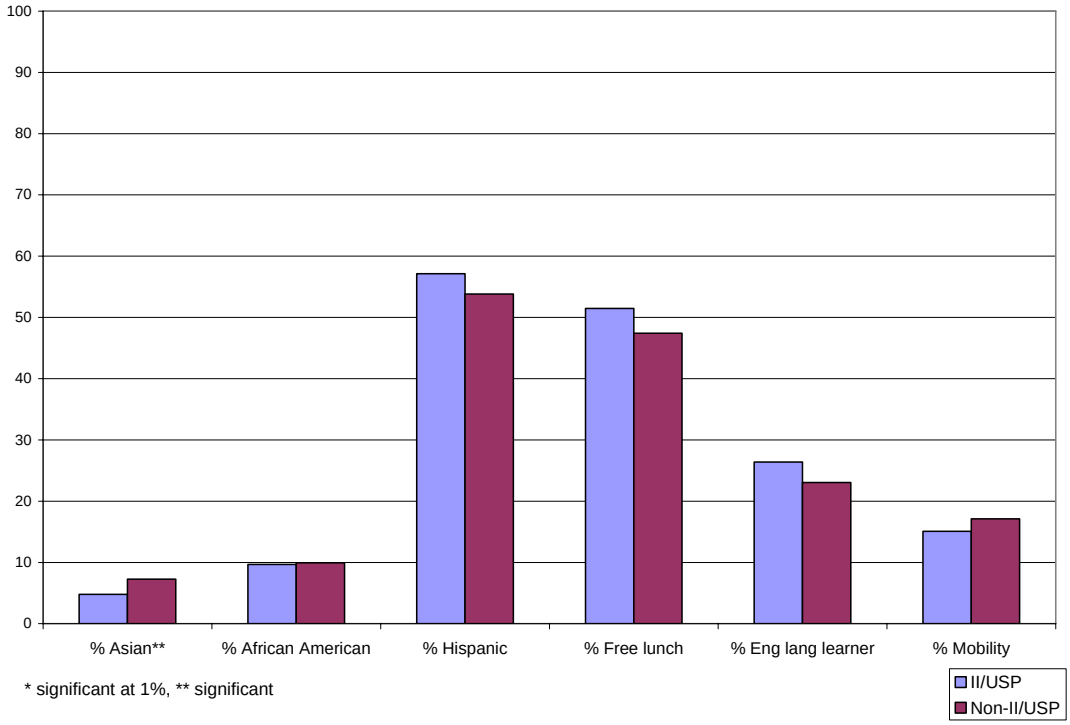
* significant at 1%, ** significant at

Demographic Characteristics of Cohort 2 II/USP and Comparison High Schools



* significant at 1%, ** significant at

Demographic Characteristics of Cohort 3 II/USP and Comparison High Schools



* significant at 1%, ** significant

Appendix A-7:
Overall Effect Sizes and Sub-Population Effect Sizes Calculation

Effect Sizes for API, CST, SAT-9 and CAT/6 Score Changes in II/USP Schools Relative to Comparison Schools

	Cohort 1					Cohort 2				Cohort 3		
	Planning (2000)	Imp 1 (2001)	Imp. 2 (2002)	Post II/USP (2003)	Post II/USP (2004)	Planning (2001)	Imp 1 (2002)	Imp. 2 (2003)	Post II/USP (2004)	Planning (2002)	Imp 1 (2003)	Imp. 2 (2004)
Elementary												
API effect size	0.086*	-0.098*	-0.161**	-0.055	-0.015	0.094*	0.063	0.029	-0.068	0.003	0.007	0.107*
SAT-9 and CAT/6 Reading effect size	0.006	-0.020**	-0.020**	-0.018**	-0.011**	0.032**	0.008	-0.008	-0.015**	-0.005	-0.001	0.013*
SAT-9 and CAT/6 Math effect size	0.041**	0.005	-0.030**	-0.025**	-0.008	0.015**	0.008	0.023	-0.014**	-0.008	0.010*	0.007
CST ELA effect size	-	-	-	-	-	-	-	-	-	-0.012*	0.005**	0.013**
CST math effect size	-	-	-	-	-	-	-	-	-	-0.041	-0.018**	0.014
Middle												
API effect size	0.135*	0.027	-0.035	-0.112	-0.062	-0.018	-0.014	-0.017	-0.092	-0.065	0.003	-0.044
SAT-9 and CAT/6 Reading effect size	0.010	0.016**	0.015**	-0.024**	-0.060**	0.009	-0.016**	0.012**	-0.018**	-0.016**	-0.014**	-0.012*
SAT-9 and CAT/6 Math effect size	0.036**	0.007	0.015**	-0.022**	-0.034	-0.001	-0.019**	0.013**	-0.002	-0.015**	-0.004	-0.007
CST ELA effect size	-	-	-	-	-	-	-	-	-	0.005	-0.017**	-0.033**
CST math effect size	-	-	-	-	-	-	-	-	-	-0.020	-0.058**	-0.036**
High												
API effect size	0.073	0.064	-0.037	-0.219*	0.108	0.143	-0.036	-0.006	-0.046	-0.050	0.131	-0.100
SAT-9 and CAT/6 Reading effect size	-0.004	0.020**	0.022**	-0.026**	0.051**	0.025**	0.001	0.001	-0.009	-0.020**	0.028**	-0.015**
SAT-9 and CAT/6 Math effect size	0.008	0.026**	0.020**	-0.021**	0.022**	0.018**	0.005	0.022**	-0.014**	0.003	0.026**	-0.002
CST ELA effect size	-	-	-	-	-	-	-	-	-	0.036	0.049	0.049
CST math effect size	-	-	-	-	-	-	-	-	-	0.019	0.016**	0.025

* significant at .05 level
**significant at .01 level

Effect Sizes for Subpopulations: Change in Achievement Gap Between ELs and EOs, Standardized SAT-9 and CAT/6 Reading Scale Scores by School Type and Cohort*

COHORT 1	1997-1998	1998-1999	1999-2000	2000-2001	2001-2002	2002-2003	2003-2004	1999-2002	II/USP effect on 1999-2002 gap	1999-2004	II/USP effect on 1999-2004 gap
Comparison Elementary Schools	-0.41	-0.43	-0.42	-0.41	-0.38	-0.27	-0.26	0.05**		0.17**	
II/USP Elementary Schools	-0.46	-0.43	-0.42	-0.42	-0.40	-0.28	-0.27	0.04**	-0.01	0.16**	-0.01
Comparison Middle Schools	-0.57	-0.61	-0.58	-0.60	-0.60	-0.51	-0.51	0.01		0.10**	
II/USP Middle Schools	-0.60	-0.61	-0.56	-0.56	-0.57	-0.46	-0.51	0.05**	0.04**	0.11**	0.01
Comparison High Schools	-0.67	-0.67	-0.66	-0.66	-0.66	-0.52	-0.54	0.01		0.13**	
II/USP High Schools	-0.77	-0.73	-0.67	-0.63	-0.64	-0.47	-0.50	0.08**	0.07**	0.23**	0.10**

COHORT 2	1997-1998	1998-1999	1999-2000	2000-2001	2001-2002	2002-2003	2003-2004	2000-2003	II/USP effect on 2000-2003 gap	2000-2004	II/USP effect on 2000-2004 gap
Comparison Elementary Schools	-0.46	-0.44	-0.43	-0.45	-0.43	-0.30	-0.29	0.131**		0.145**	
II/USP Elementary Schools	-0.43	-0.41	-0.39	-0.40	-0.36	-0.27	-0.24	0.127**	0.00	0.154**	0.01
Comparison Middle Schools	-0.59	-0.62	-0.60	-0.60	-0.60	-0.49	-0.524	0.103**		0.073**	
II/USP Middle Schools	-0.62	-0.65	-0.59	-0.58	-0.58	-0.47	-0.48	0.117**	0.01	0.113**	0.030**
Comparison High Schools	-0.66	-0.69	-0.66	-0.65	-0.65	-0.51	-0.558	0.149**		0.102**	
II/USP High Schools	-0.76	-0.73	-0.68	-0.69	-0.68	-0.51	-0.55	0.177**	0.027**	0.139**	0.037**

COHORT 3	1997-1998	1998-1999	1999-2000	2000-2001	2001-2002	2002-2003	2003-2004	2001-2004	II/USP effect on 2001-2004 gap
Comparison Elementary Schools	-0.48	-0.48	-0.44	-0.45	-0.44	-0.30	-0.30	0.15**	
II/USP Elementary Schools	-0.44	-0.44	-0.40	-0.42	-0.39	-0.28	-0.27	0.15**	-0.01
Comparison Middle Schools	-0.59	-0.61	-0.59	-0.59	-0.59	-0.48	-0.49	0.10**	
II/USP Middle Schools	-0.62	-0.65	-0.61	-0.59	-0.59	-0.48	-0.52	0.08**	-0.02**
Comparison High Schools	-0.71	-0.69	-0.67	-0.66	-0.67	-0.53	-0.54	0.12**	
II/USP High Schools	-0.75	-0.70	-0.69	-0.70	-0.67	-0.54	-0.56	0.14**	0.03**

*A positive II/USP effect indicates that II/USP closed the gap between ELs and EOs at a higher rate than comparison schools.

Effect Sizes for Subpopulations: Change in Achievement Gap Between Special Education and Non-Special Education Students, Standardized SAT-9 and CAT/6 Reading Scale Scores by School Type and Cohort*

COHORT 1	1997-1998	1998-1999	1999-2000	2000-2001	2001-2002	2002-2003	2003-2004	1999-2002	II/USP effect on 1999-2002 gap	1999-2004	II/USP effect on 1999-2004 gap
Elementary Reading Comparison	-0.54	-0.49	-0.52	-0.53	-0.58	-0.70	-0.71	-0.09**		-0.22**	
Elementary Reading II/USP	-0.58	-0.55	-0.55	-0.57	-0.60	-0.70	-0.73	-0.05**	0.04*	-0.18**	0.04*
Middle Reading Comparison	-0.69	-0.67	-0.67	-0.68	-0.69	-0.78	-0.76	-0.02		-0.09**	
Middle Reading II/USP	-0.73	-0.81	-0.74	-0.79	-0.79	-0.83	-0.84	0.02	0.04*	-0.03	0.06**
High Reading Comparison	-0.81	-0.74	-0.73	-0.72	-0.72	-0.73	-0.68	0.02		0.06**	
High Reading II/USP	-0.74	-0.80	-0.77	-0.80	-0.79	-0.78	-0.78	0.01	-0.01	0.02	-0.04

COHORT 2	1997-1998	1998-1999	1999-2000	2000-2001	2001-2002	2002-2003	2003-2004	2000-2003	II/USP effect on 2000-2003 gap	2000-2004	II/USP effect on 2000-2004 gap
Elementary Reading Comparison	-0.62	-0.58	-0.56	-0.58	-0.63	-0.77	-0.73	-0.21**		-0.17**	
Elementary Reading II/USP	-0.54	-0.50	-0.50	-0.53	-0.60	-0.70	-0.70	-0.19**	0.02	-0.20**	-0.03
Middle Reading Comparison	-0.74	-0.73	-0.70	-0.68	-0.70	-0.78	-0.81	-0.07*		-0.10**	
Middle Reading II/USP	-0.74	-0.73	-0.68	-0.63	-0.61	-0.68	-0.70	0.00	0.07	-0.02	0.08
High Reading Comparison	-0.86	-0.68	-0.78	-0.77	-0.77	-0.76	-0.73	0.02		0.05**	
High Reading II/USP	-0.85	-0.77	-0.74	-0.80	-0.79	-0.80	-0.76	-0.06**	-0.08**	-0.03*	-0.08**

COHORT 3	1997-1998	1998-1999	1999-2000	2000-2001	2001-2002	2002-2003	2003-2004	2001-2004	II/USP effect on 2001-2004 gap
Elementary Reading Comparison	-0.63	-0.57	-0.58	-0.57	-0.62	-0.73	-0.71	-0.14**	
Elementary Reading II/USP	-0.60	-0.54	-0.55	-0.56	-0.63	-0.76	-0.77	-0.21**	-0.07**
Middle Reading Comparison	-0.77	-0.80	-0.75	-0.76	-0.76	-0.83	-0.80	-0.04**	
Middle Reading II/USP	-0.71	-0.66	-0.64	-0.69	-0.70	-0.77	-0.79	-0.10**	-0.06**
High Reading Comparison	-0.88	-0.82	-0.80	-0.79	-0.80	-0.80	-0.747	0.04**	
High Reading II/USP	-0.78	-0.76	-0.72	-0.75	-0.76	-0.75	-0.73	0.02	-0.02

***A positive II/USP effect indicates that the II/USP schools' gap between special education students and students receiving regular education increased at a lower rate than comparison schools.**

Effect Sizes for Subpopulations: Change in Achievement Gap Between Free Lunch and Non-Free Lunch Students, Standardized SAT-9 Reading Scale Scores by School Type and Cohort*

COHORT 1	1997-1998	1998-1999	1999-2000	2000-2001	2001-2002	2002-2003	2003-2004	1999-2002	II/USP effect on 1999-2002 gap	1999-2004	II/USP effect on 1999-2004 gap
Elementary Reading Comparison	-0.14	-0.21	-0.24	-0.23	-0.21	-0.10	-0.079	0.01		0.135**	
Elementary Reading II/USP	-0.12	-0.19	-0.26	-0.21	-0.22	-0.10	-0.082	-0.020**	-0.035**	0.100**	-0.025*
Middle Reading Comparison	-0.10	-0.15	-0.17	-0.16	-0.15	-0.05	-0.03	0.00		0.122**	
Middle Reading II/USP	-0.11	-0.17	-0.21	-0.18	-0.18	-0.08	-0.07	-0.01	-0.01	0.096**	-0.026*
High Reading Comparison	-0.03	-0.06	-0.10	-0.10	-0.09	-0.04	-0.027	-0.030**		0.028**	
High Reading II/USP	-0.12	-0.16	-0.14	-0.13	-0.14	-0.03	-0.01	0.01	0.054**	0.140**	0.112**

COHORT 2	1997-1998	1998-1999	1999-2000	2000-2001	2001-2002	2002-2003	2003-2004	2000-2003	II/USP effect on 2000-2003 gap	2000-2004	II/USP effect on 2000-2004 gap
Elementary Reading Comparison	-0.14	-0.20	-0.27	-0.25	-0.25	-0.16	-0.12	0.114**		0.151**	
Elementary Reading II/USP	-0.13	-0.21	-0.24	-0.22	-0.17	-0.06	-0.06	0.183**	0.069**	0.175**	0.02
Middle Reading Comparison	-0.11	-0.13	-0.18	-0.17	-0.15	-0.07	-0.044	0.100**		0.136**	
Middle Reading II/USP	-0.11	-0.16	-0.24	-0.20	-0.20	-0.09	-0.08	0.147**	0.037**	0.163**	0.027**
High Reading Comparison	-0.09	-0.10	-0.13	-0.12	-0.12	-0.02	-0.031	0.112**		0.103**	
High Reading II/USP	-0.13	-0.15	-0.16	-0.16	-0.13	-0.04	-0.05	0.120**	0.01	0.108**	0.01

COHORT 3	1997-1998	1998-1999	1999-2000	2000-2001	2001-2002	2002-2003	2003-2004	2001-2004	II/USP effect on 2001-2004 gap
Elementary Reading Comparison	-0.16	-0.21	-0.26	-0.26	-0.25	-0.14	-0.11	0.144**	
Elementary Reading II/USP	-0.11	-0.17	-0.25	-0.23	-0.21	-0.12	-0.09	0.138**	-0.01
Middle Reading Comparison	-0.14	-0.18	-0.20	-0.19	-0.18	-0.09	-0.06	0.128**	
Middle Reading II/USP	-0.15	-0.20	-0.23	-0.22	-0.23	-0.10	-0.11	0.109**	-0.02
High Reading Comparison	-0.12	-0.13	-0.16	-0.14	-0.13	-0.05	-0.033	0.107**	
High Reading II/USP	-0.12	-0.15	-0.13	-0.13	-0.13	-0.02	0.00	0.126**	0.019*

*A positive II/USP effect indicates that II/USP closed the gap between FRLP students and non-FRLP students at a higher rate than comparison schools.

Effect Sizes for Subpopulations: Change in Achievement Gap Between ELs and EOs, Standardized SAT-9 and CAT/6 Math Scale Scores by School Type and Cohort

COHORT 1	1997-1998	1998-1999	1999-2000	2000-2001	2001-2002	2002-2003	2003-2004	1999-2002	II/USP effect on 1999-2002 gap	1999-2004	II/USP effect on 1999-2004 gap
Comparison Elementary Schools	-0.18	-0.18	-0.19	-0.21	-0.17	-0.13	-0.140	0.016**		0.045**	
II/USP Elementary Schools	-0.21	-0.19	-0.20	-0.19	-0.18	-0.13	-0.142	0.018*	0.00	0.052**	0.01
Comparison Middle Schools	-0.27	-0.29	-0.28	-0.29	-0.30	-0.38	-0.38	-0.01		-0.093**	
II/USP Middle Schools	-0.32	-0.32	-0.29	-0.29	-0.29	-0.33	-0.38	0.029**	0.039**	-0.057**	0.036**
Comparison High Schools	-0.30	-0.31	-0.32	-0.30	-0.31	-0.36	-0.407	0.00		-0.100**	
II/USP High Schools	-0.42	-0.36	-0.33	-0.33	-0.33	-0.31	-0.36	0.023*	0.02	0.00	0.103**

COHORT 2	1997-1998	1998-1999	1999-2000	2000-2001	2001-2002	2002-2003	2003-2004	2000-2003	II/USP effect on 2000-2003 gap	2000-2004	II/USP effect on 2000-2004 gap
Comparison Elementary Schools	-0.27	-0.23	-0.23	-0.23	-0.22	-0.16	-0.15	0.071**		0.082**	
II/USP Elementary Schools	-0.19	-0.17	-0.19	-0.18	-0.15	-0.11	-0.10	0.074**	0.00	0.087**	0.01
Comparison Middle Schools	-0.28	-0.29	-0.30	-0.32	-0.33	-0.35	-0.380	-0.056**		-0.084**	
II/USP Middle Schools	-0.33	-0.33	-0.30	-0.30	-0.31	-0.34	-0.36	-0.045**	0.01	-0.056**	0.028**
Comparison High Schools	-0.30	-0.33	-0.32	-0.32	-0.33	-0.37	-0.419	-0.055**		-0.099**	
II/USP High Schools	-0.36	-0.35	-0.31	-0.33	-0.33	-0.38	-0.40	-0.064**	-0.01	-0.089**	0.01

COHORT 3	1997-1998	1998-1999	1999-2000	2000-2001	2001-2002	2002-2003	2003-2004	2001-2004	II/USP effect on 2001-2004 gap
Comparison Elementary Schools	-0.24	-0.25	-0.22	-0.24	-0.22	-0.16	-0.17	0.072**	
II/USP Elementary Schools	-0.22	-0.20	-0.16	-0.19	-0.16	-0.14	-0.14	0.056**	-0.02
Comparison Middle Schools	-0.27	-0.29	-0.30	-0.31	-0.30	-0.35	-0.37	-0.063**	
II/USP Middle Schools	-0.29	-0.30	-0.31	-0.29	-0.30	-0.36	-0.38	-0.098**	-0.035**
Comparison High Schools	-0.35	-0.34	-0.34	-0.34	-0.35	-0.39	-0.416	-0.081**	
II/USP High Schools	-0.35	-0.32	-0.33	-0.33	-0.32	-0.38	-0.41	-0.085**	0.00

Effect Sizes for Subpopulations: Change in Achievement Gap Between Special Education and Non-Special Education Students, Standardized SAT-9 and CAT/6 Math Scale Scores by School Type and Cohort

COHORT 1	1997-1998	1998-1999	1999-2000	2000-2001	2001-2002	2002-2003	2003-2004	1999-2002	II/USP effect on 1999-2002 gap	1999-2004	II/USP effect on 1999-2004 gap
Elementary Reading Comparison	-0.62	-0.55	-0.58	-0.61	-0.65	-0.79	-0.794	-0.107**		-0.247**	
Elementary Reading II/USP	-0.63	-0.59	-0.61	-0.63	-0.65	-0.73	-0.755	-0.065**	0.043*	-0.169**	0.078**
Middle Reading Comparison	-0.51	-0.50	-0.52	-0.54	-0.58	-0.87	-0.86	-0.079**		-0.362**	
Middle Reading II/USP	-0.54	-0.60	-0.57	-0.63	-0.64	-0.88	-0.88	-0.035*	0.044*	-0.277**	0.085**
High Reading Comparison	-0.51	-0.49	-0.48	-0.48	-0.47	-0.78	-0.754	0.02		-0.265**	
High Reading II/USP	-0.50	-0.56	-0.57	-0.58	-0.58	-0.84	-0.87	-0.01	-0.03	-0.307**	-0.04

COHORT 2	1997-1998	1998-1999	1999-2000	2000-2001	2001-2002	2002-2003	2003-2004	2000-2003	II/USP effect on 2000-2003 gap	2000-2004	II/USP effect on 2000-2004 gap
Elementary Reading Comparison	-0.64	-0.60	-0.59	-0.64	-0.68	-0.81	-0.77	-0.217**		-0.181**	
Elementary Reading II/USP	-0.61	-0.55	-0.56	-0.59	-0.67	-0.79	-0.77	-0.230**	-0.01	-0.203**	-0.02
Middle Reading Comparison	-0.54	-0.55	-0.55	-0.59	-0.67	-1.02	-1.356	-0.466**		-0.803**	
Middle Reading II/USP	-0.56	-0.56	-0.55	-0.54	-0.57	-0.92	-1.20	-0.370**	0.096*	-0.648**	0.155**
High Reading Comparison	-0.56	-0.48	-0.53	-0.54	-0.54	-0.81	-0.801	-0.271**		-0.266**	
High Reading II/USP	-0.60	-0.53	-0.50	-0.55	-0.56	-0.83	-0.81	-0.327**	-0.056**	-0.306**	-0.040**

COHORT 3	1997-1998	1998-1999	1999-2000	2000-2001	2001-2002	2002-2003	2003-2004	2001-2004	II/USP effect on 2001-2004 gap
Elementary Reading Comparison	-0.67	-0.59	-0.61	-0.62	-0.67	-0.75	-0.74	-0.119**	
Elementary Reading II/USP	-0.65	-0.57	-0.61	-0.64	-0.72	-0.83	-0.81	-0.171**	-0.052**
Middle Reading Comparison	-0.57	-0.59	-0.59	-0.60	-0.63	-0.92	-0.88	-0.283**	
Middle Reading II/USP	-0.54	-0.52	-0.55	-0.56	-0.57	-0.85	-0.85	-0.285**	0.00
High Reading Comparison	-0.58	-0.57	-0.56	-0.54	-0.57	-0.83	-0.817	-0.273**	
High Reading II/USP	-0.54	-0.52	-0.52	-0.53	-0.53	-0.79	-0.79	-0.257**	0.02

Effect Sizes for Subpopulations: Change in Achievement Gap Between Free Lunch and Non-Free Lunch Students, Standardized SAT-9 and CAT/6 Math Scale Scores by School Type and Cohort

COHORT 1	1997-1998	1998-1999	1999-2000	2000-2001	2001-2002	2002-2003	2003-2004	1999-2002	II/USP effect on 1999-2002 gap	1999-2004	II/USP effect on 1999-2004 gap
Elementary Reading Comparison	-0.06	-0.14	-0.20	-0.21	-0.19	-0.12	-0.103	-0.053**		0.035**	
Elementary Reading II/USP	-0.08	-0.14	-0.22	-0.19	-0.19	-0.11	-0.092	-0.044**	0.01	0.0524**	0.02
Middle Reading Comparison	-0.04	-0.09	-0.13	-0.11	-0.09	-0.09	-0.06	-0.01		0.030**	
Middle Reading II/USP	-0.05	-0.07	-0.16	-0.14	-0.11	-0.10	-0.10	-0.042**	-0.037**	-0.023*	-0.053**
High Reading Comparison	0.00	0.00	-0.04	-0.03	-0.03	-0.06	-0.049	-0.025**		-0.044**	
High Reading II/USP	-0.05	-0.09	-0.06	-0.05	-0.06	-0.05	-0.02	0.029**	0.054**	0.074**	0.119**

COHORT 2	1997-1998	1998-1999	1999-2000	2000-2001	2001-2002	2002-2003	2003-2004	2000-2003	II/USP effect on 2000-2003 gap	2000-2004	II/USP effect on 2000-2004 gap
Elementary Reading Comparison	-0.10	-0.15	-0.23	-0.21	-0.21	-0.17	-0.13	0.057**		0.101**	
Elementary Reading II/USP	-0.07	-0.13	-0.19	-0.18	-0.14	-0.08	-0.07	0.117**	0.061**	0.127**	0.03
Middle Reading Comparison	-0.04	-0.08	-0.12	-0.12	-0.11	-0.09	-0.057	0.028**		0.061**	
Middle Reading II/USP	-0.04	-0.10	-0.18	-0.15	-0.15	-0.12	-0.10	0.062**	0.035**	0.082**	0.021*
High Reading Comparison	-0.02	-0.04	-0.06	-0.06	-0.06	-0.06	-0.058	0.00		0.00	
High Reading II/USP	-0.07	-0.07	-0.08	-0.07	-0.08	-0.06	-0.07	0.018*	0.018*	0.01	0.01

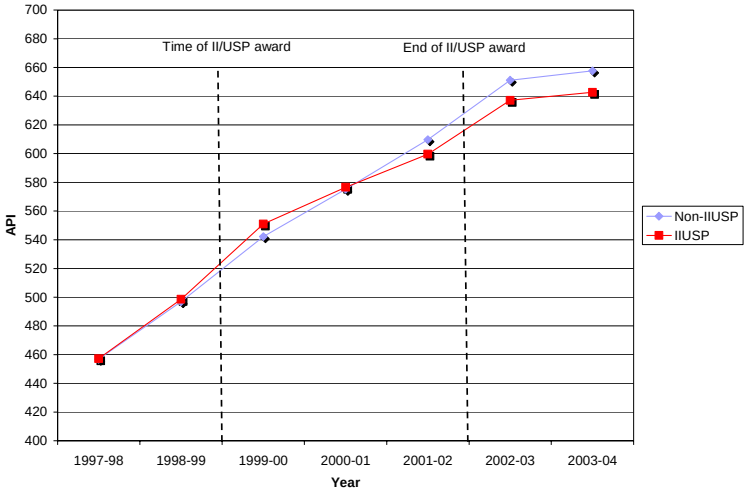
COHORT 3	1997-1998	1998-1999	1999-2000	2000-2001	2001-2002	2002-2003	2003-2004	2001-2004	II/USP effect on 2001-2004 gap
Elementary Reading Comparison	-0.11	-0.16	-0.22	-0.22	-0.20	-0.14	-0.12	0.103**	
Elementary Reading II/USP	-0.05	-0.13	-0.20	-0.19	-0.17	-0.12	-0.10	0.090**	-0.01
Middle Reading Comparison	-0.06	-0.11	-0.15	-0.14	-0.12	-0.11	-0.08	0.060**	
Middle Reading II/USP	-0.05	-0.14	-0.18	-0.16	-0.19	-0.13	-0.11	0.044**	-0.02
High Reading Comparison	-0.05	-0.06	-0.08	-0.08	-0.07	-0.08	-0.068	0.01	
High Reading II/USP	-0.03	-0.05	-0.04	-0.05	-0.04	-0.04	-0.03	0.015*	0.01

Appendix A-8:
HLM Results Elementary, Middle, and High Schools

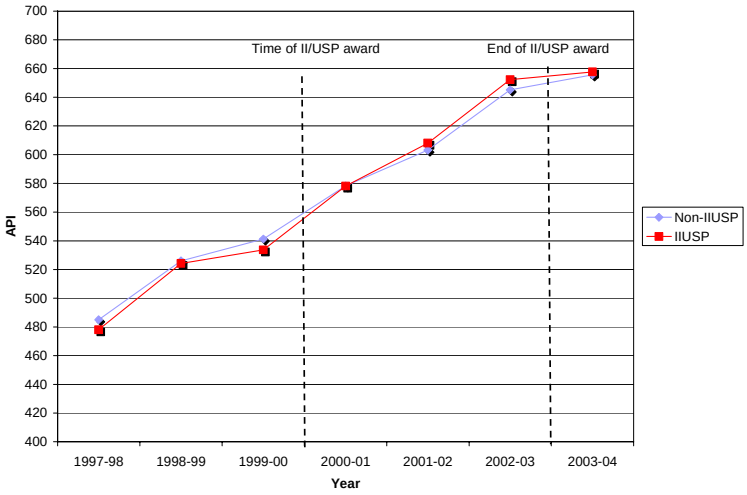
HLM Regression for Elementary Schools, API scores

	Cohort 1		Cohort 2		Cohort 3	
Variables	**Estimate**	**P value**	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	605.640	<.0001	584.600	<.0001	559.030	<.0001
IMP_PCT_AF_AM	-1.746	<.0001	-1.536	<.0001	-1.325	<.0001
IMP_PCT_ASIAN	-0.265	0.1286	0.075	0.7563	0.115	0.4804
IMP_PCT_HISP	-1.299	<.0001	-0.951	<.0001	-0.546	<.0001
IMP_PCT_MEALS	-0.507	<.0001	-0.368	<.0001	-0.372	<.0001
IMP_PCT_EL	-0.699	<.0001	-0.923	<.0001	-0.776	<.0001
IMP_MOBILITY	-0.066	0.1681	-0.067	0.3794	-0.165	0.0005
IMP_AVG_ED	3.424	0.0339	11.413	<.0001	10.894	<.0001
IMP_PCT_FULL	0.278	<.0001	0.120	0.1292	0.293	<.0001
IIUSP1	0.361	0.9329	-6.958	0.2045	-7.616	0.0301
CSRD1	-23.173	0.0078	-15.782	0.1261	14.104	0.3445
YEAR99	40.041	<.0001	40.988	<.0001	-73.957	<.0001
YEAR00	45.353	<.0001	15.237	<.0001	38.247	<.0001
YEAR01	33.022	<.0001	37.067	<.0001	47.324	<.0001
YEAR02	34.567	<.0001	24.879	<.0001	6.611	0.0001
YEAR03	41.391	<.0001	42.024	<.0001	24.867	<.0001
YEAR04	6.473	0.0009	10.433	0.0002	40.120	<.0001
IIUSP1*YEAR99	1.404	0.6508	5.186	0.1708	5.072	0.004
IIUSP1*YEAR00	7.026	0.0209	-5.781	0.1158	3.401	0.274
IIUSP1*YEAR01	-7.472	0.0175	7.467	0.0429	1.972	0.5177
IIUSP1*YEAR02	-11.436	0.0005	4.951	0.1879	-5.222	0.0875
IIUSP1*YEAR03	-3.563	0.2856	2.146	0.5747	0.199	0.9482
IIUSP1*YEAR04	-0.911	0.7842	-4.945	0.1979	0.499	0.8716
CSRD1*YEAR99	12.805	0.0425	-7.453	0.302	7.123	0.0218
CSRD1*YEAR00	-9.760	0.1132	-1.706	0.8126	8.902	0.4977
CSRD1*YEAR01	3.805	0.5426	0.019	0.9979	-13.540	0.2851
CSRD1*YEAR02	4.958	0.4356	4.708	0.5145	-26.172	0.046
CSRD1*YEAR03	2.844	0.6591	8.580	0.2388	28.625	0.0271
CSRD1*YEAR04	9.872	0.1266	-1.460	0.8410	9.255	0.4457

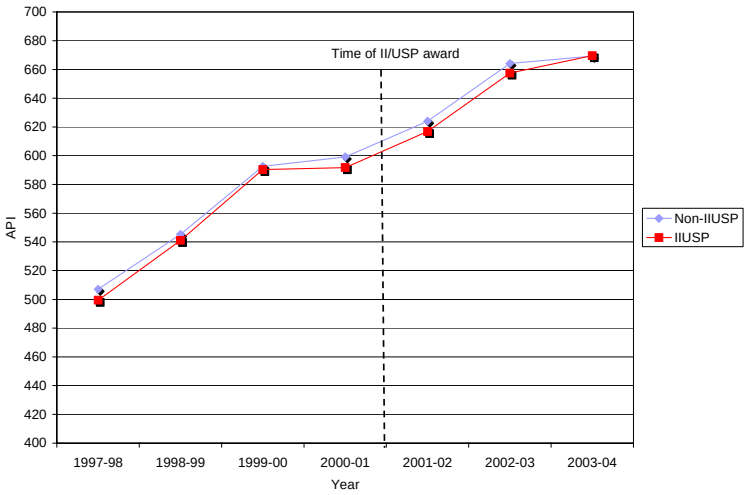
Relative API Performance, Cohort 1 Elementary Schools



Relative API Performance, Cohort 2 Elementary Schools



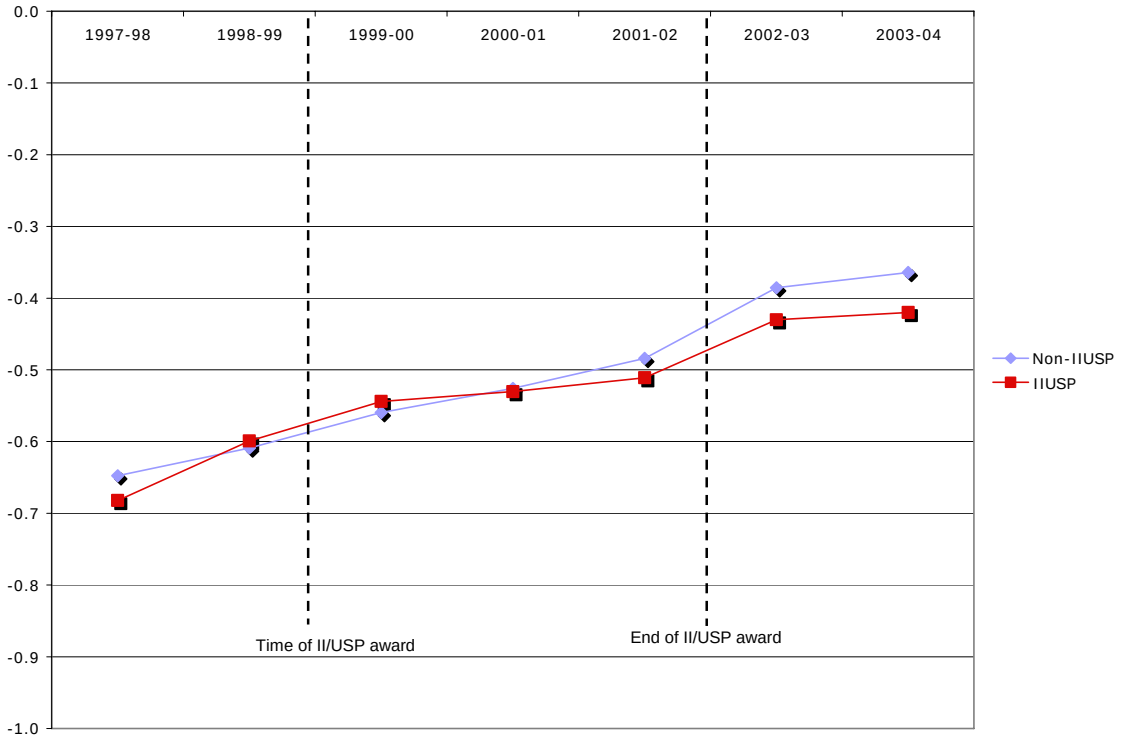
Relative API Performance, Cohort 3 Elementary Schools



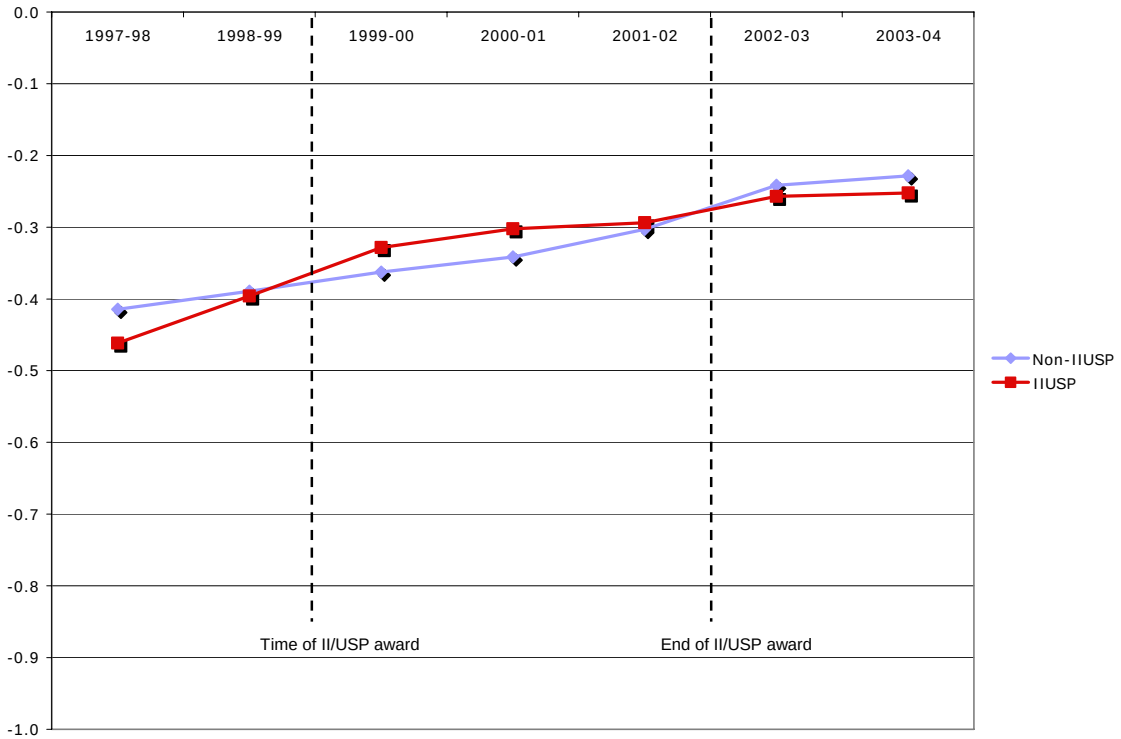
HLM Regression for Cohort 1 Elementary Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.112	<.0001	0.058	0.0216
GIRL	0.108	<.0001	-0.015	<.0001
ASIAN	-0.007	0.0295	0.196	<.0001
HISPANIC	-0.258	<.0001	-0.235	<.0001
BLACK	-0.406	<.0001	-0.454	<.0001
OTHERS	-0.091	<.0001	-0.052	<.0001
ELMISN	-0.119	<.0001	-0.082	<.0001
EL	-0.374	<.0001	-0.174	<.0001
R_FEP	0.345	<.0001	0.484	<.0001
FEP	0.185	<.0001	0.231	<.0001
FLUNCH	-0.175	<.0001	-0.143	<.0001
HIGHPARED	0.222	<.0001	0.203	<.0001
PARED_MISSING	-0.012	<.0001	-0.034	<.0001
SPECED	-0.592	<.0001	-0.657	<.0001
YEAR99	0.039	<.0001	0.025	<.0001
YEAR00	0.049	<.0001	0.027	<.0001
YEAR01	0.034	<.0001	0.021	<.0001
YEAR02	0.042	<.0001	0.039	<.0001
YEAR03	0.099	<.0001	0.061	<.0001
YEAR04	0.021	<.0001	0.013	<.0001
GRADE3	-0.031	<.0001	-0.023	<.0001
GRADE4	-0.062	<.0001	-0.060	<.0001
GRADE5	-0.089	<.0001	-0.101	<.0001
PCT_MEALS	-0.185	<.0001	-0.092	0.0046
IIUSP1	-0.035	0.0032	-0.047	0.0004
CSR1	-0.052	0.0282	0.000	0.9947
YEAR99*IIUSP1	0.044	<.0001	0.041	<.0001
YEAR00*IIUSP1	0.006	0.2135	0.041	<.0001
YEAR01*IIUSP1	-0.020	<.0001	0.005	0.2719
YEAR02*IIUSP1	-0.022	<.0001	-0.030	<.0001
YEAR03*IIUSP1	-0.018	<.0001	-0.025	<.0001
YEAR04*IIUSP1	-0.011	0.0121	-0.008	0.0715
YEAR99*CSR1	-0.012	0.2416	-0.023	0.0254
YEAR00*CSR1	-0.013	0.1502	-0.039	<.0001
YEAR01*CSR1	0.017	0.0388	-0.018	0.0378
YEAR02*CSR1	-0.011	0.1609	0.010	0.2433
YEAR03*CSR1	0.063	<.0001	0.047	<.0001
YEAR04*CSR1	0.001	0.9010	0.018	0.0361

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores), Cohort 1 Elementary Schools



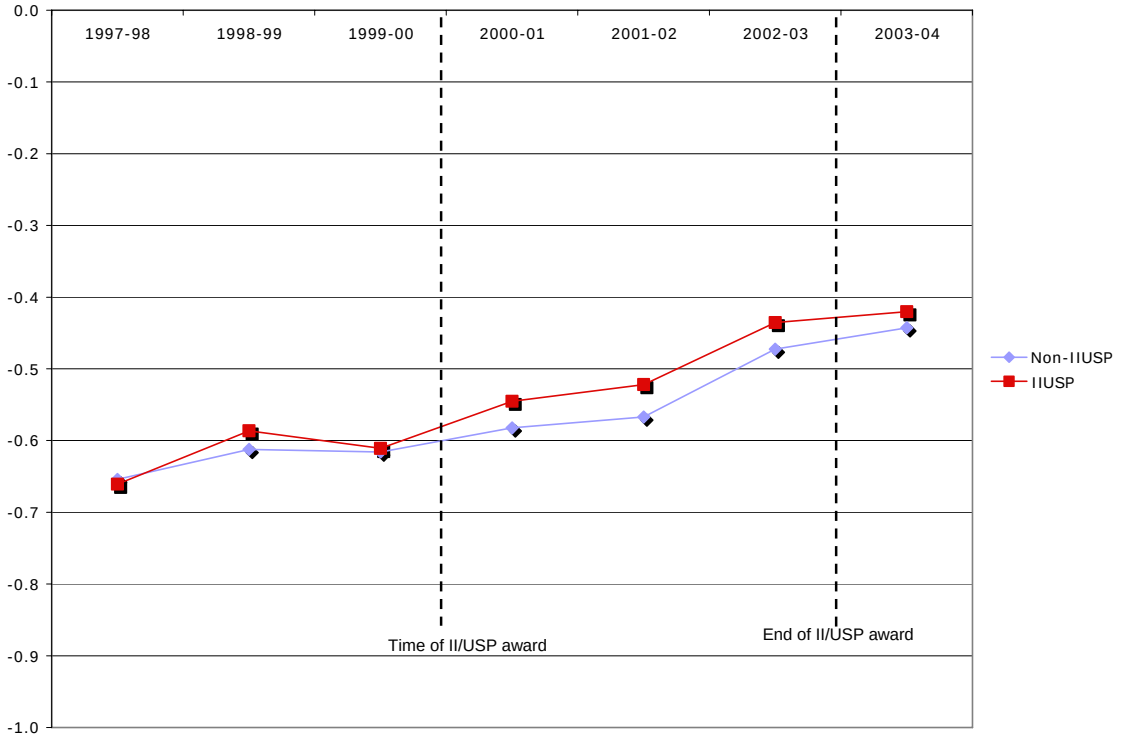
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores), Cohort 1 Elementary Schools



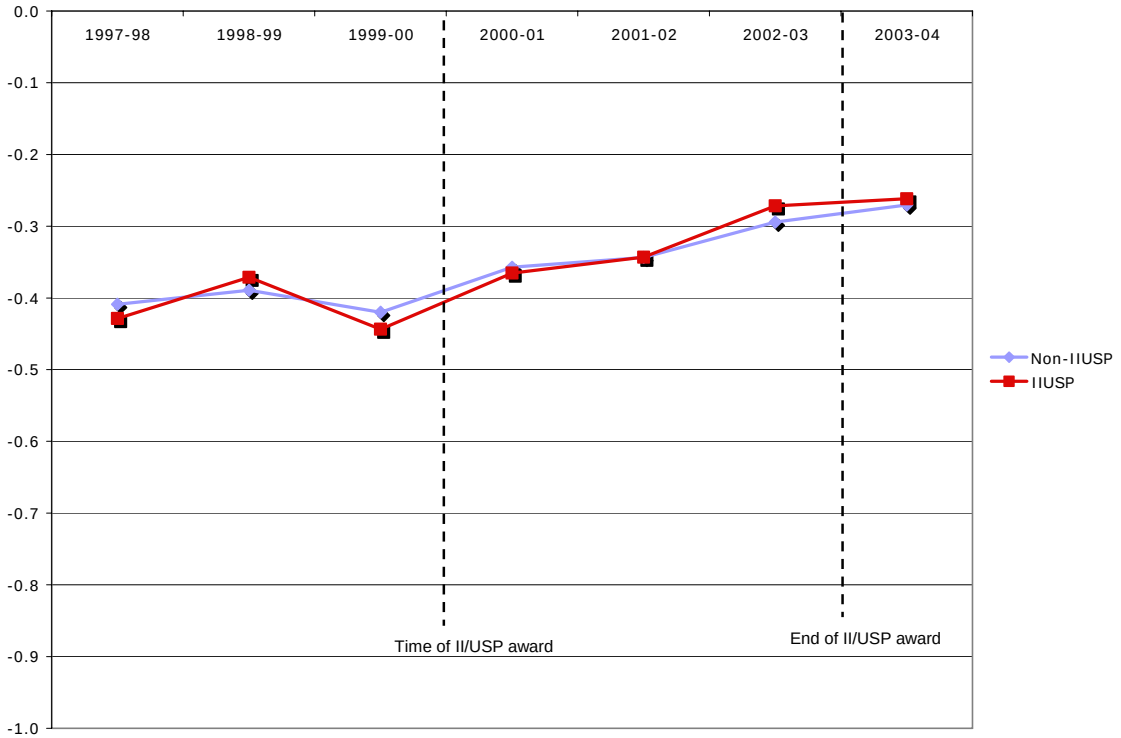
HLM Regression for Cohort 2 Elementary Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.160	<.0001	0.111	0.0007
GIRL	0.114	<.0001	-0.011	<.0001
ASIAN	-0.030	<.0001	0.164	<.0001
HISPANIC	-0.267	<.0001	-0.240	<.0001
BLACK	-0.419	<.0001	-0.467	<.0001
OTHERS	-0.124	<.0001	-0.077	<.0001
ELMISN	-0.161	<.0001	-0.106	<.0001
EL	-0.371	<.0001	-0.181	<.0001
R_FEP	0.364	<.0001	0.500	<.0001
FEP	0.197	<.0001	0.241	<.0001
FLUNCH	-0.178	<.0001	-0.146	<.0001
HIGHPARED	0.237	<.0001	0.218	<.0001
PARED_MISSING	-0.016	<.0001	-0.035	<.0001
SPECED	-0.613	<.0001	-0.667	<.0001
YEAR99	0.042	<.0001	0.020	<.0001
YEAR00	-0.004	0.3691	-0.031	<.0001
YEAR01	0.034	<.0001	0.063	<.0001
YEAR02	0.015	0.0002	0.014	0.0008
YEAR03	0.095	<.0001	0.049	<.0001
YEAR04	0.030	<.0001	0.024	<.0001
GRADE3	-0.028	<.0001	-0.039	<.0001
GRADE4	-0.057	<.0001	-0.070	<.0001
GRADE5	-0.074	<.0001	-0.094	<.0001
PCT_MEALS	-0.266	<.0001	-0.150	0.0005
IIUSP2	-0.007	0.6112	-0.020	0.2019
CSR2	0.009	0.7240	-0.024	0.433
YEAR99*IIUSP2	0.033	<.0001	0.038	<.0001
YEAR00*IIUSP2	-0.021	0.0005	-0.041	<.0001
YEAR01*IIUSP2	0.032	<.0001	0.015	0.0053
YEAR02*IIUSP2	0.008	0.1357	0.008	0.1392
YEAR03*IIUSP2	-0.008	0.1479	0.023	<.0001
YEAR04*IIUSP2	-0.015	0.0084	-0.014	0.0134
YEAR99*CSR2	-0.068	<.0001	-0.053	<.0001
YEAR00*CSR2	0.034	0.0010	0.003	0.7906
YEAR01*CSR2	-0.022	0.0233	-0.027	0.0051
YEAR02*CSR2	0.028	0.0031	0.037	0.0001
YEAR03*CSR2	0.044	<.0001	0.051	<.0001
YEAR04*CSR2	0.008	0.3685	0.013	0.1866

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores), Cohort 2 Elementary Schools



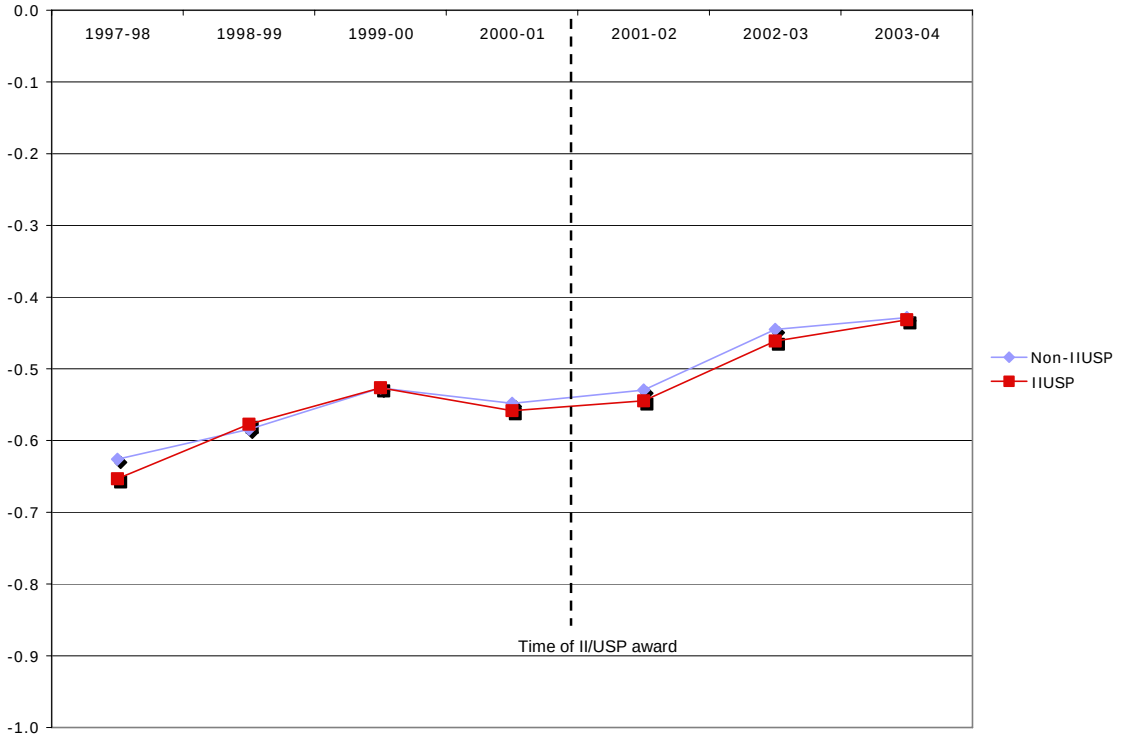
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores), Cohort 2 Elementary Schools



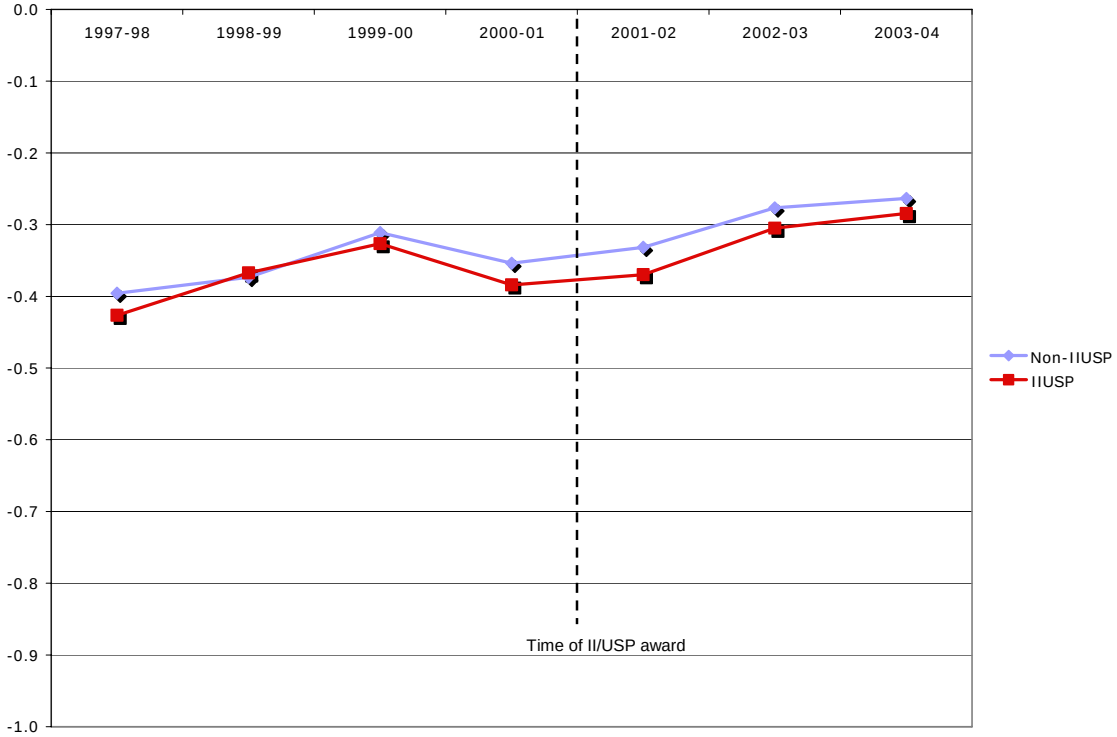
HLM Regression for Cohort 3 Elementary Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.168	<.0001	0.105	<.0001
GIRL	0.118	<.0001	-0.014	<.0001
ASIAN	-0.009	0.0099	0.184	<.0001
HISPANIC	-0.254	<.0001	-0.230	<.0001
BLACK	-0.410	<.0001	-0.449	<.0001
OTHERS	-0.100	<.0001	-0.055	<.0001
ELMISN	-0.135	<.0001	-0.098	<.0001
EL	-0.391	<.0001	-0.196	<.0001
R_FEP	0.319	<.0001	0.451	<.0001
FEP	0.164	<.0001	0.217	<.0001
FLUNCH	-0.188	<.0001	-0.158	<.0001
HIGHPARED	0.239	<.0001	0.220	<.0001
PARED_MISSING	-0.003	0.0849	-0.031	<.0001
SPECED	-0.639	<.0001	-0.680	<.0001
YEAR99	0.042	<.0001	0.022	<.0001
YEAR00	0.057	<.0001	0.062	<.0001
YEAR01	-0.021	<.0001	-0.043	<.0001
YEAR02	0.018	<.0001	0.022	<.0001
YEAR03	0.085	<.0001	0.055	<.0001
YEAR04	0.016	<.0001	0.013	<.0001
GRADE3	-0.018	<.0001	-0.024	<.0001
GRADE4	-0.044	<.0001	-0.052	<.0001
GRADE5	-0.060	<.0001	-0.079	<.0001
PCT_MEALS	-0.249	<.0001	-0.124	<.0001
IIUSP3	-0.027	0.0076	-0.031	0.0099
CSR3	-0.036	0.4156	-0.031	0.5419
YEAR99*IIUSP3	0.034	<.0001	0.037	<.0001
YEAR00*IIUSP3	-0.006	0.2241	-0.022	<.0001
YEAR01*IIUSP3	-0.011	0.0213	-0.015	0.0024
YEAR02*IIUSP3	-0.005	0.3313	-0.008	0.1145
YEAR03*IIUSP3	-0.001	0.7696	0.010	0.0476
YEAR04*IIUSP3	0.013	0.0044	0.007	0.1307
YEAR99*CSR3	0.061	0.0052	0.045	0.0396
YEAR00*CSR3	-0.074	0.0002	-0.056	0.0049
YEAR01*CSR3	-0.120	<.0001	-0.101	<.0001
YEAR02*CSR3	0.062	0.0006	0.109	<.0001
YEAR03*CSR3	0.060	0.0007	0.048	0.0086
YEAR04*CSR3	-0.015	0.4056	-0.013	0.4943
YEAR99*IIUSP3*HP	-0.096	<.0001	-0.073	<.0001
YEAR00*IIUSP3*HP	0.035	0.0002	0.039	<.0001
YEAR01*IIUSP3*HP	0.010	0.242	-0.005	0.5598
YEAR02*IIUSP3*HP	0.032	<.0001	0.027	0.0012
YEAR03*IIUSP3*HP	0.064	<.0001	0.048	<.0001
YEAR04*IIUSP3*HP	-0.002	0.8319	-0.002	0.7746

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores), Cohort 3 Elementary Schools



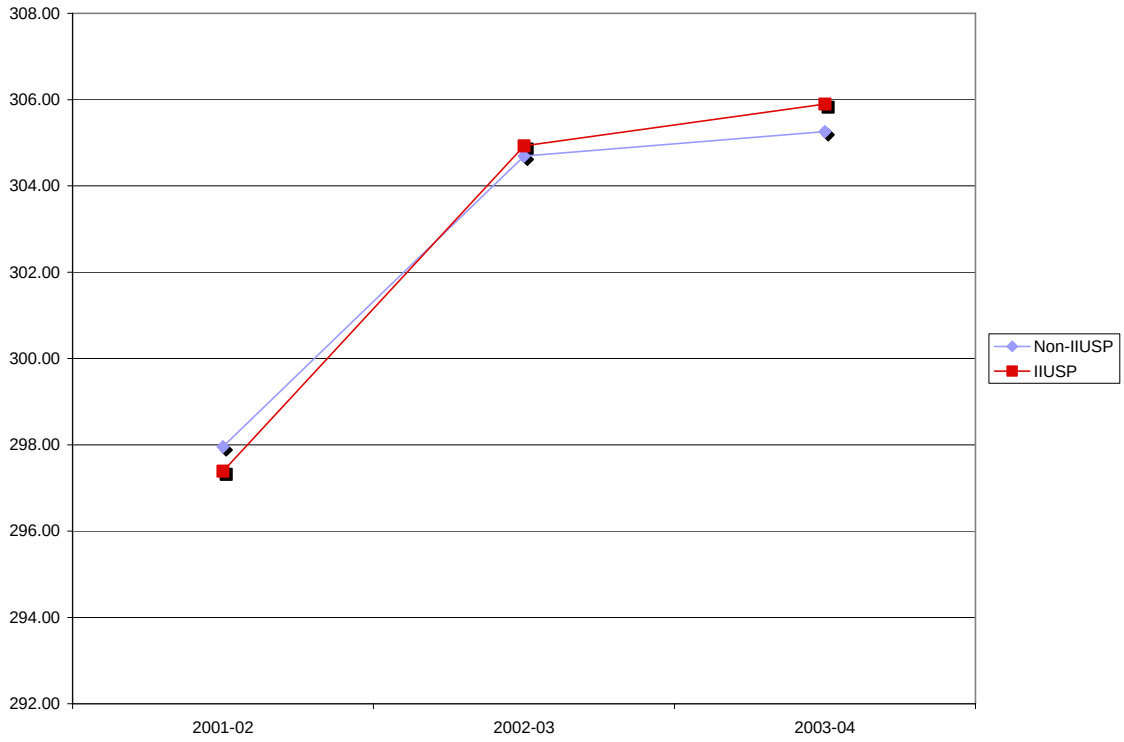
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores), Cohort 3 Elementary Schools



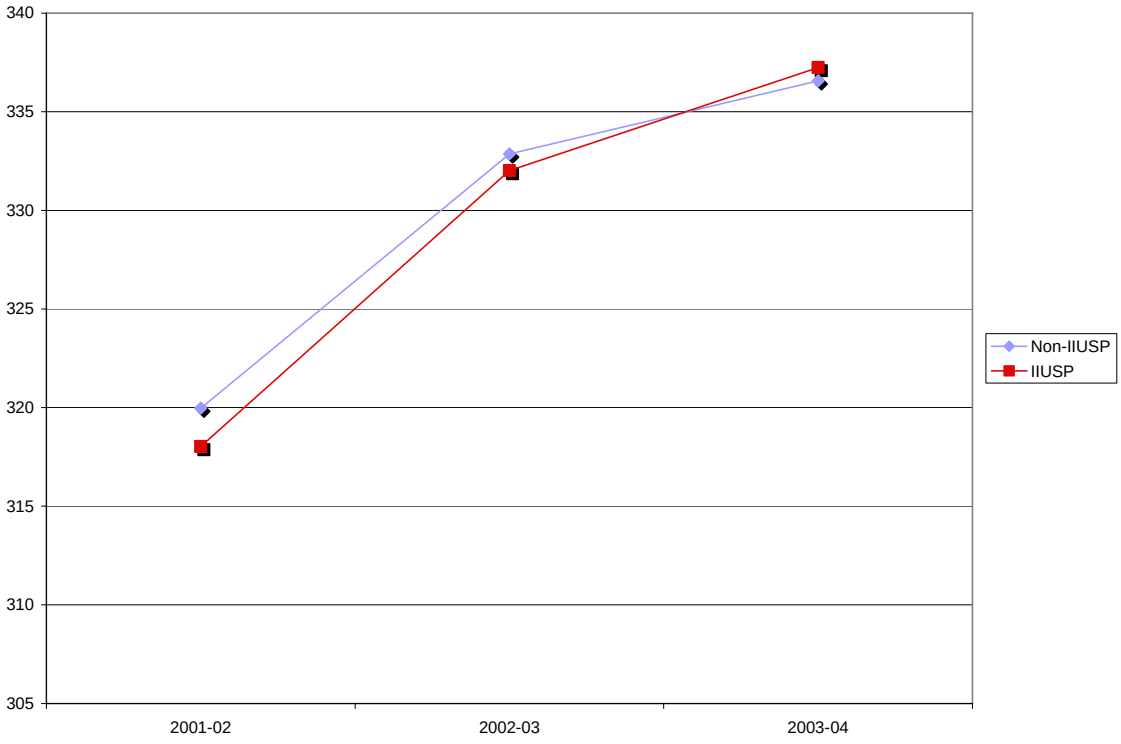
HLM Regression for Cohort 3 Elementary Schools, CST ELA and Math Scores

	ELA		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	334.640	<.0001	351.170	<.0001
GIRL	7.786	<.0001	-2.906	<.0001
ASIAN	5.722	<.0001	14.571	<.0001
HISPANIC	-12.023	<.0001	-15.987	<.0001
BLACK	-19.698	<.0001	-30.142	<.0001
OTHERS	-1.431	<.0001	-2.161	<.0001
ELMISN	-11.662	<.0001	-12.704	<.0001
EL	-18.168	<.0001	-12.245	<.0001
R_FEP	22.052	<.0001	33.237	<.0001
FEP	12.587	<.0001	18.032	<.0001
FLUNCH	-11.924	<.0001	-12.858	<.0001
HIGHPARED	11.492	<.0001	13.516	<.0001
PARED_MISSING	-1.586	<.0001	-3.370	<.0001
SPECED	-32.069	<.0001	-42.553	<.0001
YEAR03	6.733	<.0001	12.887	<.0001
YEAR04	0.569	<.0001	3.708	<.0001
GRADE3	-8.781	<.0001	-9.935	<.0001
GRADE4	5.723	<.0001	-11.748	<.0001
GRADE5	1.143	<.0001	-26.564	<.0001
PCT_MEALS	-10.828	<.0001	-4.947	0.0755
IIUSP3	-0.566	0.3826	-1.943	0.0547
CSR3	-4.711	0.0904	-1.826	0.6734
YEAR03*IIUSP3	0.804	0.0013	1.095	0.0012
YEAR04*IIUSP3	0.398	0.1092	1.523	<.0001
YEAR03*CSR3	2.929	0.002	2.847	0.0261
YEAR04*CSR3	0.956	0.3041	4.145	0.0012
YEAR03*IIUSP3*HP	0.249	0.5544	0.081	0.8871
YEAR04*IIUSP3*HP	-1.263	0.0029	-0.520	0.3706

Relative CST ELA Performance, Cohort 3 Elementary Schools



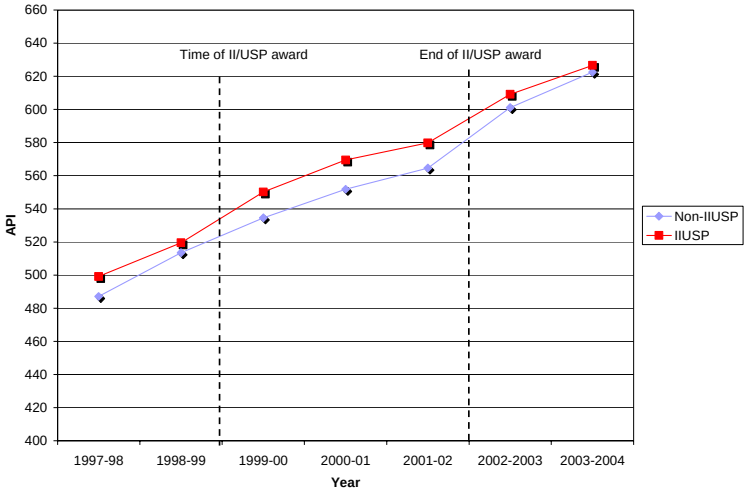
Relative CST Math Performance, Cohort 3 Elementary Schools



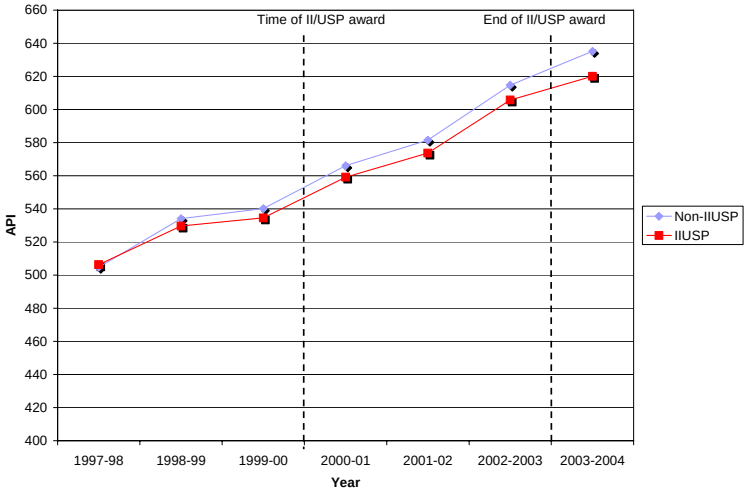
HLM Regression for Middle Schools, API scores

	Cohort 1		Cohort 2		Cohort 3	
Variables	**Estimate**	**P value**	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	597.990	<.0001	632.810	<.0001	605.860	<.0001
IMP_PCT_AF_AM	-2.612	<.0001	-2.614	<.0001	-1.701	<.0001
IMP_PCT_ASIAN	-0.881	0.0022	-1.050	0.0003	-0.465	0.0712
IMP_PCT_HISP	-1.907	<.0001	-2.059	<.0001	-1.076	<.0001
IMP_PCT_MEALS	-0.233	0.0055	-0.234	0.0024	-0.231	0.002
IMP_PCT_EL	-0.244	0.0401	-0.458	<.0001	-0.636	<.0001
IMP_MOBILITY	-0.123	0.0336	0.123	0.02	-0.054	0.2336
IMP_AVG_ED	20.578	<.0001	10.403	0.0101	7.399	0.0174
IMP_PCT_FULL	0.159	0.105	0.385	<.0001	0.253	0.0039
IIUSP1	12.054	0.0868	1.475	0.8226	-18.802	0.0014
CSRD1	-11.304	0.4849	-3.637	0.8188	0.573	0.9872
YEAR99	26.401	<.0001	29.324	<.0001	-67.326	<.0001
YEAR00	20.982	<.0001	6.074	0.0216	23.027	<.0001
YEAR01	17.271	<.0001	25.845	<.0001	25.271	<.0001
YEAR02	12.823	<.0001	15.496	<.0001	3.327	0.1251
YEAR03	36.471	<.0001	33.214	<.0001	11.638	<.0001
YEAR04	21.368	<.0001	20.357	<.0001	29.974	<.0001
IIUSP1*YEAR99	-6.118	0.1863	-5.992	0.1494	19.480	<.0001
IIUSP1*YEAR00	9.721	0.0332	-1.129	0.7794	10.161	0.0467
IIUSP1*YEAR01	1.987	0.677	-1.229	0.7606	8.140	0.1021
IIUSP1*YEAR02	-2.420	0.6325	-0.991	0.8119	-1.305	0.7927
IIUSP1*YEAR03	-7.156	0.1696	-1.142	0.7881	-3.925	0.4419
IIUSP1*YEAR04	-3.896	0.4571	-6.003	0.1583	0.175	0.973
CSRD1*YEAR99	14.009	0.1961	-3.011	0.7632	-2.444	0.6323
CSRD1*YEAR00	-16.403	0.1192	1.355	0.8887	22.083	0.4369
CSRD1*YEAR01	-12.432	0.2528	16.937	0.0813	-1.410	0.9604
CSRD1*YEAR02	0.371	0.9732	-8.493	0.3818	-22.993	0.4203
CSRD1*YEAR03	1.821	0.8693	-3.860	0.6914	19.666	0.4890
CSRD1*YEAR04	-0.037	0.9973	-2.755	0.7768	6.039	0.8319

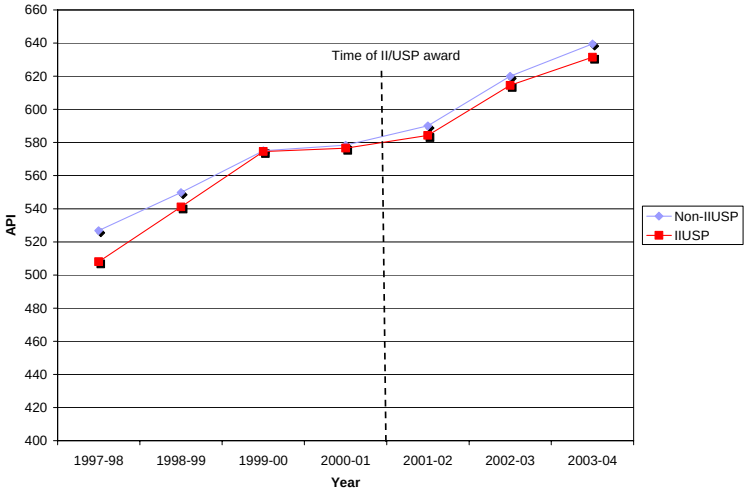
Relative API Performance, Cohort 1 Middle Schools



Relative API Performance, Cohort 2 Middle Schools



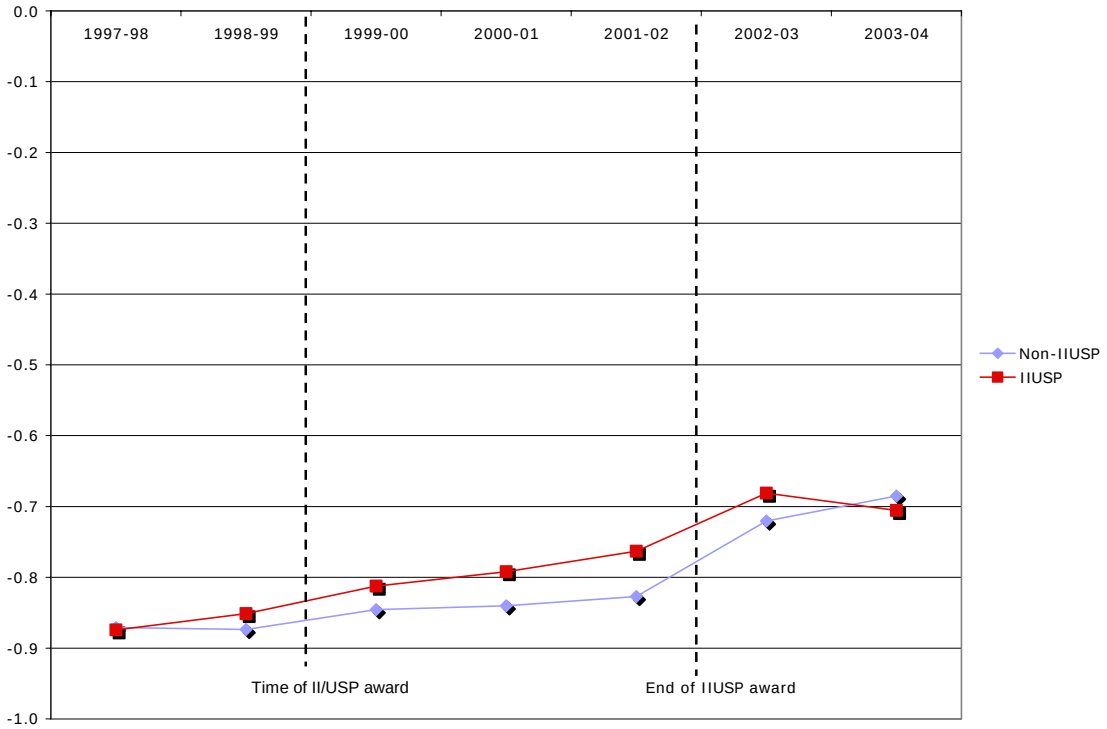
Relative API Performance, Cohort 3 Middle Schools



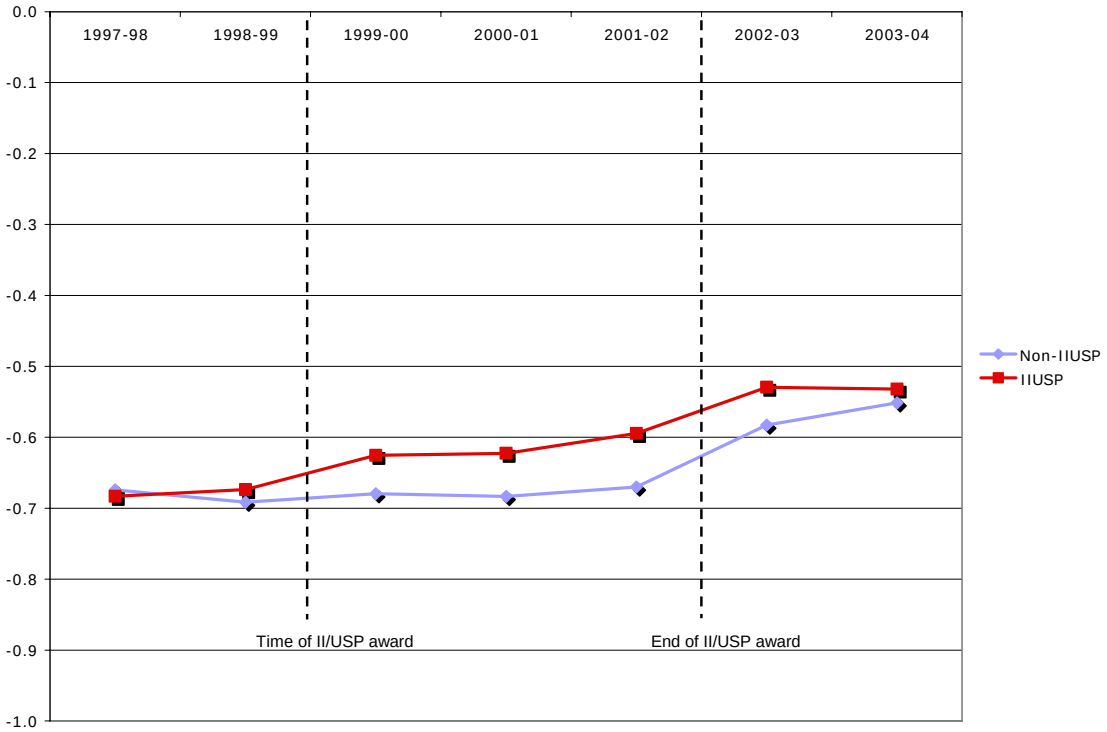
HLM Regression for Cohort 1 Middle Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.218	<.0001	0.069	0.0387
GIRL	0.103	<.0001	-0.007	<.0001
ASIAN	-0.065	<.0001	0.262	<.0001
HISPANIC	-0.338	<.0001	-0.315	<.0001
BLACK	-0.518	<.0001	-0.510	<.0001
OTHERS	-0.185	<.0001	-0.096	<.0001
ELMISN	-0.184	<.0001	-0.071	<.0001
EL	-0.564	<.0001	-0.311	<.0001
R_FEP	0.214	<.0001	0.283	<.0001
FEP	0.167	<.0001	0.183	<.0001
FLUNCH	-0.129	<.0001	-0.096	<.0001
HIGHPARED	0.217	<.0001	0.181	<.0001
PARED_MISSING	-0.036	<.0001	-0.036	<.0001
SPECED	-0.729	<.0001	-0.636	<.0001
YEAR99	-0.003	0.3466	-0.017	<.0001
YEAR00	0.028	<.0001	0.012	<.0001
YEAR01	0.005	0.086	-0.004	0.1706
YEAR02	0.013	<.0001	0.014	<.0001
YEAR03	0.107	<.0001	0.087	<.0001
YEAR04	0.035	<.0001	0.031	<.0001
GRADE7	-0.010	<.0001	0.020	<.0001
GRADE8	-0.026	<.0001	-0.003	0.1513
PCT_MEALS	-0.320	<.0001	-0.202	<.0001
IIUSP1	-0.003	0.8612	-0.009	0.6455
CSR1	0.088	0.0357	0.147	0.0007
YEAR99*IIUSP1	0.026	0.0001	0.027	<.0001
YEAR00*IIUSP1	0.010	0.0868	0.036	<.0001
YEAR01*IIUSP1	0.016	0.0071	0.007	0.2157
YEAR02*IIUSP1	0.015	0.0101	0.015	0.0103
YEAR03*IIUSP1	-0.024	<.0001	-0.022	<.0001
YEAR04*IIUSP1	-0.060	<.0001	-0.034	<.0001
YEAR99*CSR1	-0.088	<.0001	-0.088	<.0001
YEAR00*CSR1	0.012	0.3451	-0.043	0.0007
YEAR01*CSR1	-0.010	0.4122	-0.050	<.0001
YEAR02*CSR1	-0.042	0.0002	-0.038	0.0004
YEAR03*CSR1	0.006	0.5775	0.010	0.3437
YEAR04*CSR1	0.079	<.0001	0.038	0.0002

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores), Cohort 1 Middle Schools



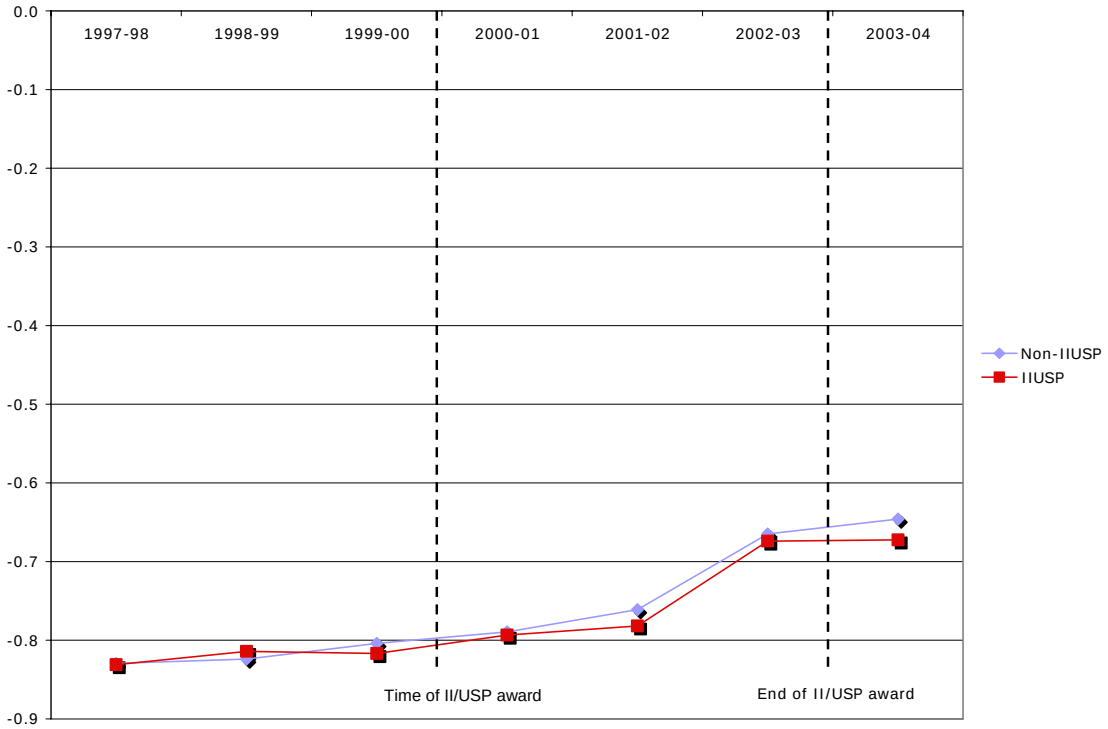
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores), Cohort 1 Middle Schools



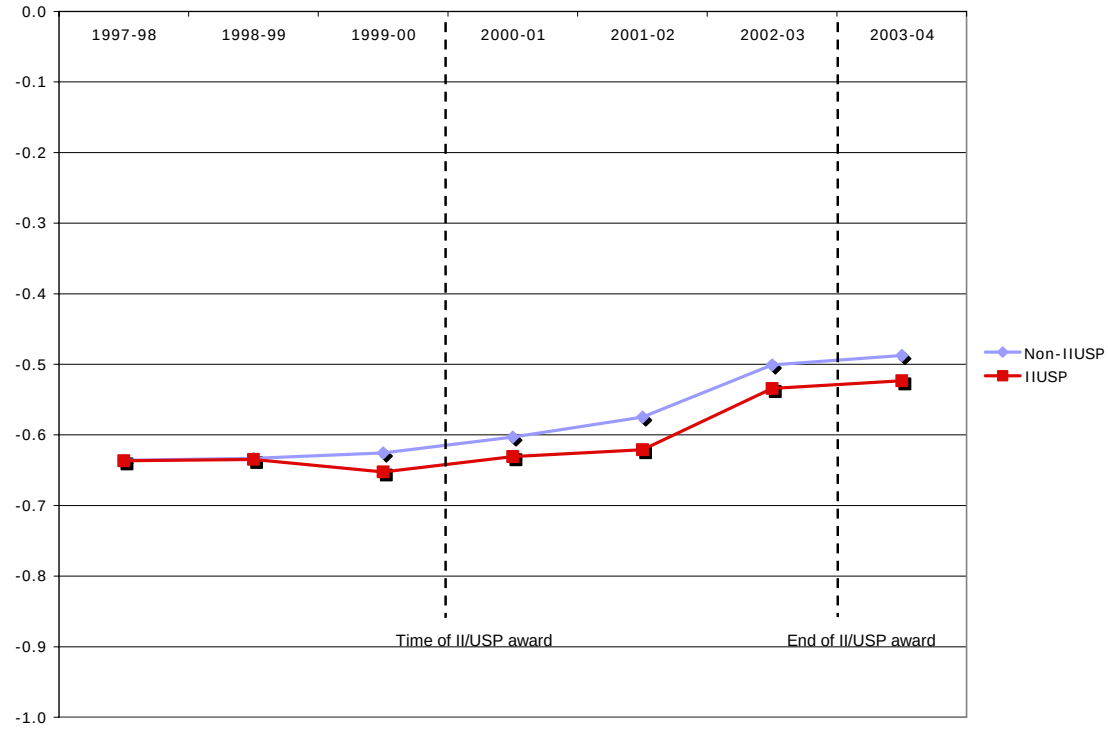
HLM Regression for Cohort 2 Middle Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.200	<.0001	0.140	<.0001
GIRL	0.102	<.0001	-0.012	<.0001
ASIAN	-0.016	<.0001	0.268	<.0001
HISPANIC	-0.298	<.0001	-0.292	<.0001
BLACK	-0.489	<.0001	-0.491	<.0001
OTHERS	-0.151	<.0001	-0.077	<.0001
ELMISN	-0.161	<.0001	-0.075	<.0001
EL	-0.566	<.0001	-0.323	<.0001
R_FEP	0.218	<.0001	0.299	<.0001
FEP	0.182	<.0001	0.217	<.0001
FLUNCH	-0.139	<.0001	-0.104	<.0001
HIGHPARED	0.220	<.0001	0.189	<.0001
PARED_MISSING	-0.027	<.0001	-0.032	<.0001
SPECED	-0.747	<.0001	-0.667	<.0001
YEAR99	0.006	0.1149	0.002	0.4749
YEAR00	0.020	<.0001	0.008	0.0144
YEAR01	0.015	<.0001	0.022	<.0001
YEAR02	0.028	<.0001	0.029	<.0001
YEAR03	0.096	<.0001	0.074	<.0001
YEAR04	0.019	<.0001	0.013	<.0001
GRADE7	-0.007	<.0001	0.007	<.0001
GRADE8	-0.023	<.0001	-0.015	<.0001
PCT_MEALS	-0.302	<.0001	-0.292	<.0001
IIUSP2	-0.002	0.9211	-0.001	0.9687
CSR2	0.063	0.1059	0.131	0.0031
YEAR99*IIUSP2	0.011	0.0390	-0.001	0.8981
YEAR00*IIUSP2	-0.022	<.0001	-0.025	<.0001
YEAR01*IIUSP2	0.009	0.0703	-0.001	0.8432
YEAR02*IIUSP2	-0.016	0.0004	-0.019	<.0001
YEAR03*IIUSP2	0.012	0.0121	0.013	0.0042
YEAR04*IIUSP2	-0.018	0.0001	-0.002	0.6111
YEAR99*CSR2	-0.034	0.0075	-0.084	<.0001
YEAR00*CSR2	-0.028	0.0151	-0.045	<.0001
YEAR01*CSR2	0.049	<.0001	0.026	0.0092
YEAR02*CSR2	-0.034	0.0009	-0.015	0.1342
YEAR03*CSR2	-0.055	<.0001	-0.034	0.0004
YEAR04*CSR2	0.036	0.0002	0.033	0.0007

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores), Cohort 2 Middle Schools



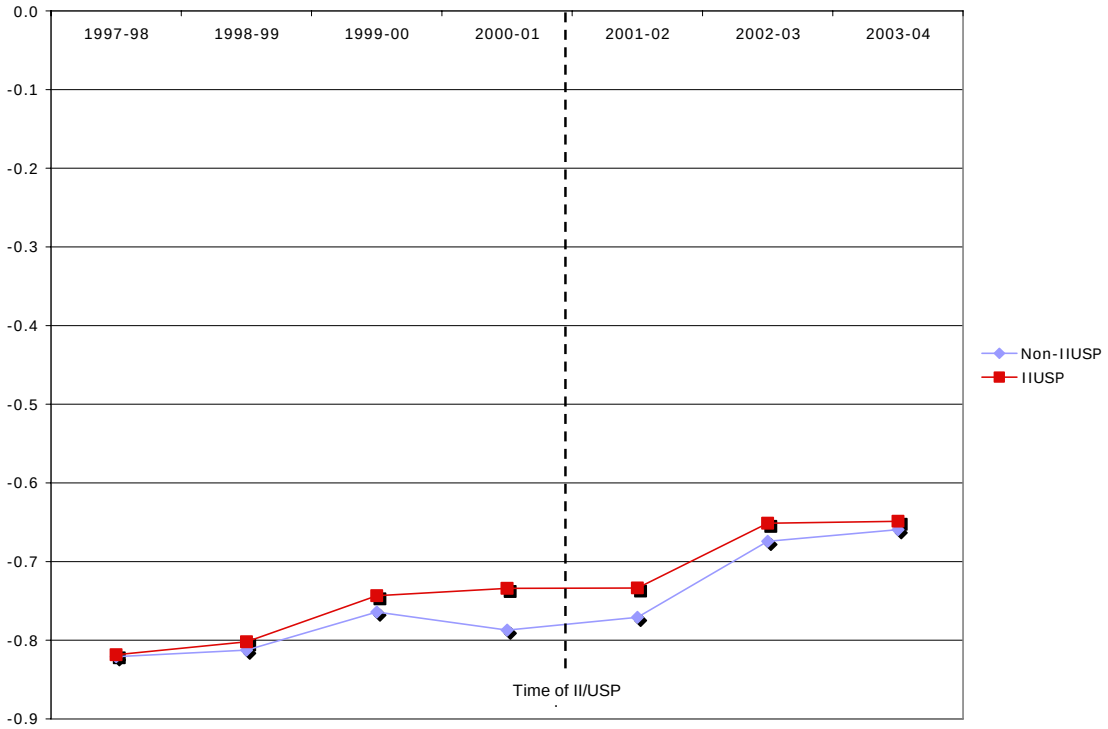
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores), Cohort 2 Middle Schools



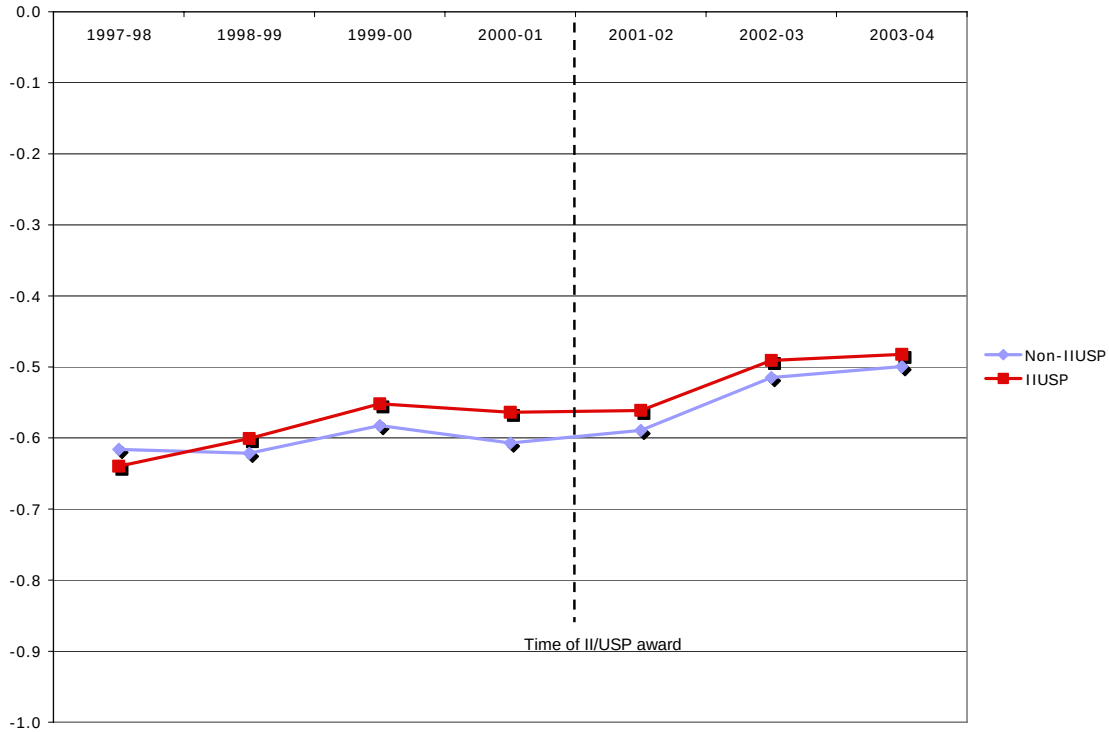
HLM Regression for Cohort 3 Middle Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.237	<.0001	0.149	<.0001
GIRL	0.104	<.0001	-0.014	<.0001
ASIAN	-0.066	<.0001	0.228	<.0001
HISPANIC	-0.317	<.0001	-0.292	<.0001
BLACK	-0.483	<.0001	-0.484	<.0001
OTHERS	-0.165	<.0001	-0.072	<.0001
ELMISN	-0.153	<.0001	-0.104	<.0001
EL	-0.560	<.0001	-0.320	<.0001
R_FEP	0.226	<.0001	0.293	<.0001
FEP	0.167	<.0001	0.180	<.0001
FLUNCH	-0.154	<.0001	-0.120	<.0001
HIGHPARED	0.214	<.0001	0.183	<.0001
PARED_MISSING	-0.030	<.0001	-0.030	<.0001
SPECED	-0.766	<.0001	-0.691	<.0001
YEAR99	0.008	0.008	-0.006	0.0683
YEAR00	0.048	<.0001	0.039	<.0001
YEAR01	-0.023	<.0001	-0.024	<.0001
YEAR02	0.016	<.0001	0.018	<.0001
YEAR03	0.097	<.0001	0.074	<.0001
YEAR04	0.015	<.0001	0.016	<.0001
GRADE7	-0.006	0.0002	0.006	0.0001
GRADE8	-0.023	<.0001	-0.017	<.0001
PCT_MEALS	-0.331	<.0001	-0.280	<.0001
IIUSP3	0.003	0.8808	-0.024	0.2109
CSR3	0.030	0.8315	0.072	0.629
YEAR99*IIUSP3	0.008	0.2881	0.044	<.0001
YEAR00*IIUSP3	0.010	0.1503	0.010	0.1439
YEAR01*IIUSP3	0.033	<.0001	0.013	0.0387
YEAR02*IIUSP3	-0.016	0.0101	-0.015	0.0114
YEAR03*IIUSP3	-0.014	0.0197	-0.004	0.5244
YEAR04*IIUSP3	-0.012	0.0445	-0.007	0.2262
YEAR99*CSR3	0.026	0.8022	-0.005	0.9578
YEAR00*CSR3	-0.102	0.0568	-0.101	0.0512
YEAR01*CSR3	-0.056	0.2433	0.054	0.2533
YEAR02*CSR3	-0.079	0.0947	0.057	0.2149
YEAR03*CSR3	0.069	0.1462	-0.163	0.0004
YEAR04*CSR3	0.055	0.2543	-0.034	0.4652
YEAR99*IIUSP3*HP	-0.062	<.0001	-0.083	<.0001
YEAR00*IIUSP3*HP	-0.034	0.0036	-0.062	<.0001
YEAR01*IIUSP3*HP	-0.050	<.0001	-0.047	<.0001
YEAR02*IIUSP3*HP	0.026	0.0121	0.040	<.0001
YEAR03*IIUSP3*HP	0.079	<.0001	0.011	0.2514
YEAR04*IIUSP3*HP	0.033	0.0011	0.037	0.0002

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores), Cohort 3 Middle Schools



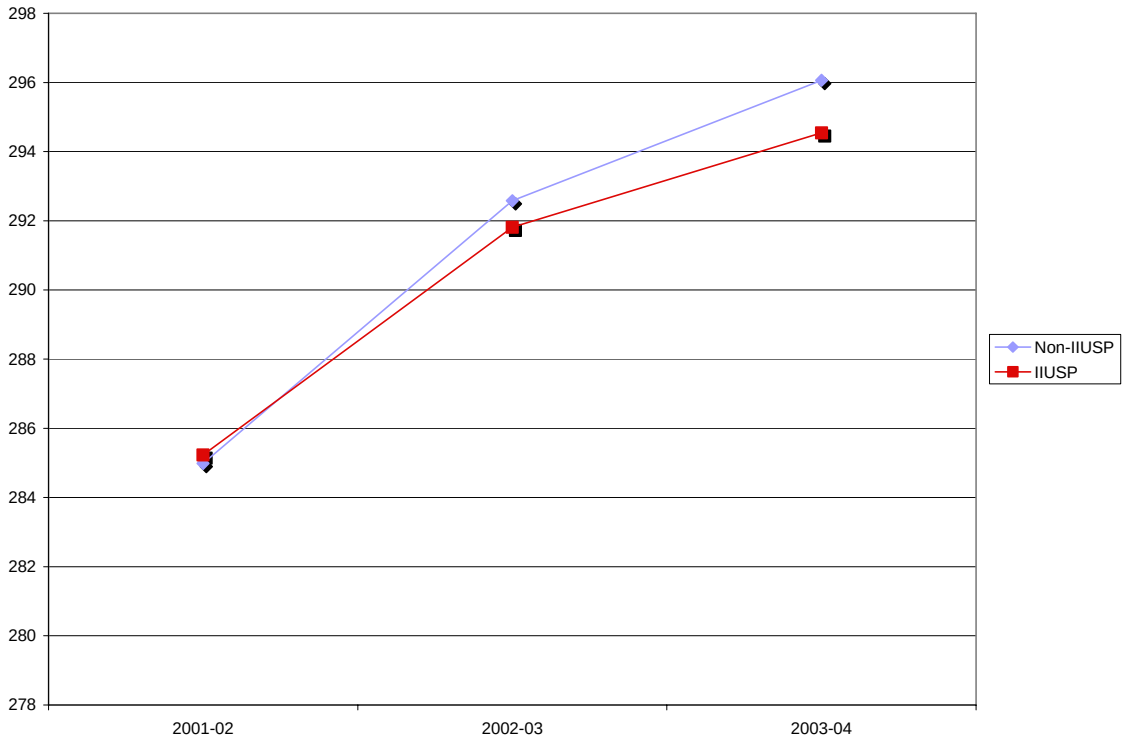
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores), Cohort 3 Middle Schools



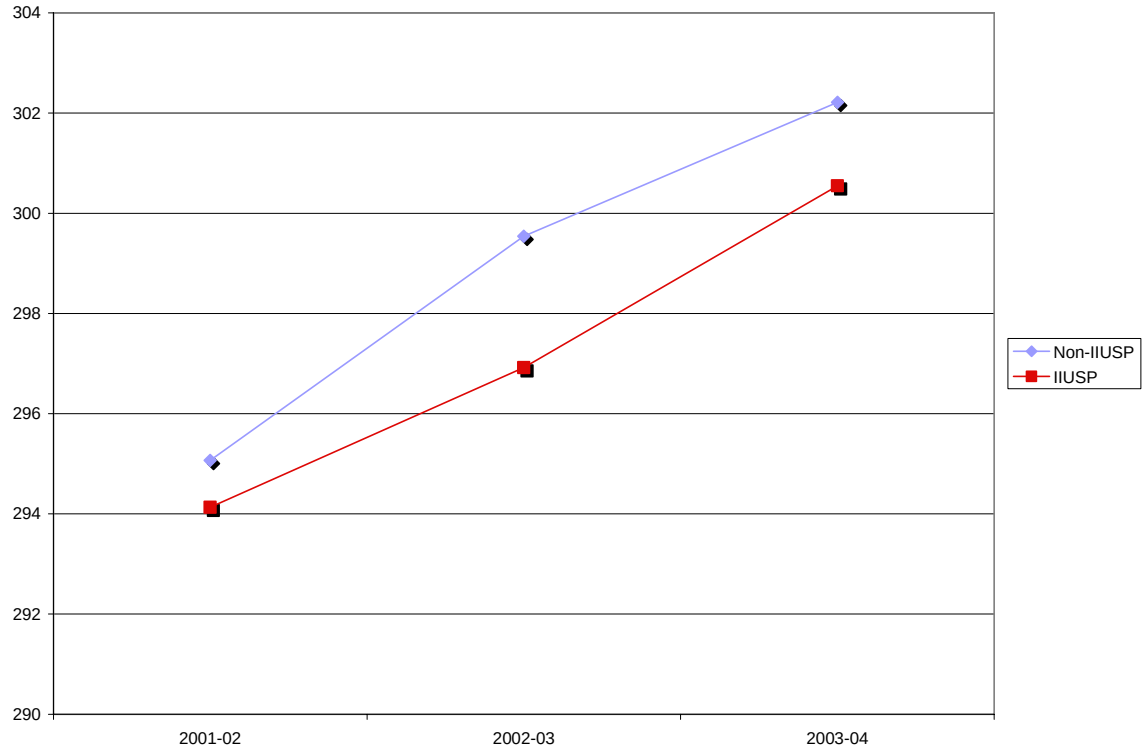
HLM Regression for Cohort 3 Middle Schools, CST ELA and Math Scores

	ELA		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	333.840	<.0001	332.750	<.0001
GIRL	8.915	<.0001	-0.667	<.0001
ASIAN	0.382	0.1572	15.395	<.0001
HISPANIC	-15.310	<.0001	-15.059	<.0001
BLACK	-22.915	<.0001	-23.678	<.0001
OTHERS	-5.745	<.0001	-2.115	<.0001
ELMISN	-11.629	<.0001	-9.522	<.0001
EL	-25.171	<.0001	-14.586	<.0001
R_FEP	16.015	<.0001	19.160	<.0001
FEP	11.492	<.0001	12.617	<.0001
FLUNCH	-8.576	<.0001	-6.796	<.0001
HIGHPARED	10.472	<.0001	9.696	<.0001
PARED_MISSING	-1.691	<.0001	-1.475	<.0001
SPECED	-36.155	<.0001	-33.545	<.0001
YEAR03	7.598	<.0001	4.473	<.0001
YEAR04	3.482	<.0001	2.672	<.0001
GRADE7	-2.952	<.0001	-5.365	<.0001
GRADE8	-5.037	<.0001	-13.142	<.0001
PCT_MEALS	-15.287	<.0001	-14.662	<.0001
IIUSP3	0.248	0.8092	-0.937	0.4988
CSR3	-9.585	0.1472	-2.250	0.7994
YEAR03*IIUSP3	-1.015	0.0008	-1.687	<.0001
YEAR04*IIUSP3	-0.755	0.0124	0.957	0.0053
YEAR03*CSR3	2.817	0.2375	0.723	0.7839
YEAR04*CSR3	6.357	0.0072	-1.944	0.4678
YEAR03*IIUSP3*HP	2.882	<.0001	4.903	<.0001
YEAR04*IIUSP3*HP	0.782	0.1194	-1.713	0.0027

Relative CST ELA Performance, Cohort 3 Middle Schools



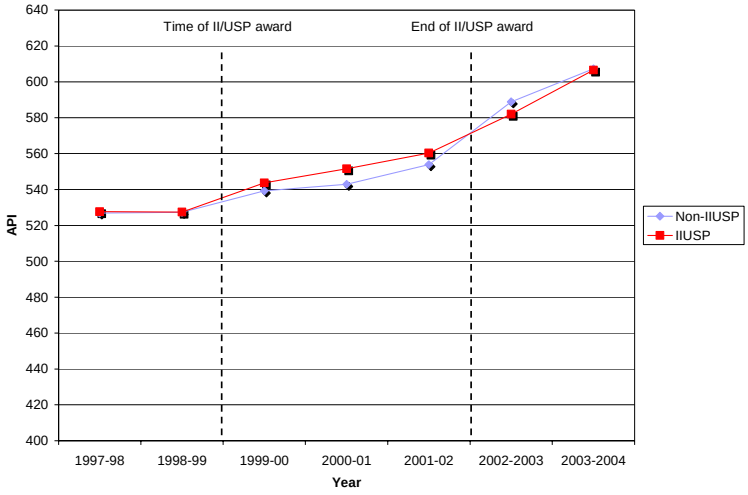
Relative CST Math Performance, Cohort 3 Middle Schools



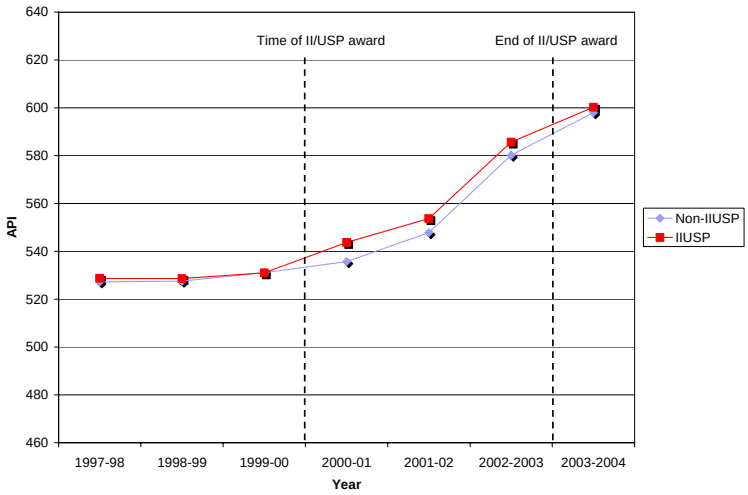
HLM Regression for High Schools, API scores

	Cohort 1		Cohort 2		Cohort 3	
Variables	**Estimate**	**P value**	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	573.440	<.0001	516.660	<.0001	588.310	<.0001
IMP_PCT_AF_AM	-1.886	<.0001	-2.162	<.0001	-1.932	<.0001
IMP_PCT_ASIAN	-0.150	0.5322	0.652	0.0092	0.095	0.6522
IMP_PCT_HISP	-1.172	<.0001	-1.003	<.0001	-1.118	<.0001
IMP_PCT_MEALS	-0.059	0.4943	-0.050	0.4891	-0.023	0.7556
IMP_PCT_EL	-0.871	<.0001	-0.247	0.0366	-0.467	<.0001
IMP_MOBILITY	-0.187	0.0078	0.007	0.9114	-0.127	0.0018
IMP_AVG_ED	12.985	0.0019	32.137	<.0001	19.639	<.0001
IMP_PCT_FULL	0.376	0.0071	0.206	0.064	-0.003	0.9733
IIUSP1	0.756	0.9212	1.455	0.7883	-3.842	0.4233
CSRD1	-15.396	0.3267	-2.307	0.9058	-3.258	0.8329
YEAR99	0.398	0.8673	0.255	0.8998	-46.115	<.0001
YEAR00	11.941	<.0001	3.571	0.0779	0.103	0.9582
YEAR01	3.672	0.1829	4.618	0.0247	15.974	<.0001
YEAR02	10.942	0.001	12.129	<.0001	-2.791	0.1592
YEAR03	35.034	<.0001	32.303	<.0001	13.248	<.0001
YEAR04	18.378	<.0001	17.834	<.0001	31.185	<.0001
IIUSP1*YEAR99	-0.628	0.8923	-0.402	0.9124	17.930	<.0001
IIUSP1*YEAR00	4.345	0.3482	-1.156	0.7518	0.632	0.8688
IIUSP1*YEAR01	4.173	0.39	8.113	0.0276	1.347	0.7272
IIUSP1*YEAR02	-2.191	0.6755	-2.118	0.5796	-0.774	0.8424
IIUSP1*YEAR03	-13.390	0.0182	-0.376	0.9252	-2.880	0.4732
IIUSP1*YEAR04	6.217	0.2874	-3.251	0.4269	7.211	0.0981
CSRD1*YEAR99	0.230	0.9806	-4.549	0.7409	-5.332	0.2434
CSRD1*YEAR00	14.270	0.132	9.835	0.4367	-0.838	0.9409
CSRD1*YEAR01	-2.946	0.7566	-23.576	0.0664	2.479	0.8267
CSRD1*YEAR02	4.814	0.6241	7.272	0.6019	7.898	0.4864
CSRD1*YEAR03	16.573	0.1038	4.503	0.7451	1.219	0.9201
CSRD1*YEAR04	-1.344	0.8946	7.322	0.5647	-7.834	0.5373

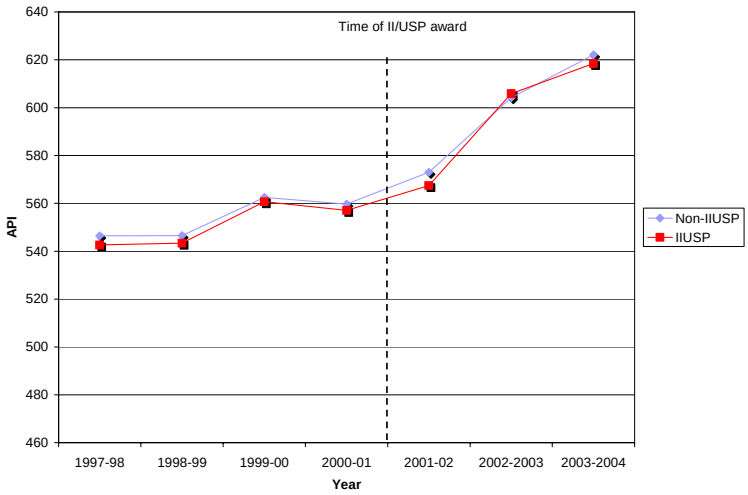
Relative API Performance, Cohort 1 High Schools



Relative API Performance, Cohort 2 High Schools



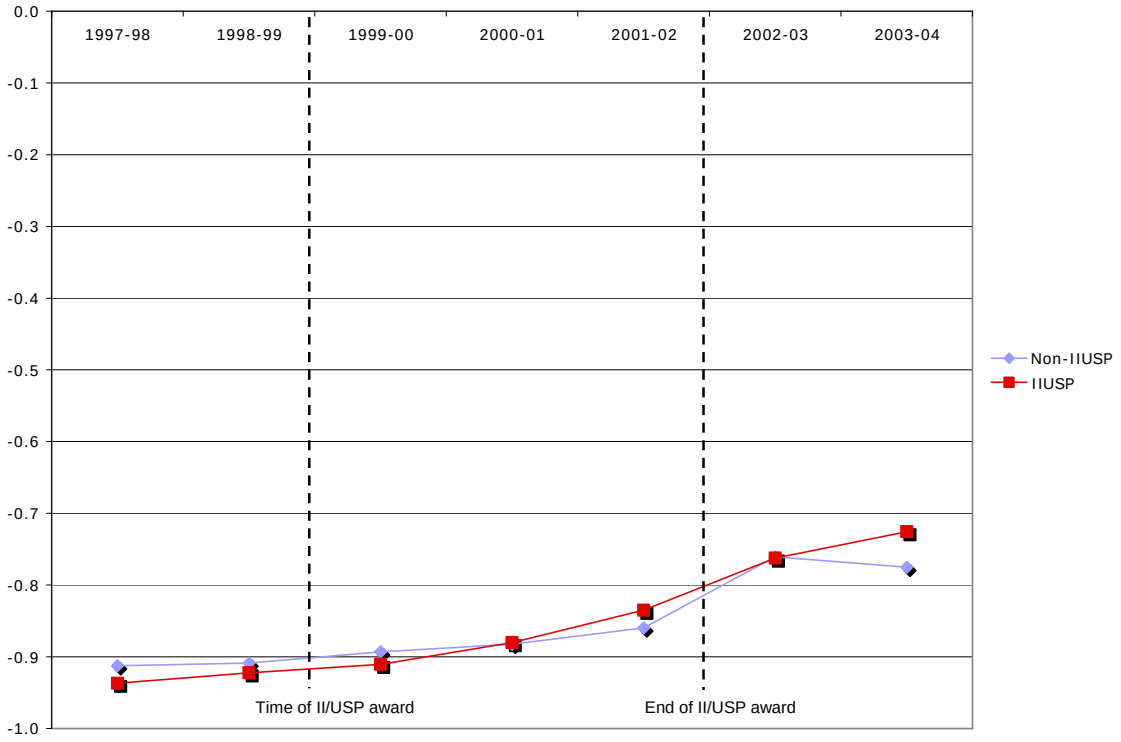
Relative API Performance, Cohort 3 High Schools



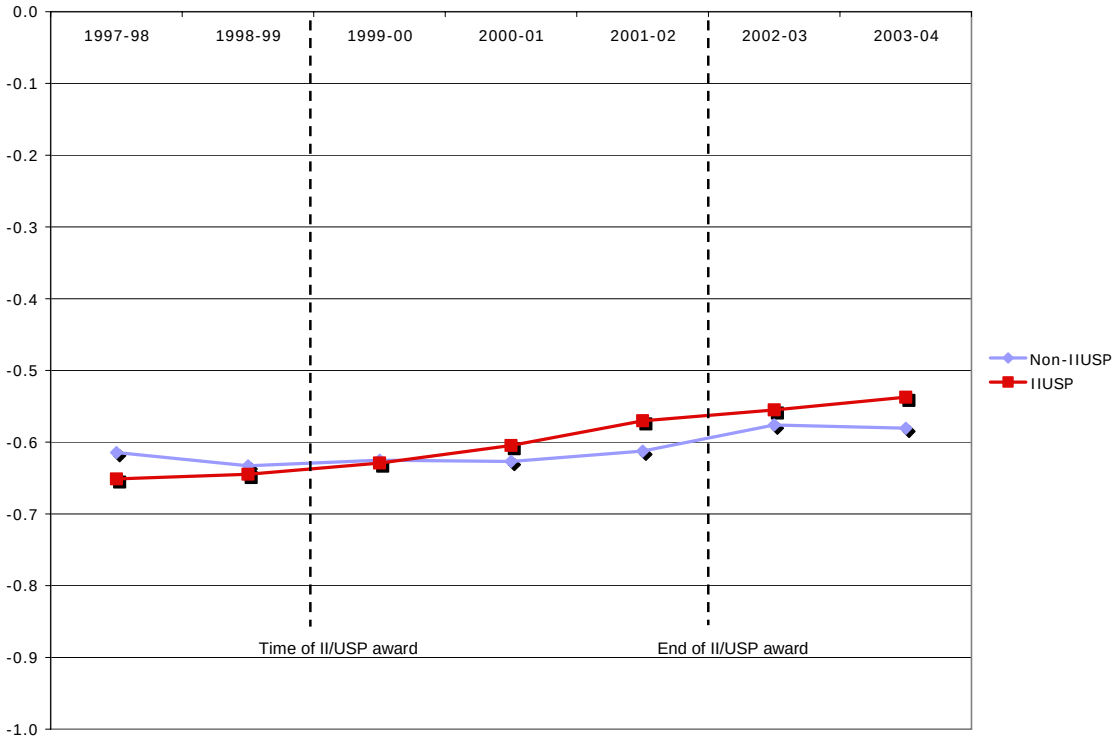
HLM Regression for Cohort 1 High Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.075	0.0381	0.121	0.0002
GIRL	0.110	<.0001	-0.066	<.0001
ASIAN	-0.120	<.0001	0.257	<.0001
HISPANIC	-0.327	<.0001	-0.326	<.0001
BLACK	-0.535	<.0001	-0.529	<.0001
OTHERS	-0.213	<.0001	-0.119	<.0001
ELMISN	-0.038	<.0001	0.040	<.0001
EL	-0.616	<.0001	-0.336	<.0001
R_FEP	0.068	<.0001	0.112	<.0001
FEP	0.101	<.0001	0.108	<.0001
FLUNCH	-0.078	<.0001	-0.041	<.0001
HIGHPARED	0.245	<.0001	0.194	<.0001
PARED_MISSING	-0.016	<.0001	-0.024	<.0001
SPECED	-0.746	<.0001	-0.611	<.0001
YEAR99	0.004	0.4156	-0.019	<.0001
YEAR00	0.016	<.0001	0.008	0.0418
YEAR01	0.011	0.0051	-0.002	0.672
YEAR02	0.023	<.0001	0.014	<.0001
YEAR03	0.099	<.0001	0.036	<.0001
YEAR04	-0.014	<.0001	-0.005	0.2063
GRADE10	0.005	0.0163	0.016	<.0001
GRADE11	0.017	<.0001	0.003	0.1719
PCT_MEALS	-0.109	0.0908	-0.181	0.0018
IIUSP1	-0.024	0.3708	-0.037	0.1314
CSR1	-0.108	0.0234	-0.078	0.0647
YEAR99*IIUSP1	0.011	0.1949	0.025	0.0022
YEAR00*IIUSP1	-0.004	0.5726	0.008	0.2444
YEAR01*IIUSP1	0.020	0.0029	0.026	<.0001
YEAR02*IIUSP1	0.022	0.0006	0.020	0.0015
YEAR03*IIUSP1	-0.026	<.0001	-0.021	0.0009
YEAR04*IIUSP1	0.051	<.0001	0.022	0.0005
YEAR99*CSR1	0.012	0.4282	-0.012	0.4200
YEAR00*CSR1	0.067	<.0001	0.103	<.0001
YEAR01*CSR1	0.003	0.7926	0.013	0.2467
YEAR02*CSR1	-0.028	0.0180	0.002	0.8508
YEAR03*CSR1	0.080	<.0001	0.029	0.0111
YEAR04*CSR1	0.021	0.0698	0.042	0.0002

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores), Cohort 1 High Schools



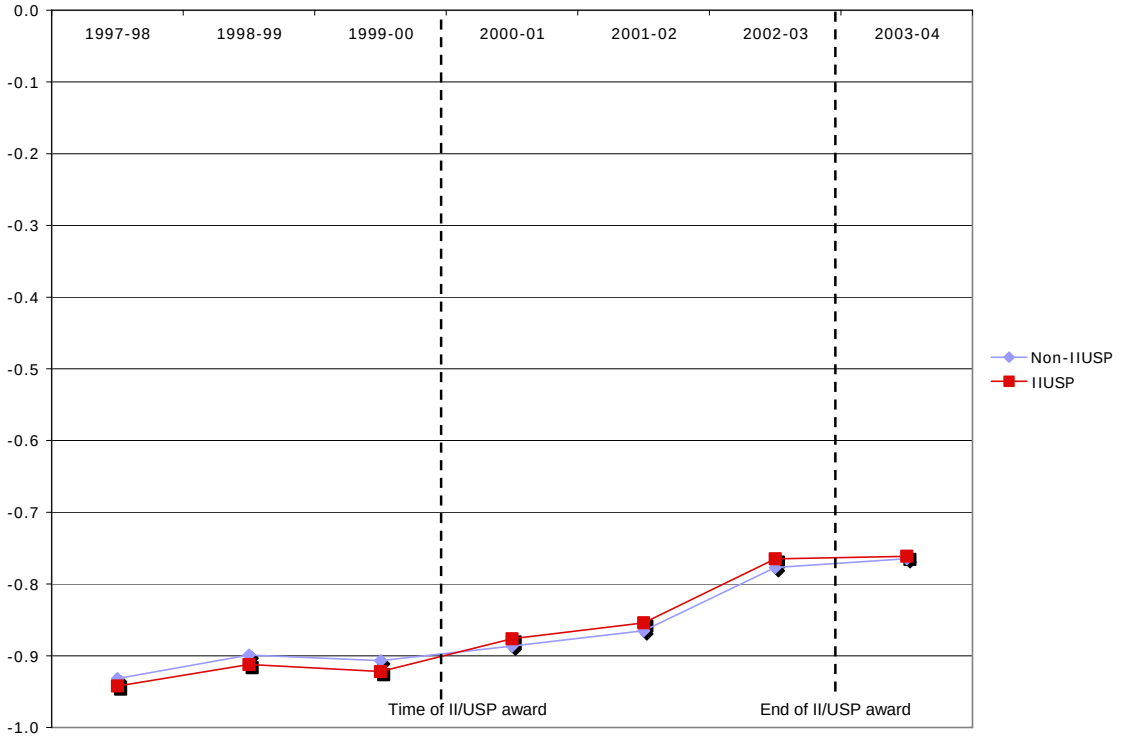
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores), Cohort 1 High Schools



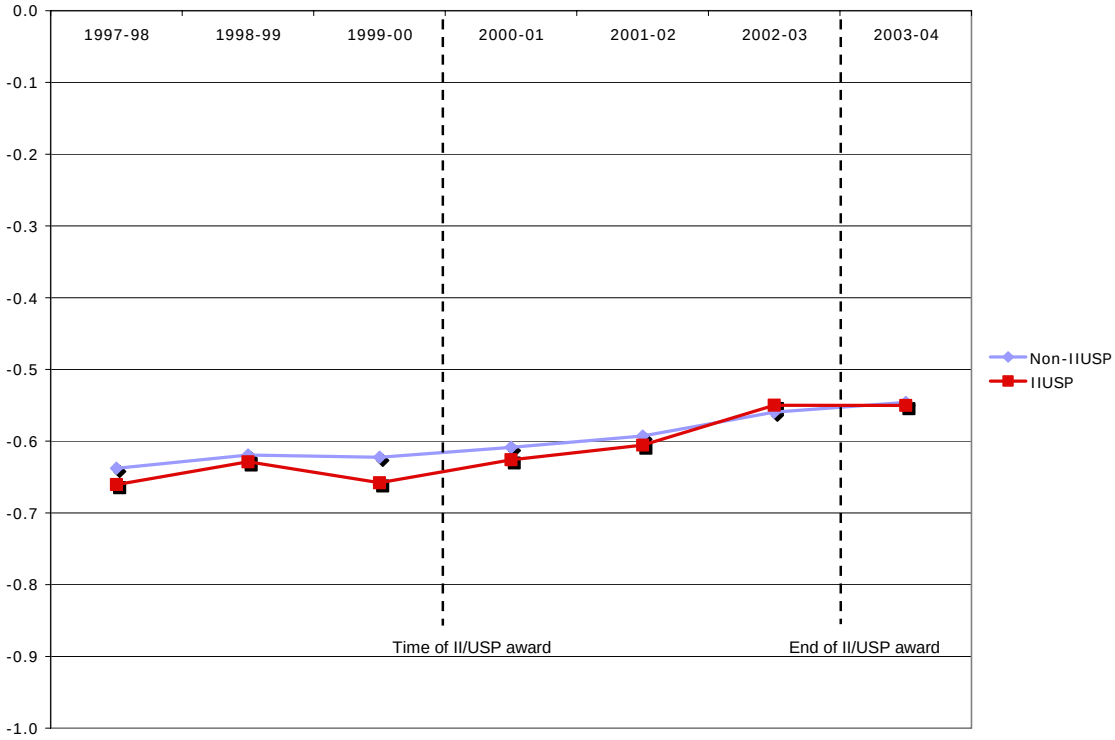
HLM Regression for Cohort 2 High Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.061	0.011	0.092	0.0001
GIRL	0.116	<.0001	-0.067	<.0001
ASIAN	-0.083	<.0001	0.252	<.0001
HISPANIC	-0.312	<.0001	-0.324	<.0001
BLACK	-0.515	<.0001	-0.520	<.0001
OTHERS	-0.181	<.0001	-0.106	<.0001
ELMISN	0.005	0.5198	0.051	<.0001
EL	-0.626	<.0001	-0.348	<.0001
R_FEP	0.077	<.0001	0.138	<.0001
FEP	0.129	<.0001	0.147	<.0001
FLUNCH	-0.095	<.0001	-0.060	<.0001
HIGHPARED	0.246	<.0001	0.203	<.0001
PARED_MISSING	-0.008	<.0001	-0.020	<.0001
SPECED	-0.770	<.0001	-0.629	<.0001
YEAR99	0.033	<.0001	0.019	<.0001
YEAR00	-0.008	0.0047	-0.003	0.2526
YEAR01	0.021	<.0001	0.014	<.0001
YEAR02	0.021	<.0001	0.016	<.0001
YEAR03	0.088	<.0001	0.033	<.0001
YEAR04	0.012	<.0001	0.013	<.0001
GRADE10	-0.007	<.0001	0.006	<.0001
GRADE11	-0.001	0.6558	-0.011	<.0001
PCT_MEALS	-0.135	0.0044	-0.143	0.0021
IIUSP2	-0.010	0.5752	-0.023	0.2076
CSR2	0.021	0.7380	0.011	0.8582
YEAR99*IIUSP2	-0.003	0.6088	0.013	0.0166
YEAR00*IIUSP2	-0.002	0.7027	-0.026	<.0001
YEAR01*IIUSP2	0.025	<.0001	0.018	<.0001
YEAR02*IIUSP2	0.001	0.8251	0.005	0.2909
YEAR03*IIUSP2	0.001	0.8199	0.022	<.0001
YEAR04*IIUSP2	-0.009	0.0550	-0.014	0.0018
YEAR99*CSR2	0.004	0.8393	-0.023	0.2331
YEAR00*CSR2	-0.006	0.7174	-0.011	0.5377
YEAR01*CSR2	-0.013	0.4286	-0.034	0.0387
YEAR02*CSR2	0.014	0.3847	-0.003	0.8659
YEAR03*CSR2	0.105	<.0001	0.046	0.0032
YEAR04*CSR2	-0.044	0.0047	0.050	0.0012

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores), Cohort 2 High Schools



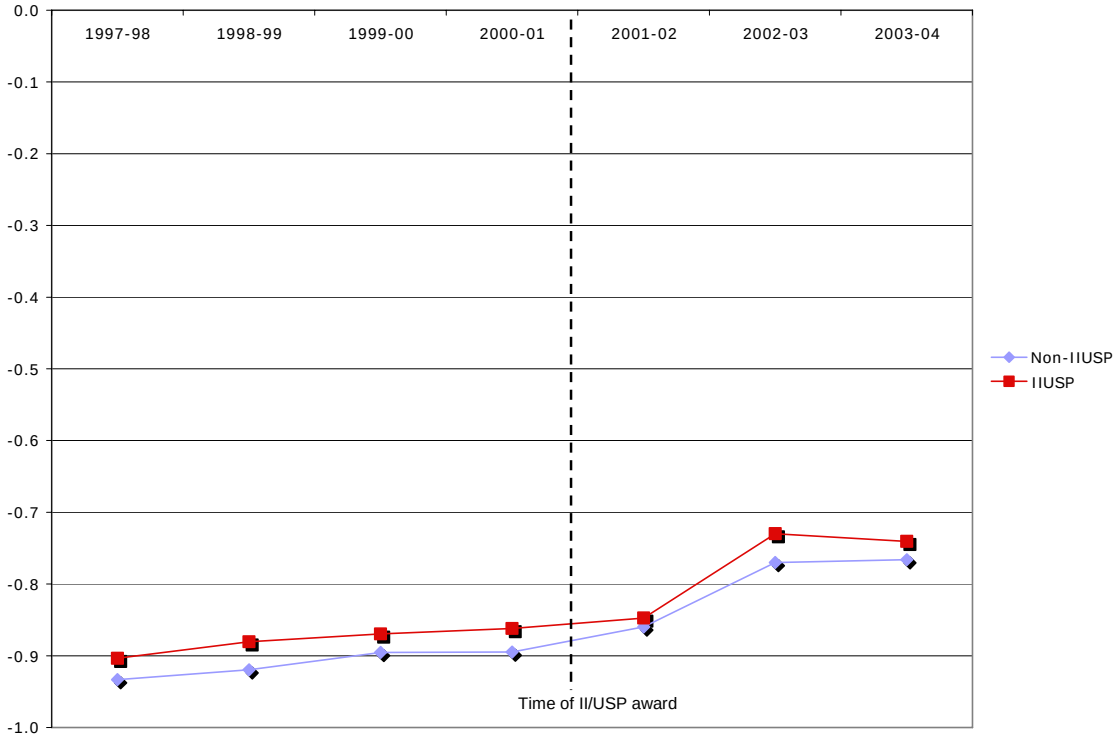
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores), Cohort 2 High Schools



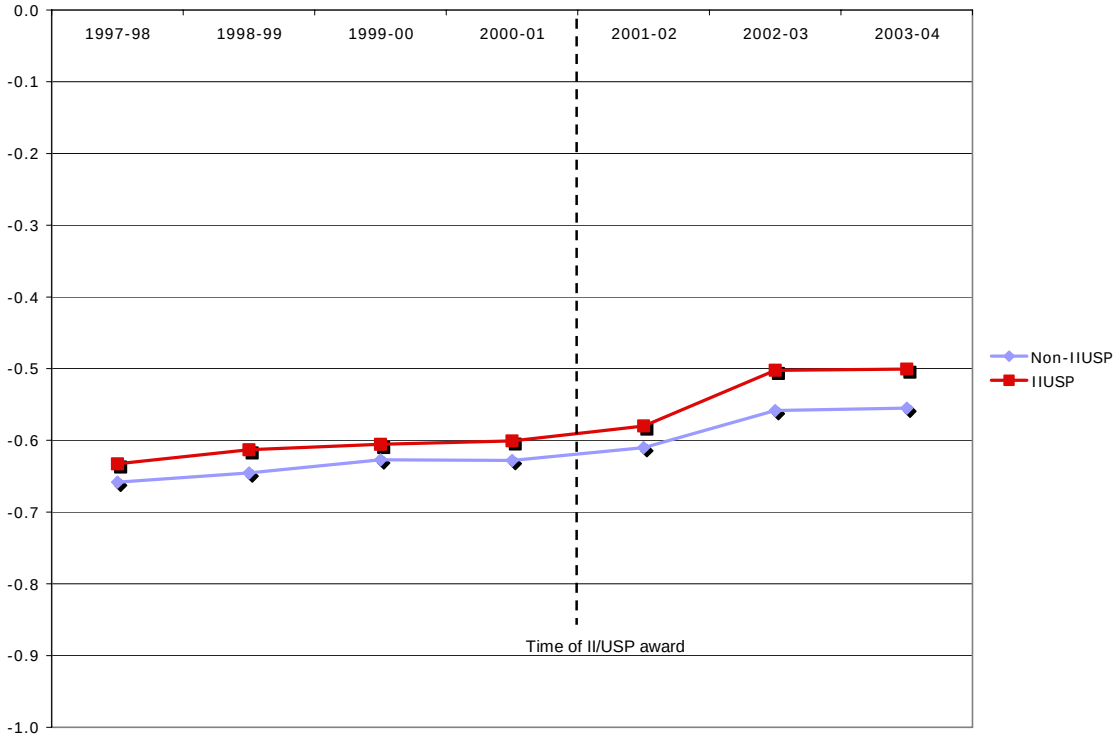
HLM Regression for Cohort 3 High Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.037	0.0908	0.034	0.141
GIRL	0.114	<.0001	-0.073	<.0001
ASIAN	-0.096	<.0001	0.249	<.0001
HISPANIC	-0.319	<.0001	-0.325	<.0001
BLACK	-0.520	<.0001	-0.523	<.0001
OTHERS	-0.196	<.0001	-0.112	<.0001
ELMISN	-0.114	<.0001	-0.018	0.0048
EL	-0.636	<.0001	-0.359	<.0001
R_FEP	0.075	<.0001	0.132	<.0001
FEP	0.102	<.0001	0.122	<.0001
FLUNCH	-0.100	<.0001	-0.062	<.0001
HIGHPARED	0.245	<.0001	0.205	<.0001
PARED_MISSING	-0.006	<.0001	-0.020	<.0001
SPECED	-0.781	<.0001	-0.641	<.0001
YEAR99	0.014	<.0001	0.013	<.0001
YEAR00	0.024	<.0001	0.018	<.0001
YEAR01	0.001	0.7814	-0.001	0.7136
YEAR02	0.035	<.0001	0.018	<.0001
YEAR03	0.090	<.0001	0.052	<.0001
YEAR04	0.004	0.0738	0.003	0.1548
GRADE10	-0.010	<.0001	0.001	0.5053
GRADE11	-0.005	<.0001	-0.017	<.0001
PCT_MEALS	-0.044	0.3299	-0.025	0.5993
IIUSP3	0.030	0.1124	0.026	0.1955
CSR3	0.051	0.405	0.089	0.1678
YEAR99*IIUSP3	0.009	0.1277	0.007	0.2524
YEAR00*IIUSP3	-0.013	0.0139	-0.010	0.0411
YEAR01*IIUSP3	0.007	0.1512	0.006	0.231
YEAR02*IIUSP3	-0.020	<.0001	0.003	0.5268
YEAR03*IIUSP3	0.028	<.0001	0.026	<.0001
YEAR04*IIUSP3	-0.015	0.0019	-0.002	0.7202
YEAR99*CSR3	-0.037	0.0052	-0.071	<.0001
YEAR00*CSR3	0.004	0.7319	-0.006	0.633
YEAR01*CSR3	0.020	0.0941	-0.017	0.1341
YEAR02*CSR3	0.005	0.6566	0.037	0.001
YEAR03*CSR3	0.030	0.0095	-0.059	<.0001
YEAR04*CSR3	-0.007	0.5537	0.006	0.6033
YEAR99*IIUSP3*HP	0.032	0.0012	0.025	0.0112
YEAR00*IIUSP3*HP	-0.055	<.0001	-0.040	<.0001
YEAR01*IIUSP3*HP	0.009	0.2212	-0.004	0.6265
YEAR02*IIUSP3*HP	0.021	0.0052	-0.007	0.3567
YEAR03*IIUSP3*HP	0.040	<.0001	0.009	0.2406
YEAR04*IIUSP3*HP	0.050	<.0001	-0.003	0.6305

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores), Cohort 3 High Schools



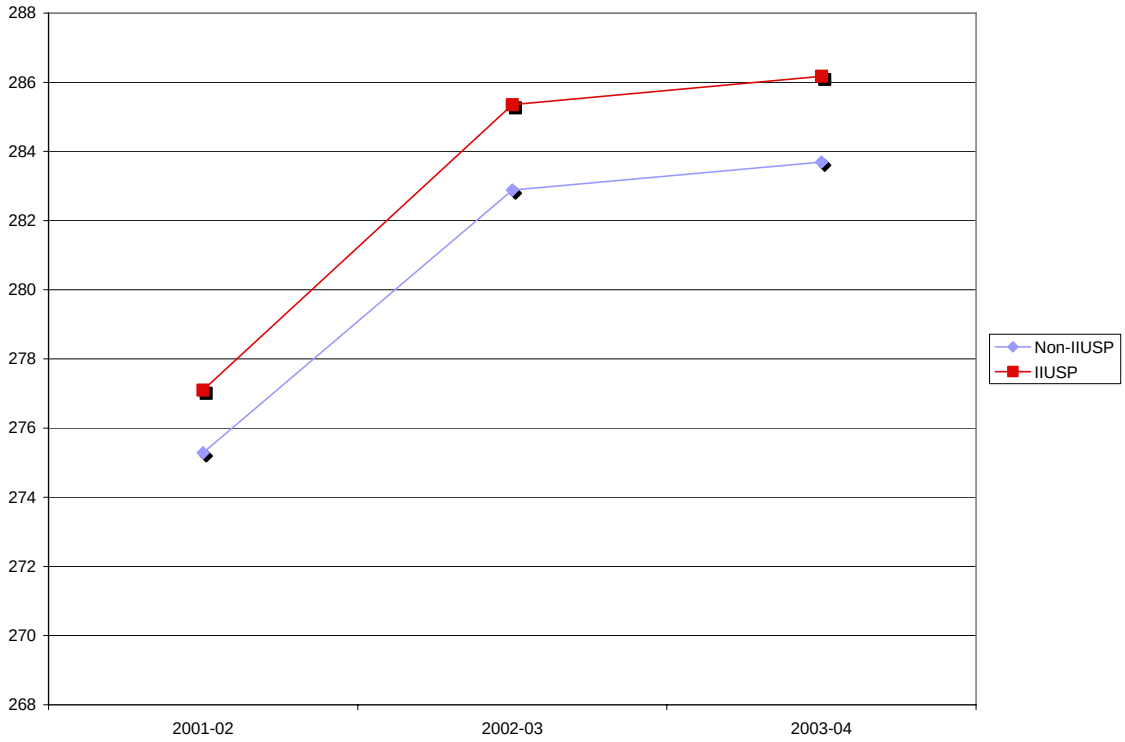
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores), Cohort 3 High Schools



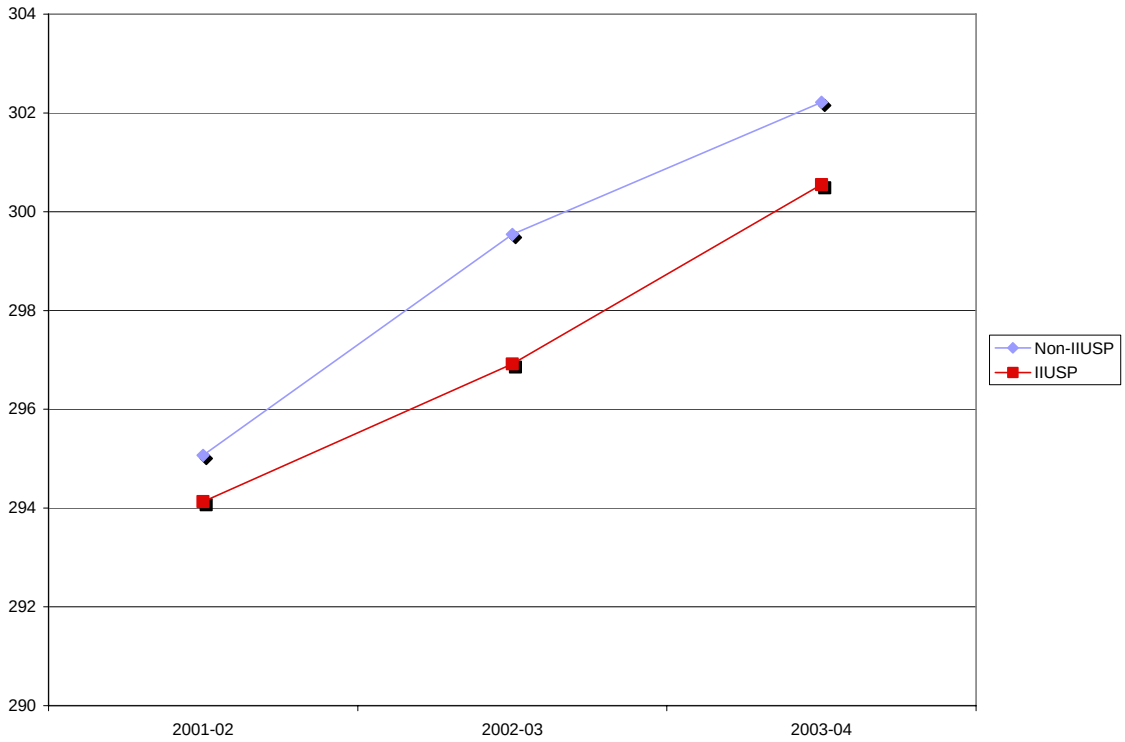
HLM Regression for Cohort 3 High Schools, CST ELA and Math Scores

	ELA		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	326.890	<.0001	312.710	<.0001
GIRL	9.101	<.0001	-2.780	<.0001
ASIAN	0.405	0.0556	13.378	<.0001
HISPANIC	-18.227	<.0001	-13.179	<.0001
BLACK	-27.899	<.0001	-20.069	<.0001
OTHERS	-6.854	<.0001	-3.669	<.0001
ELMISN	-9.406	<.0001	-1.638	0.2492
EL	-33.519	<.0001	-11.489	<.0001
R_FEP	6.587	<.0001	4.647	<.0001
FEP	7.619	<.0001	5.298	<.0001
FLUNCH	-6.083	<.0001	-1.249	<.0001
HIGHPARED	13.054	<.0001	7.041	<.0001
PARED_MISSING	-0.878	<.0001	-0.073	0.554
SPECED	-42.688	<.0001	-26.020	<.0001
YEAR03	7.594	<.0001	0.666	<.0001
YEAR04	0.807	<.0001	-0.701	<.0001
GRADE10	-3.498	<.0001	-6.429	<.0001
GRADE11	-9.322	<.0001	-13.375	<.0001
PCT_MEALS	0.404	0.8925	-12.410	0.0053
IIUSP3	1.806	0.138	0.974	0.5861
CSR3	5.126	0.2016	-0.769	0.8965
YEAR03*IIUSP3	0.660	0.0101	-0.162	0.5635
YEAR04*IIUSP3	0.014	0.9542	0.480	0.0782
YEAR03*CSR3	-0.222	0.721	-2.990	<.0001
YEAR04*CSR3	-1.509	0.0122	0.439	0.4853
YEAR03*IIUSP3*HP	2.761	<.0001	1.987	<.0001
YEAR04*IIUSP3*HP	0.162	0.68	-0.325	0.4434

Relative CST ELA Performance, Cohort 3 High Schools



Relative CST Math Performance, Cohort 3 High Schools

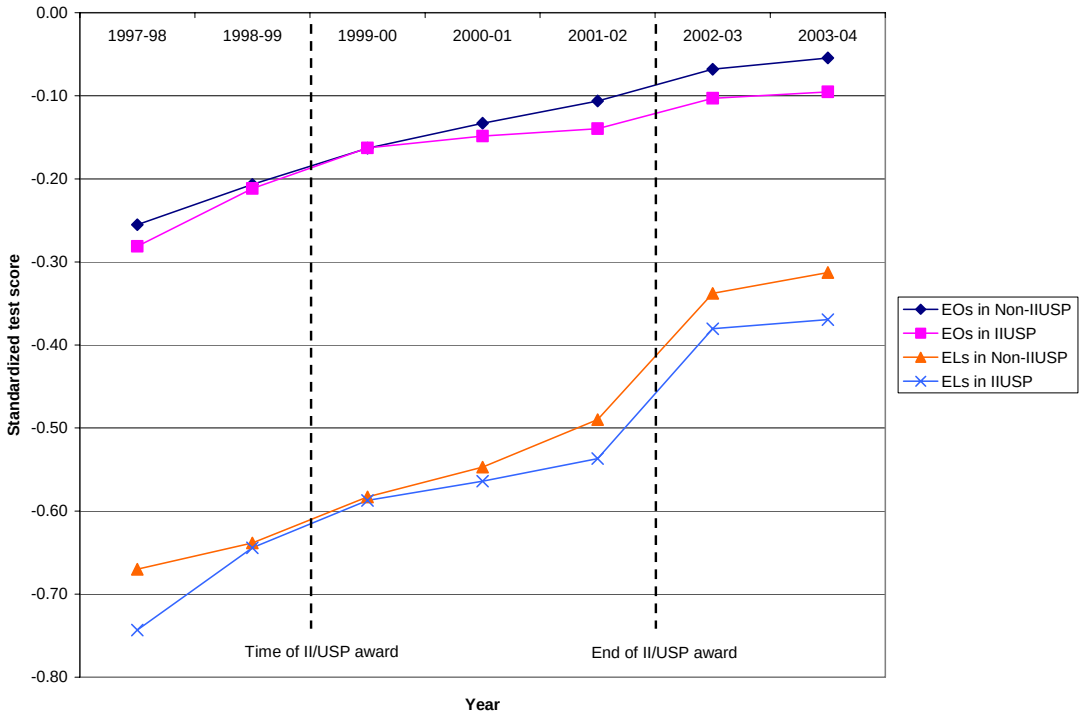


Appendix A-9:
Sub Population Regression Results

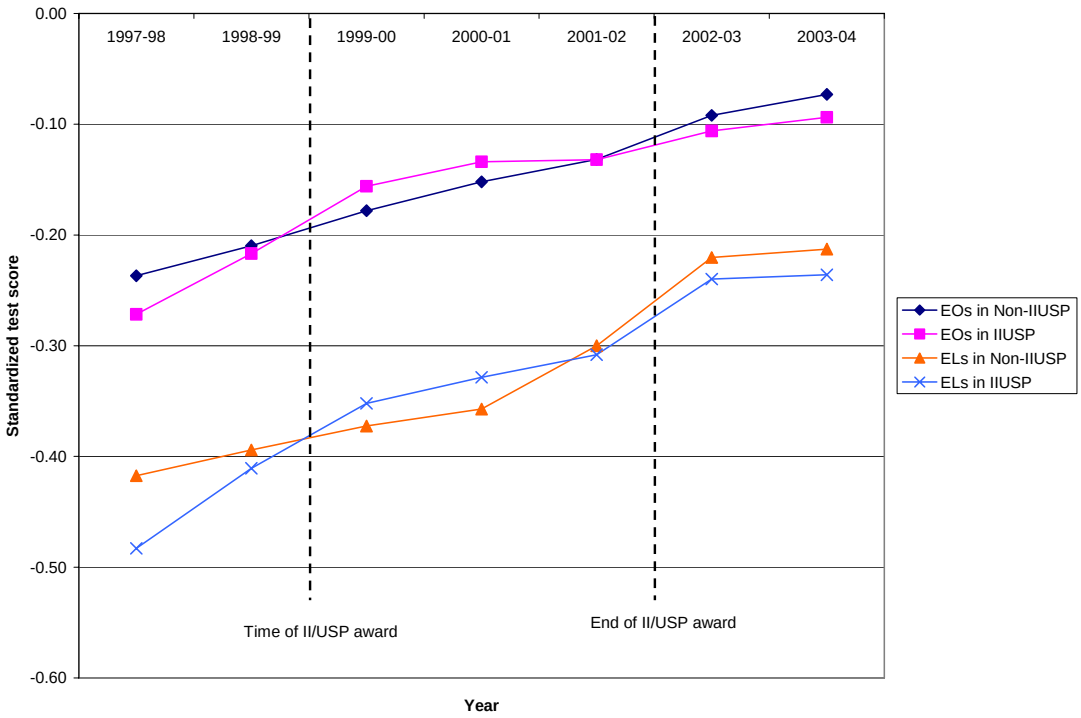
HLM Regression for ELs and EOs in Cohort 1 Elementary Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.129	<.0001	0.063	0.0126
GIRL	0.108	<.0001	-0.015	<.0001
ASIAN	-0.001	0.8718	0.198	<.0001
HISPANIC	-0.255	<.0001	-0.233	<.0001
BLACK	-0.406	<.0001	-0.454	<.0001
OTHERS	-0.089	<.0001	-0.052	<.0001
ELMISN	-0.141	<.0001	-0.090	<.0001
EL	-0.415	<.0001	-0.181	<.0001
R_FEP	0.348	<.0001	0.485	<.0001
FEP	0.186	<.0001	0.231	<.0001
FLUNCH	-0.174	<.0001	-0.143	<.0001
HIGHPARED	0.223	<.0001	0.203	<.0001
PARED_MISSING	-0.010	<.0001	-0.033	<.0001
SPECED	-0.594	<.0001	-0.658	<.0001
YEAR99	0.049	<.0001	0.027	<.0001
YEAR00	0.044	<.0001	0.032	<.0001
YEAR01	0.030	<.0001	0.026	<.0001
YEAR02	0.027	<.0001	0.020	<.0001
YEAR03	0.038	<.0001	0.040	<.0001
YEAR04	0.014	0.0001	0.019	<.0001
GRADE3	-0.031	<.0001	-0.023	<.0001
GRADE4	-0.062	<.0001	-0.061	<.0001
GRADE5	-0.089	<.0001	-0.101	<.0001
PCT_MEALS	-0.189	<.0001	-0.097	0.0028
IIUSP1	-0.026	0.0193	-0.035	0.0056
YEAR99*IIUSP1	0.021	0.0005	0.028	<.0001
YEAR00*IIUSP1	0.005	0.3496	0.029	<.0001
YEAR01*IIUSP1	-0.016	0.0035	-0.004	0.4782
YEAR02*IIUSP1	-0.018	0.0013	-0.018	0.0017
YEAR03*IIUSP1	-0.002	0.7927	-0.014	0.0189
YEAR04*IIUSP1	-0.006	0.2949	-0.007	0.2817
EL*YEAR99	-0.017	0.0004	-0.004	0.425
EL*YEAR00	0.012	0.0063	-0.010	0.0239
EL*YEAR01	0.006	0.1723	-0.011	0.0162
EL*YEAR02	0.030	<.0001	0.037	<.0001
EL*YEAR03	0.114	<.0001	0.040	<.0001
EL*YEAR04	0.011	0.0208	-0.011	0.0276
EL*IIUSP1	-0.047	<.0001	-0.031	<.0001
EL*YEAR99*IIUSP1	0.046	<.0001	0.021	0.0231
EL*YEAR00*IIUSP1	-0.004	0.677	0.008	0.3744
EL*YEAR01*IIUSP1	0.003	0.711	0.012	0.1389
EL*YEAR02*IIUSP1	-0.012	0.1414	-0.018	0.0287
EL*YEAR03*IIUSP1	0.006	0.4958	0.002	0.7703
EL*YEAR04*IIUSP1	-0.008	0.3291	0.003	0.7345

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores) for ELs and EOs in Cohort 1 II/USP and Comparison Elementary Schools



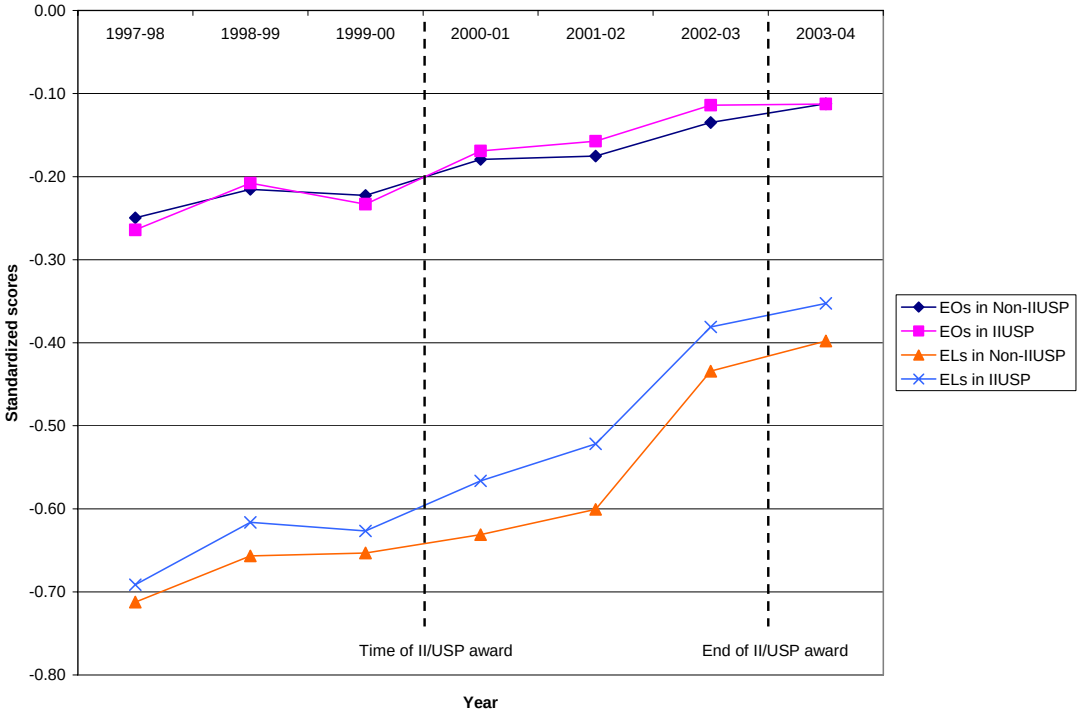
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores) for ELs and EOs in Cohort 1 II/USP and Comparison Elementary Schools



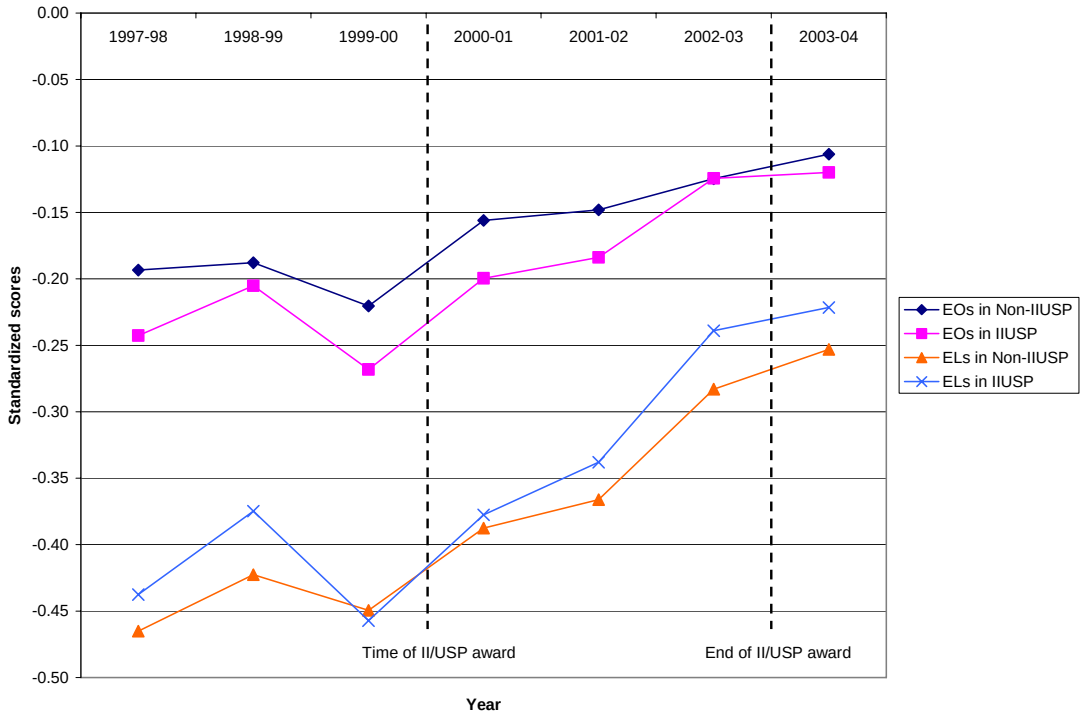
HLM Regression for ELs and EOs in Cohort 2 Elementary Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.194	<.0001	0.153	<.0001
GIRL	0.114	<.0001	-0.011	<.0001
ASIAN	-0.020	<.0001	0.172	<.0001
HISPANIC	-0.263	<.0001	-0.237	<.0001
BLACK	-0.412	<.0001	-0.459	<.0001
OTHERS	-0.119	<.0001	-0.073	<.0001
ELMISN	-0.176	<.0001	-0.112	<.0001
EL	-0.463	<.0001	-0.272	<.0001
R_FEP	0.370	<.0001	0.505	<.0001
FEP	0.198	<.0001	0.241	<.0001
FLUNCH	-0.176	<.0001	-0.145	<.0001
HIGHPARED	0.236	<.0001	0.217	<.0001
PARED_MISSING	-0.016	<.0001	-0.036	<.0001
SPECED	-0.614	<.0001	-0.668	<.0001
YEAR99	0.034	<.0001	0.006	0.3289
YEAR00	-0.007	0.1813	-0.033	<.0001
YEAR01	0.043	<.0001	0.064	<.0001
YEAR02	0.004	0.4324	0.008	0.1469
YEAR03	0.040	<.0001	0.024	<.0001
YEAR04	0.023	<.0001	0.018	0.0011
GRADE3	-0.028	<.0001	-0.039	<.0001
GRADE4	-0.057	<.0001	-0.070	<.0001
GRADE5	-0.073	<.0001	-0.094	<.0001
PCT_MEALS	-0.273	<.0001	-0.166	0.0001
IIUSP2	-0.014	0.2907	-0.049	0.0016
YEAR99*IIUSP2	0.022	0.0055	0.032	<.0001
YEAR00*IIUSP2	-0.018	0.015	-0.031	<.0001
YEAR01*IIUSP2	0.021	0.0028	0.004	0.5473
YEAR02*IIUSP2	0.007	0.2975	0.008	0.2774
YEAR03*IIUSP2	0.003	0.6472	0.036	<.0001
YEAR04*IIUSP2	-0.021	0.0038	-0.014	0.0655
EL*YEAR99	0.021	0.0199	0.037	<.0001
EL*YEAR00	0.011	0.2112	0.006	0.5130
EL*YEAR01	-0.021	0.0068	-0.002	0.7548
EL*YEAR02	0.026	0.0011	0.014	0.0973
EL*YEAR03	0.126	<.0001	0.060	<.0001
EL*YEAR04	0.014	0.0925	0.011	0.1798
EL*IIUSP2	0.035	<.0001	0.077	<.0001
EL*YEAR99*IIUSP2	-0.002	0.8472	-0.012	0.3442
EL*YEAR00*IIUSP2	0.004	0.7279	-0.025	0.0312
EL*YEAR01*IIUSP2	0.018	0.0939	0.013	0.2091
EL*YEAR02*IIUSP2	0.007	0.5143	0.010	0.3439
EL*YEAR03*IIUSP2	-0.029	0.0067	-0.020	0.0666
EL*YEAR04*IIUSP2	0.013	0.2171	0.002	0.8808

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores) for ELs and EOs in Cohort 2 II/USP and Comparison Elementary Schools



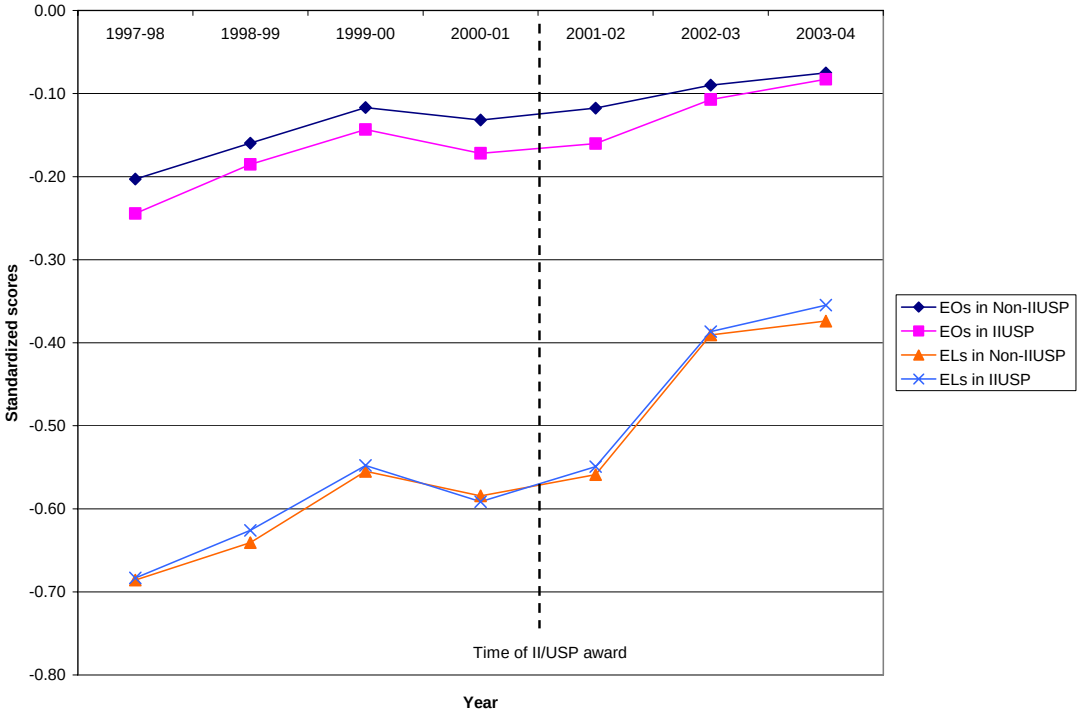
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores) for ELs and EOs in Cohort 2 II/USP and Comparison Elementary Schools



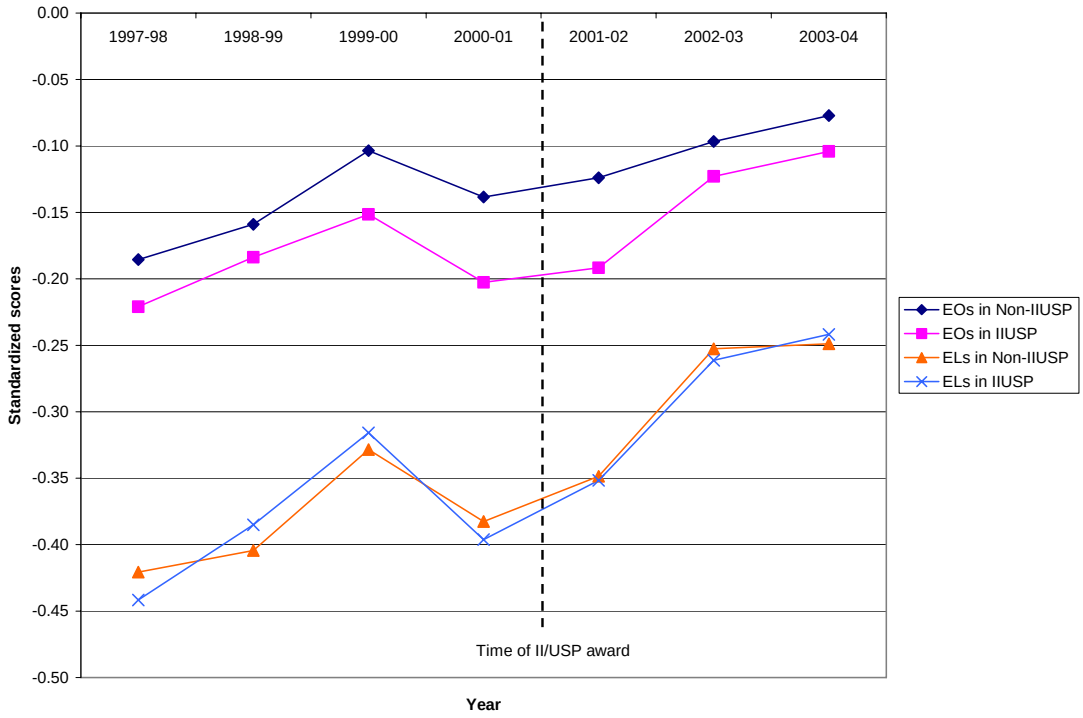
HLM Regression for ELs and EOs in Cohort 3 Elementary Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.199	<.0001	0.119	<.0001
GIRL	0.118	<.0001	-0.014	<.0001
ASIAN	-0.001	0.7679	0.188	<.0001
HISPANIC	-0.251	<.0001	-0.228	<.0001
BLACK	-0.407	<.0001	-0.447	<.0001
OTHERS	-0.097	<.0001	-0.053	<.0001
ELMISN	-0.155	<.0001	-0.107	<.0001
EL	-0.483	<.0001	-0.235	<.0001
R_FEP	0.322	<.0001	0.452	<.0001
FEP	0.164	<.0001	0.217	<.0001
FLUNCH	-0.189	<.0001	-0.158	<.0001
HIGHPARED	0.238	<.0001	0.220	<.0001
PARED_MISSING	-0.003	0.0341	-0.032	<.0001
SPECED	-0.640	<.0001	-0.681	<.0001
YEAR99	0.043	<.0001	0.027	<.0001
YEAR00	0.043	<.0001	0.055	<.0001
YEAR01	-0.015	<.0001	-0.035	<.0001
YEAR02	0.014	<.0001	0.015	<.0001
YEAR03	0.028	<.0001	0.027	<.0001
YEAR04	0.015	<.0001	0.020	<.0001
GRADE3	-0.018	<.0001	-0.024	<.0001
GRADE4	-0.044	<.0001	-0.052	<.0001
GRADE5	-0.060	<.0001	-0.078	<.0001
PCT_MEALS	-0.253	<.0001	-0.127	<.0001
IIUSP3	-0.041	<.0001	-0.035	0.0038
YEAR99*IIUSP3	0.016	0.0144	0.011	0.1064
YEAR00*IIUSP3	-0.001	0.8843	-0.023	0.0002
YEAR01*IIUSP3	-0.014	0.0151	-0.016	0.0046
YEAR02*IIUSP3	-0.002	0.6799	-0.004	0.5295
YEAR03*IIUSP3	0.025	<.0001	0.042	<.0001
YEAR04*IIUSP3	0.009	0.0973	-0.001	0.9033
EL*YEAR99	0.002	0.7838	-0.010	0.1184
EL*YEAR00	0.043	<.0001	0.021	0.0007
EL*YEAR01	-0.015	0.0072	-0.019	0.0006
EL*YEAR02	0.012	0.0312	0.019	0.0004
EL*YEAR03	0.140	<.0001	0.069	<.0001
EL*YEAR04	0.002	0.714	-0.016	0.0042
EL*IIUSP3	0.044	<.0001	0.014	0.0650
EL*YEAR99*IIUSP3	-0.004	0.7317	0.030	0.0052
EL*YEAR00*IIUSP3	-0.007	0.4798	0.016	0.0999
EL*YEAR01*IIUSP3	0.000	0.9662	-0.010	0.2687
EL*YEAR02*IIUSP3	0.019	0.0309	0.014	0.1094
EL*YEAR03*IIUSP3	-0.031	0.0004	-0.047	<.0001
EL*YEAR04*IIUSP3	0.006	0.5200	0.016	0.0654

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores) for ELs and EOs in Cohort 3 II/USP and Comparison Elementary Schools



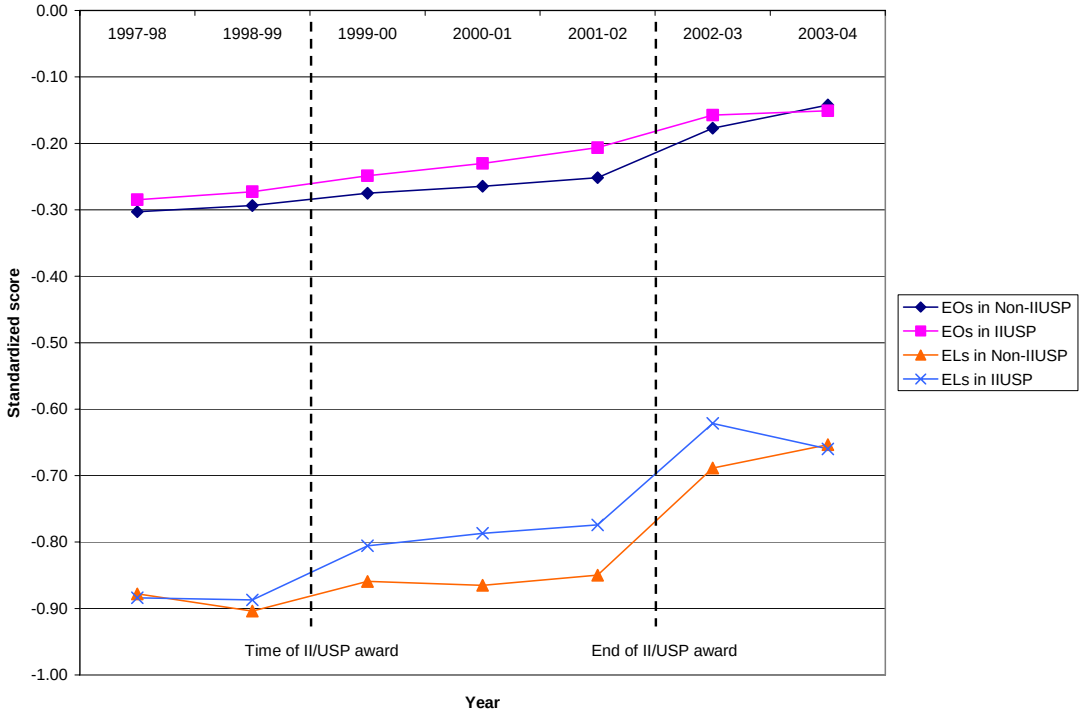
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores) for ELs and EOs in Cohort 3 II/USP and Comparison Elementary Schools



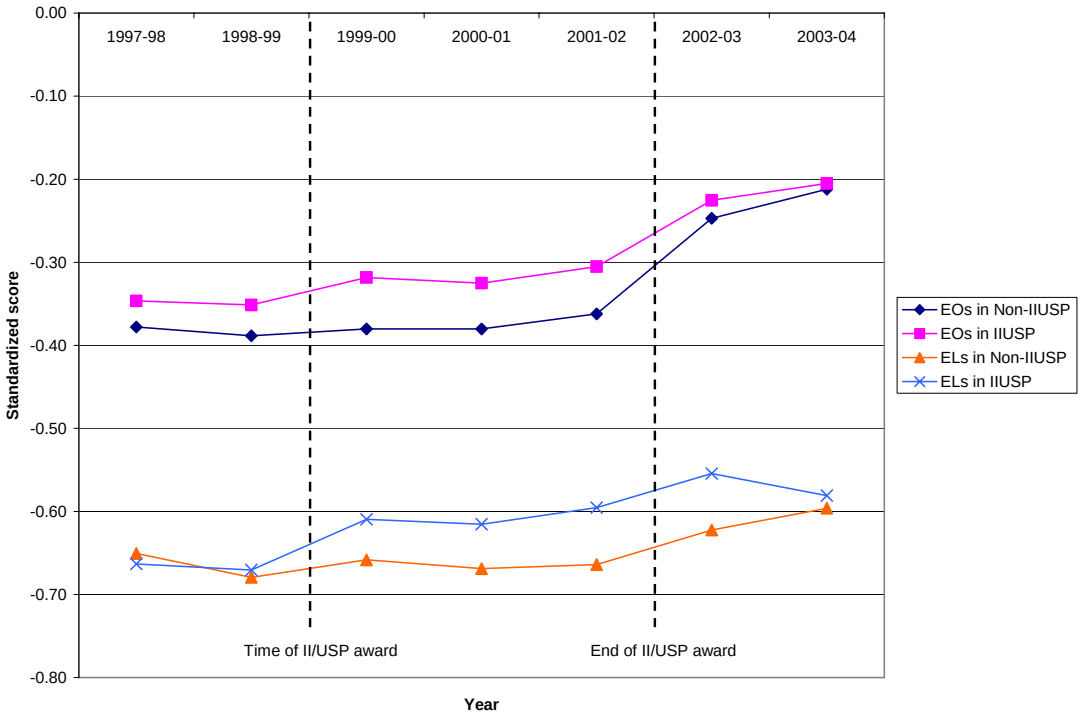
HLM Regression for ELs and EOs in Cohort 1 Middle Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.221	<.0001	0.056	0.0914
GIRL	0.103	<.0001	-0.007	<.0001
ASIAN	-0.063	<.0001	0.260	<.0001
HISPANIC	-0.338	<.0001	-0.315	<.0001
BLACK	-0.518	<.0001	-0.510	<.0001
OTHERS	-0.185	<.0001	-0.097	<.0001
ELMISN	-0.188	<.0001	-0.057	<.0001
EL	-0.575	<.0001	-0.273	<.0001
R_FEP	0.213	<.0001	0.282	<.0001
FEP	0.165	<.0001	0.184	<.0001
FLUNCH	-0.129	<.0001	-0.095	<.0001
HIGHPARED	0.216	<.0001	0.182	<.0001
PARED_MISSING	-0.036	<.0001	-0.037	<.0001
SPECED	-0.730	<.0001	-0.635	<.0001
YEAR99	0.009	0.0124	-0.011	0.0034
YEAR00	0.019	<.0001	0.008	0.0141
YEAR01	0.010	0.004	0.000	0.9709
YEAR02	0.013	0.0019	0.018	<.0001
YEAR03	0.074	<.0001	0.115	<.0001
YEAR04	0.035	<.0001	0.035	<.0001
GRADE7	-0.010	<.0001	0.020	<.0001
GRADE8	-0.026	<.0001	-0.003	0.1223
PCT_MEALS	-0.320	<.0001	-0.204	<.0001
IIUSP1	0.018	0.3171	0.031	0.0944
YEAR99*IIUSP1	0.003	0.7278	0.006	0.4368
YEAR00*IIUSP1	0.005	0.4434	0.025	0.0001
YEAR01*IIUSP1	0.008	0.2032	-0.007	0.2914
YEAR02*IIUSP1	0.011	0.1024	0.002	0.7849
YEAR03*IIUSP1	-0.026	0.0002	-0.035	<.0001
YEAR04*IIUSP1	-0.028	<.0001	-0.015	0.0245
EL*YEAR99	-0.035	<.0001	-0.018	0.0022
EL*YEAR00	0.026	<.0001	0.013	0.0207
EL*YEAR01	-0.016	0.0081	-0.010	0.0745
EL*YEAR02	0.002	0.7371	-0.013	0.0472
EL*YEAR03	0.087	<.0001	-0.074	<.0001
EL*YEAR04	0.000	0.9486	-0.009	0.1813
EL*IIUSP1	-0.024	0.0133	-0.044	<.0001
EL*YEAR99*IIUSP1	0.020	0.1406	0.016	0.2094
EL*YEAR00*IIUSP1	0.032	0.0094	0.015	0.2100
EL*YEAR01*IIUSP1	0.016	0.1543	0.011	0.3000
EL*YEAR02*IIUSP1	-0.013	0.2518	0.013	0.2369
EL*YEAR03*IIUSP1	0.017	0.1438	0.035	0.0018
EL*YEAR04*IIUSP1	-0.045	<.0001	-0.038	0.0006

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores) for ELs and EOs in Cohort 1 II/USP and Comparison Middle Schools



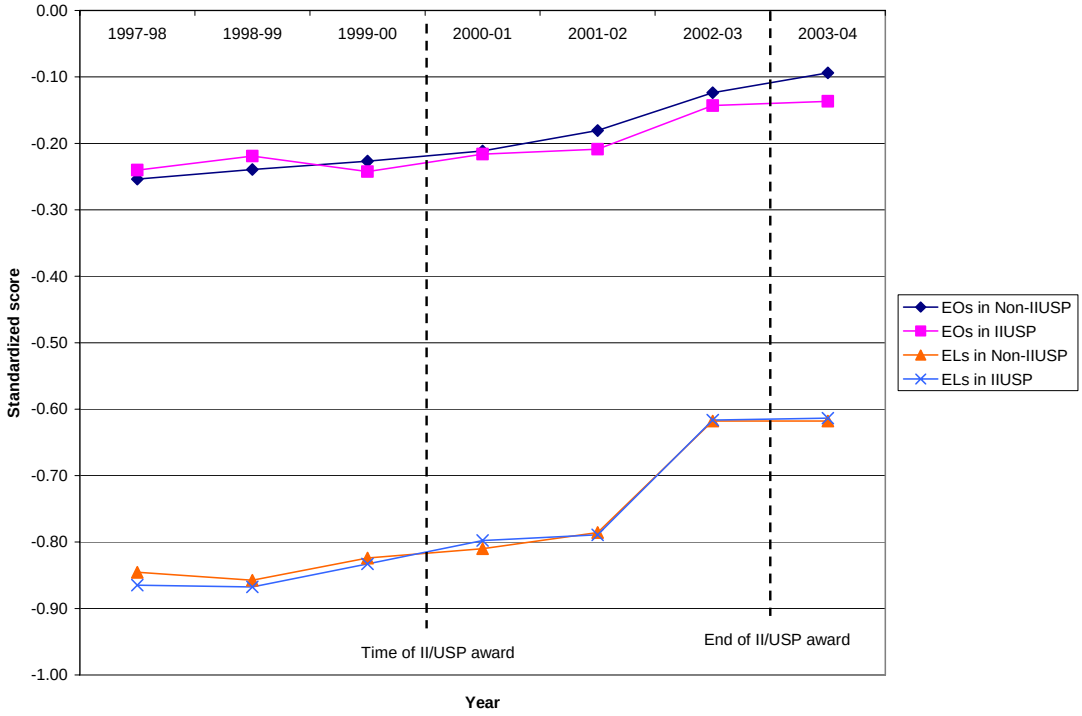
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores) for ELs and EOs in Cohort 1 II/USP and Comparison Middle Schools



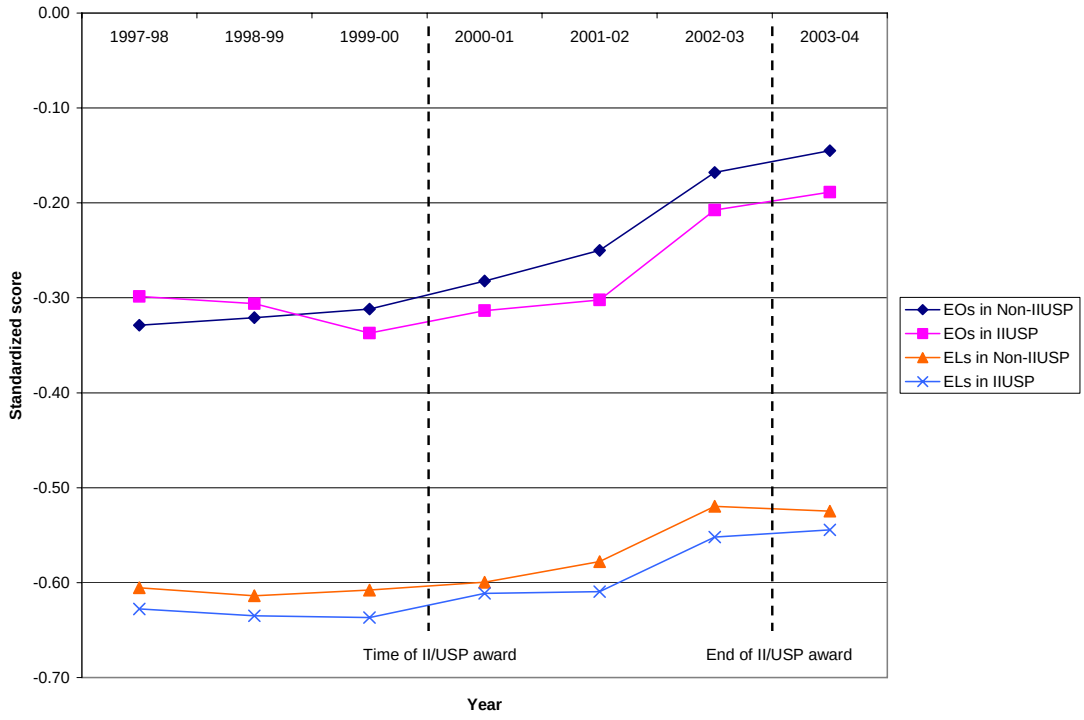
HLM Regression for ELs and EOs in Cohort 2 Middle Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.210	<.0001	0.124	0.0003
GIRL	0.102	<.0001	-0.012	<.0001
ASIAN	-0.013	0.0002	0.266	<.0001
HISPANIC	-0.297	<.0001	-0.293	<.0001
BLACK	-0.487	<.0001	-0.492	<.0001
OTHERS	-0.150	<.0001	-0.077	<.0001
ELMISN	-0.176	<.0001	-0.071	<.0001
EL	-0.592	<.0001	-0.277	<.0001
R_FEP	0.219	<.0001	0.299	<.0001
FEP	0.181	<.0001	0.217	<.0001
FLUNCH	-0.139	<.0001	-0.104	<.0001
HIGHPARED	0.219	<.0001	0.189	<.0001
PARED_MISSING	-0.027	<.0001	-0.032	<.0001
SPECED	-0.748	<.0001	-0.667	<.0001
YEAR99	0.014	0.001	0.008	0.0582
YEAR00	0.013	0.0023	0.009	0.0234
YEAR01	0.015	<.0001	0.029	<.0001
YEAR02	0.031	<.0001	0.032	<.0001
YEAR03	0.057	<.0001	0.082	<.0001
YEAR04	0.030	<.0001	0.023	<.0001
GRADE7	-0.007	0.0001	0.007	<.0001
GRADE8	-0.023	<.0001	-0.015	<.0001
PCT_MEALS	-0.304	<.0001	-0.292	<.0001
IIUSP2	0.013	0.3964	0.030	0.088
YEAR99*IIUSP2	0.007	0.2786	-0.016	0.0118
YEAR00*IIUSP2	-0.036	<.0001	-0.040	<.0001
YEAR01*IIUSP2	0.011	0.0454	-0.006	0.2692
YEAR02*IIUSP2	-0.023	<.0001	-0.021	<.0001
YEAR03*IIUSP2	0.009	0.1143	0.012	0.0206
YEAR04*IIUSP2	-0.023	<.0001	-0.004	0.4486
EL*YEAR99	-0.026	0.0005	-0.017	0.0224
EL*YEAR00	0.021	0.0032	-0.003	0.6640
EL*YEAR01	-0.001	0.8427	-0.021	0.0007
EL*YEAR02	-0.006	0.3383	-0.011	0.0950
EL*YEAR03	0.111	<.0001	-0.024	0.0002
EL*YEAR04	-0.030	<.0001	-0.028	<.0001
EL*IIUSP2	-0.033	<.0001	-0.053	<.0001
EL*YEAR99*IIUSP2	0.002	0.8293	0.017	0.1249
EL*YEAR00*IIUSP2	0.038	0.0003	0.032	0.0015
EL*YEAR01*IIUSP2	0.010	0.3059	0.023	0.0127
EL*YEAR02*IIUSP2	0.007	0.4372	0.001	0.9171
EL*YEAR03*IIUSP2	-0.004	0.6932	-0.013	0.1535
EL*YEAR04*IIUSP2	0.026	0.0051	0.017	0.0652

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores) for ELs and EOs in Cohort 2 II/USP and Comparsions Middle Schools



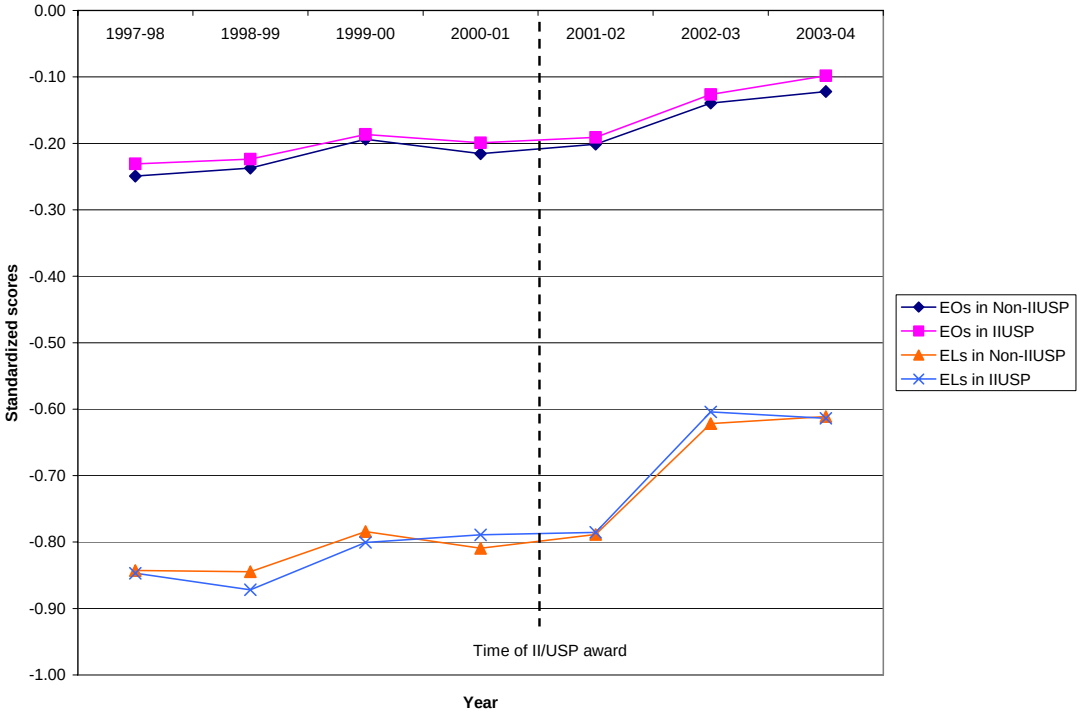
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores) for ELs and EOs in Cohort 2 II/USP and Comparison Middle Schools



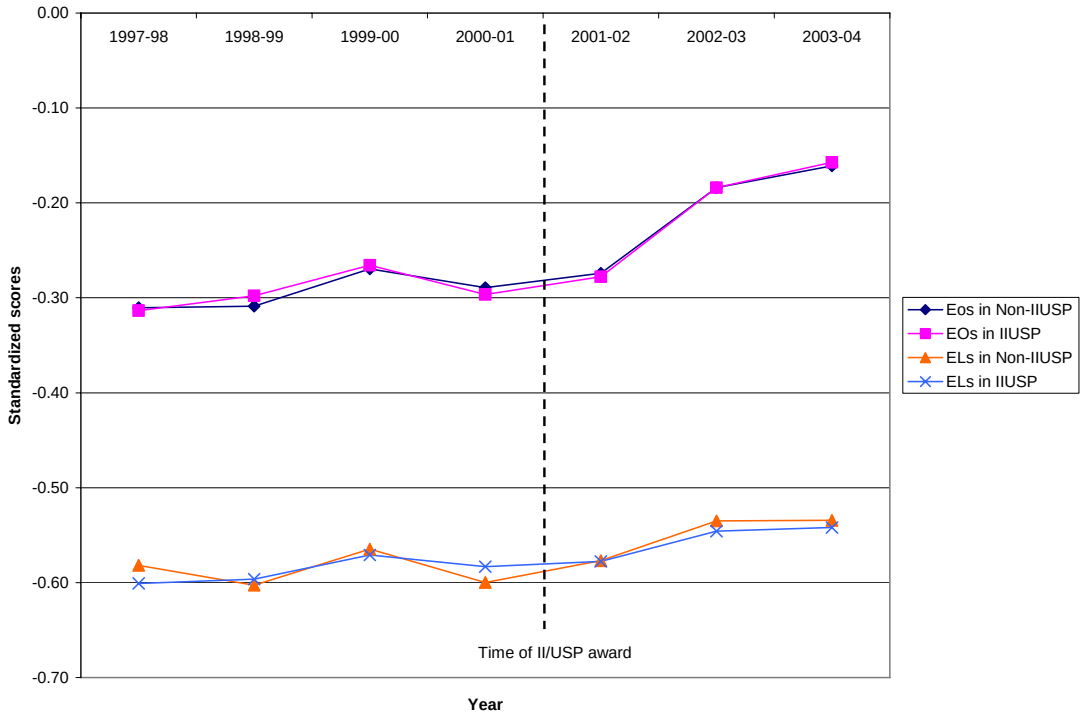
HLM Regression for ELs and EOs in Cohort 3 Middle Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.258	<.0001	0.153	<.0001
GIRL	0.104	<.0001	-0.014	<.0001
ASIAN	-0.063	<.0001	0.226	<.0001
HISPANIC	-0.316	<.0001	-0.292	<.0001
BLACK	-0.481	<.0001	-0.485	<.0001
OTHERS	-0.164	<.0001	-0.072	<.0001
ELMISN	-0.168	<.0001	-0.099	<.0001
EL	-0.594	<.0001	-0.271	<.0001
R_FEP	0.226	<.0001	0.292	<.0001
FEP	0.165	<.0001	0.180	<.0001
FLUNCH	-0.154	<.0001	-0.119	<.0001
HIGHPARED	0.213	<.0001	0.183	<.0001
PARED_MISSING	-0.031	<.0001	-0.031	<.0001
SPECED	-0.768	<.0001	-0.690	<.0001
YEAR99	0.012	0.0014	0.002	0.6754
YEAR00	0.043	<.0001	0.039	<.0001
YEAR01	-0.022	<.0001	-0.020	<.0001
YEAR02	0.015	<.0001	0.015	<.0001
YEAR03	0.062	<.0001	0.090	<.0001
YEAR04	0.017	<.0001	0.023	<.0001
GRADE7	-0.006	0.0003	0.007	<.0001
GRADE8	-0.023	<.0001	-0.017	<.0001
PCT_MEALS	-0.349	<.0001	-0.312	<.0001
IIUSP3	0.018	0.3064	-0.003	0.8667
YEAR99*IIUSP3	-0.005	0.5416	0.014	0.0631
YEAR00*IIUSP3	-0.006	0.3817	-0.007	0.3146
YEAR01*IIUSP3	0.009	0.1569	-0.011	0.0911
YEAR02*IIUSP3	-0.006	0.3338	0.003	0.6035
YEAR03*IIUSP3	0.003	0.6646	0.004	0.5631
YEAR04*IIUSP3	0.011	0.104	0.004	0.5407
EL*YEAR99	-0.014	0.0401	-0.023	0.0005
EL*YEAR00	0.017	0.0068	-0.001	0.8715
EL*YEAR01	-0.003	0.5690	-0.015	0.0055
EL*YEAR02	0.006	0.2793	0.008	0.1594
EL*YEAR03	0.105	<.0001	-0.048	<.0001
EL*YEAR04	-0.007	0.1827	-0.022	<.0001
EL*IIUSP3	-0.023	0.0258	-0.016	0.1005
EL*YEAR99*IIUSP3	-0.018	0.2009	0.012	0.3908
EL*YEAR00*IIUSP3	0.017	0.1856	-0.006	0.6401
EL*YEAR01*IIUSP3	0.027	0.022	0.034	0.0030
EL*YEAR02*IIUSP3	-0.010	0.3652	-0.020	0.0652
EL*YEAR03*IIUSP3	0.012	0.3007	-0.014	0.2015
EL*YEAR04*IIUSP3	-0.031	0.0054	-0.001	0.9437

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores) for ELs and EOs in Cohort 3 II/USP and Comparison Middle Schools



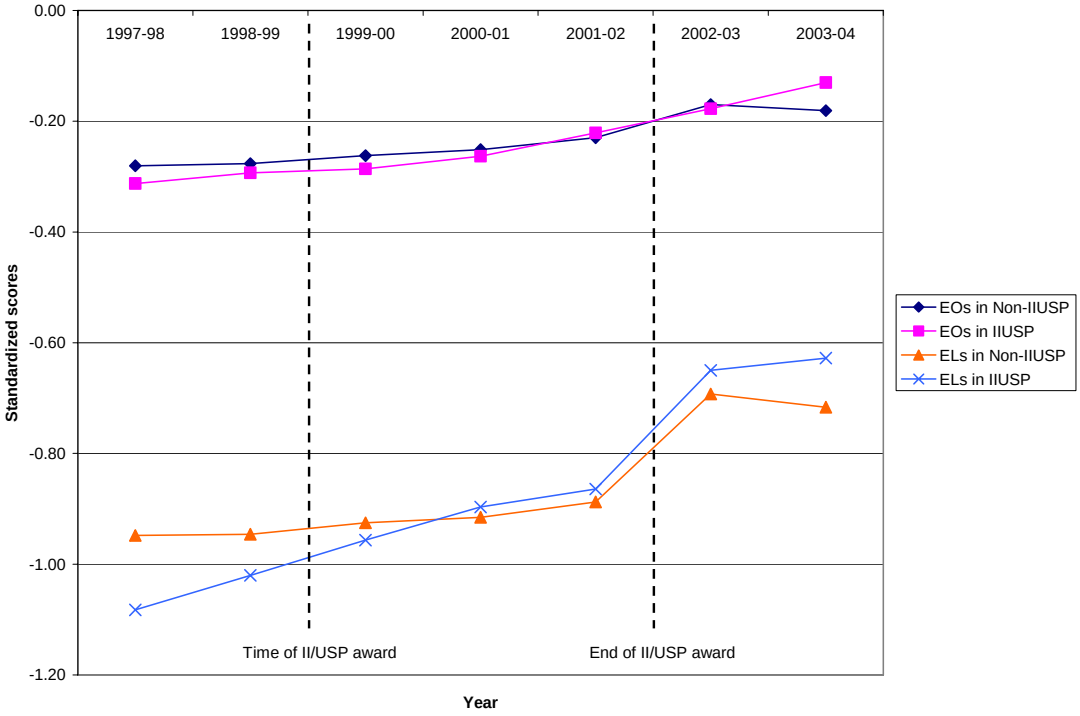
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores) for ELs and EOs in Cohort 3 II/USP and Comparison Middle Schools



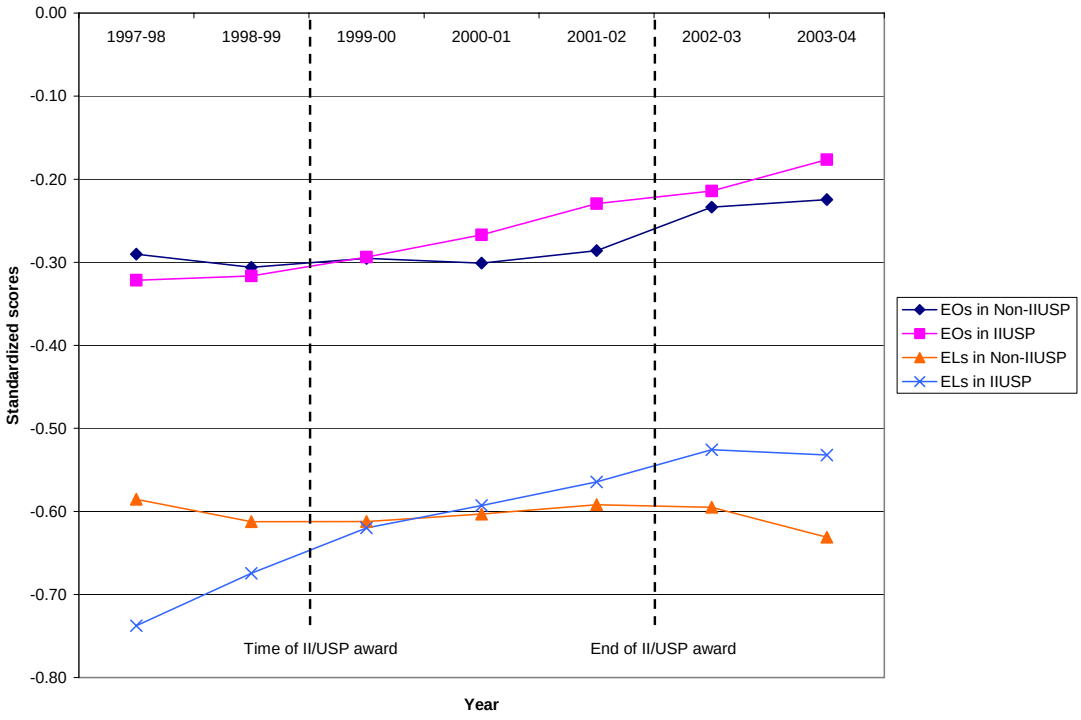
HLM Regression for ELs and EOs in Cohort 1 High Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.095	0.0069	0.106	0.0008
GIRL	0.110	<.0001	-0.066	<.0001
ASIAN	-0.117	<.0001	0.256	<.0001
HISPANIC	-0.328	<.0001	-0.326	<.0001
BLACK	-0.533	<.0001	-0.529	<.0001
OTHERS	-0.212	<.0001	-0.120	<.0001
ELMISN	-0.052	<.0001	0.047	<.0001
EL	-0.668	<.0001	-0.295	<.0001
R_FEP	0.071	<.0001	0.112	<.0001
FEP	0.097	<.0001	0.106	<.0001
FLUNCH	-0.079	<.0001	-0.041	<.0001
HIGHPARED	0.242	<.0001	0.194	<.0001
PARED_MISSING	-0.020	<.0001	-0.025	<.0001
SPECED	-0.748	<.0001	-0.611	<.0001
YEAR99	0.004	0.4391	-0.016	0.0021
YEAR00	0.014	0.0022	0.011	0.0175
YEAR01	0.011	0.0122	-0.006	0.1821
YEAR02	0.021	<.0001	0.015	0.0004
YEAR03	0.060	<.0001	0.052	<.0001
YEAR04	-0.011	0.0113	0.009	0.0301
GRADE10	0.004	0.0245	0.016	<.0001
GRADE11	0.017	<.0001	0.002	0.2395
PCT_MEALS	-0.119	0.055	-0.171	0.0021
IIUSP1	-0.032	0.2043	-0.031	0.1642
YEAR99*IIUSP1	0.015	0.0842	0.021	0.0142
YEAR00*IIUSP1	-0.007	0.357	0.012	0.104
YEAR01*IIUSP1	0.012	0.1001	0.032	<.0001
YEAR02*IIUSP1	0.021	0.0031	0.022	0.0012
YEAR03*IIUSP1	-0.017	0.0195	-0.037	<.0001
YEAR04*IIUSP1	0.058	<.0001	0.029	<.0001
EL*YEAR99	-0.002	0.8601	-0.011	0.2584
EL*YEAR00	0.007	0.4579	-0.011	0.2315
EL*YEAR01	-0.001	0.8792	0.015	0.0743
EL*YEAR02	0.006	0.4625	-0.004	0.6175
EL*YEAR03	0.135	<.0001	-0.055	<.0001
EL*YEAR04	-0.013	0.1072	-0.046	<.0001
EL*IIUSP1	-0.102	<.0001	-0.121	<.0001
EL*YEAR99*IIUSP1	0.045	0.0126	0.069	<.0001
EL*YEAR00*IIUSP1	0.050	0.0011	0.042	0.0043
EL*YEAR01*IIUSP1	0.038	0.0066	-0.014	0.2903
EL*YEAR02*IIUSP1	-0.016	0.2362	-0.005	0.7019
EL*YEAR03*IIUSP1	0.036	0.008	0.078	<.0001
EL*YEAR04*IIUSP1	-0.012	0.3653	0.001	0.9126

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores) for ELs and EOs in Cohort 1 II/USP and Comparison High Schools



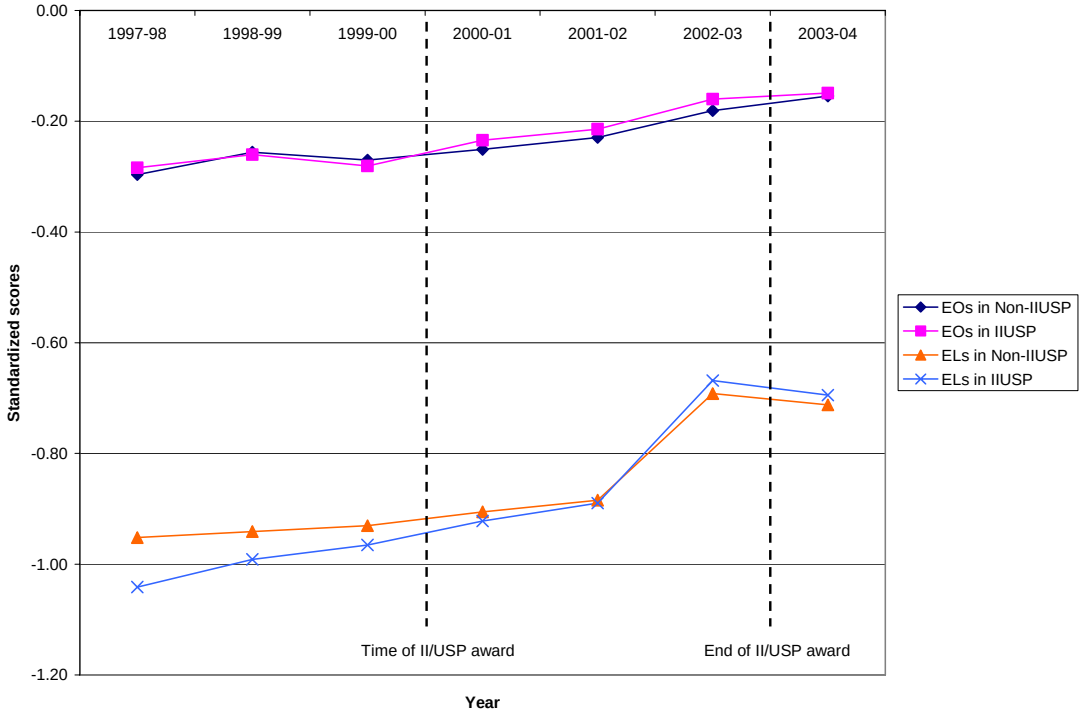
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores) for ELs and EOs in Cohort 1 II/USP and Comparison High Schools



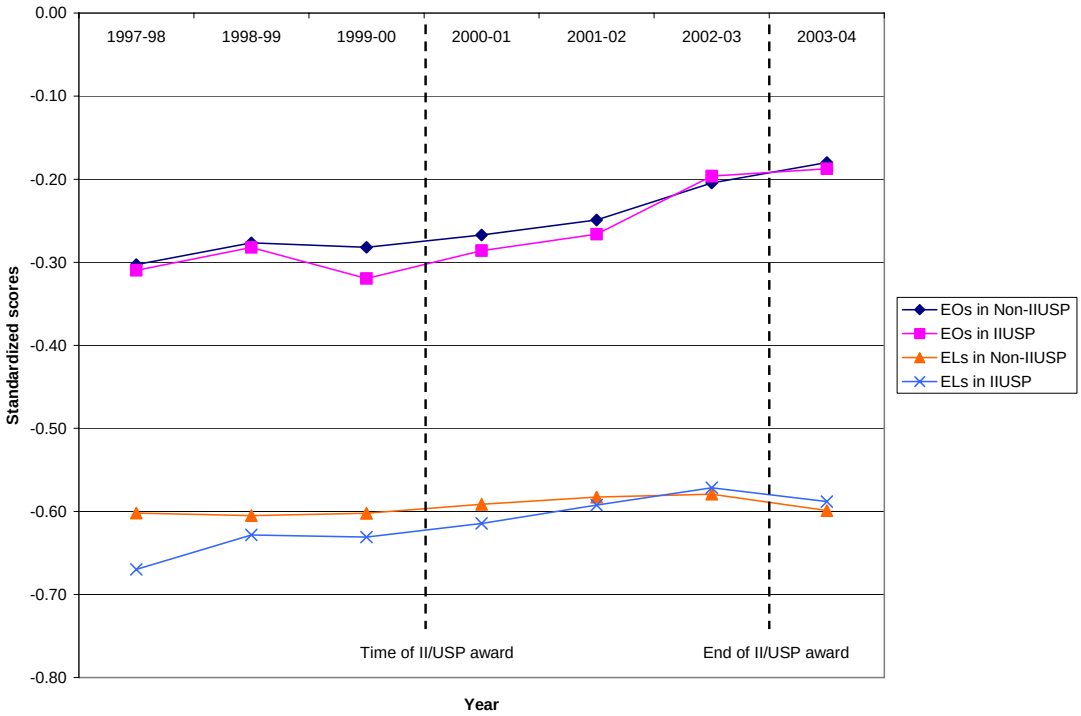
HLM Regression for ELs and EOs in Cohort 2 High Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.069	0.0042	0.079	0.0007
GIRL	0.116	<.0001	-0.067	<.0001
ASIAN	-0.080	<.0001	0.251	<.0001
HISPANIC	-0.311	<.0001	-0.324	<.0001
BLACK	-0.515	<.0001	-0.521	<.0001
OTHERS	-0.180	<.0001	-0.107	<.0001
ELMISN	-0.008	0.3084	0.057	<.0001
EL	-0.656	<.0001	-0.299	<.0001
R_FEP	0.078	<.0001	0.137	<.0001
FEP	0.127	<.0001	0.148	<.0001
FLUNCH	-0.095	<.0001	-0.059	<.0001
HIGHPARED	0.244	<.0001	0.204	<.0001
PARED_MISSING	-0.009	<.0001	-0.019	<.0001
SPECED	-0.771	<.0001	-0.628	<.0001
YEAR99	0.040	<.0001	0.026	<.0001
YEAR00	-0.014	<.0001	-0.005	0.1006
YEAR01	0.019	<.0001	0.015	<.0001
YEAR02	0.022	<.0001	0.018	<.0001
YEAR03	0.049	<.0001	0.044	<.0001
YEAR04	0.026	<.0001	0.025	<.0001
GRADE10	-0.007	<.0001	0.006	<.0001
GRADE11	-0.001	0.5305	-0.010	<.0001
PCT_MEALS	-0.134	0.0046	-0.144	0.0018
IIUSP2	0.012	0.5041	-0.007	0.7007
YEAR99*IIUSP2	-0.017	0.0082	0.001	0.8329
YEAR00*IIUSP2	-0.006	0.2815	-0.032	<.0001
YEAR01*IIUSP2	0.027	<.0001	0.019	0.0002
YEAR02*IIUSP2	-0.001	0.8082	0.002	0.7641
YEAR03*IIUSP2	0.006	0.2786	0.025	<.0001
YEAR04*IIUSP2	-0.016	0.002	-0.016	0.0015
EL*YEAR99	-0.030	<.0001	-0.029	<.0001
EL*YEAR00	0.025	0.0001	0.008	0.1972
EL*YEAR01	0.006	0.335	-0.004	0.4853
EL*YEAR02	-0.001	0.928	-0.009	0.1129
EL*YEAR03	0.144	<.0001	-0.041	<.0001
EL*YEAR04	-0.047	<.0001	-0.044	<.0001
EL*IIUSP2	-0.102	<.0001	-0.061	<.0001
EL*YEAR99*IIUSP2	0.056	<.0001	0.043	0.0007
EL*YEAR00*IIUSP2	0.021	0.065	0.027	0.019
EL*YEAR01*IIUSP2	-0.009	0.395	-0.013	0.2092
EL*YEAR02*IIUSP2	0.013	0.2249	0.012	0.2506
EL*YEAR03*IIUSP2	0.023	0.0235	-0.008	0.4365
EL*YEAR04*IIUSP2	0.010	0.3284	0.019	0.0576

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores) for ELs and EOs in Cohort 2 II/USP and Comparison High Schools



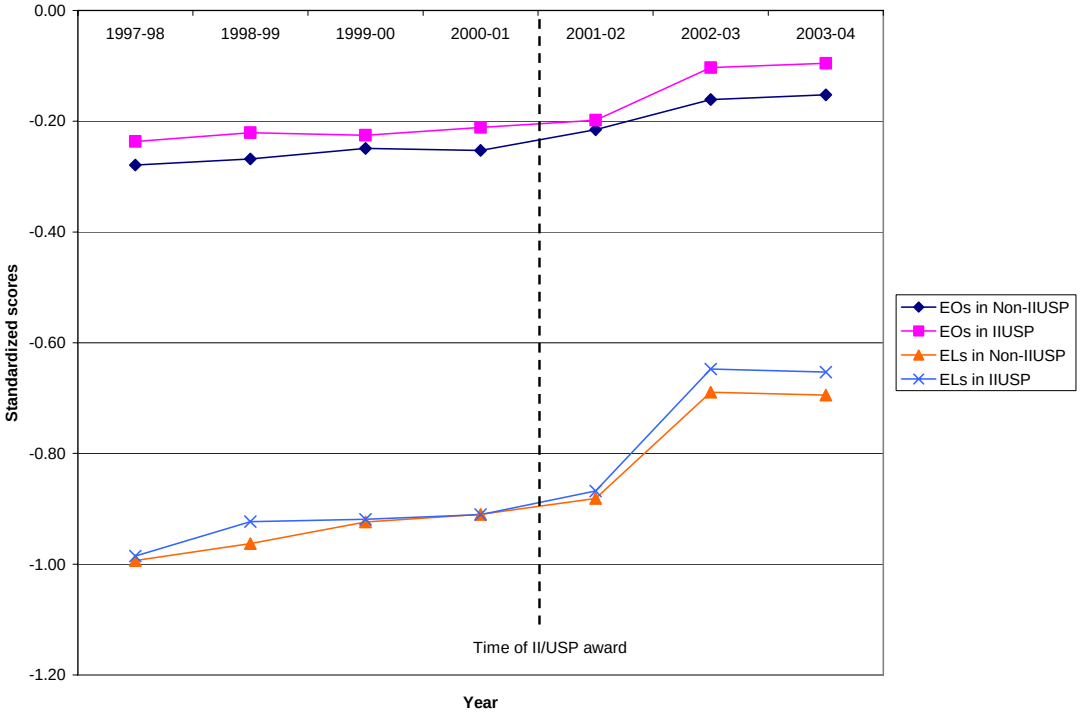
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores) for ELs and EOs in Cohort 2 II/USP and Comparison High Schools



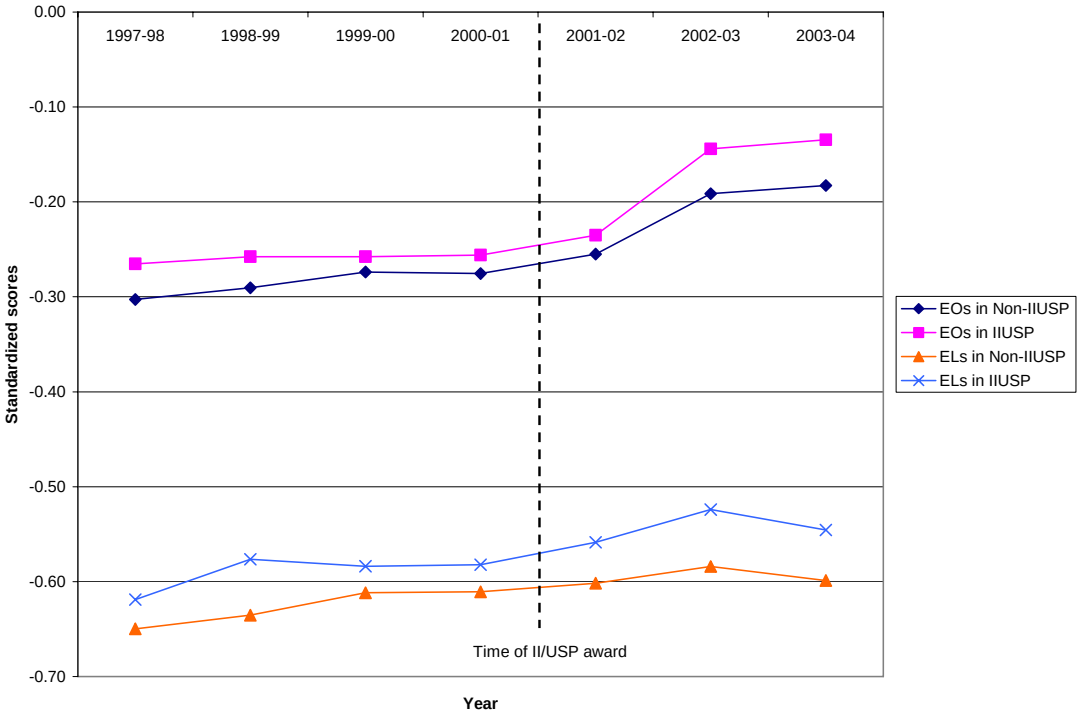
HLM Regression for ELs and EOs in Cohort 3 High Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.051	0.0189	0.032	0.1613
GIRL	0.114	<.0001	-0.073	<.0001
ASIAN	-0.093	<.0001	0.248	<.0001
HISPANIC	-0.318	<.0001	-0.325	<.0001
BLACK	-0.519	<.0001	-0.524	<.0001
OTHERS	-0.195	<.0001	-0.112	<.0001
ELMISN	-0.129	<.0001	-0.014	0.0228
EL	-0.714	<.0001	-0.347	<.0001
R_FEP	0.078	<.0001	0.131	<.0001
FEP	0.100	<.0001	0.123	<.0001
FLUNCH	-0.101	<.0001	-0.062	<.0001
HIGHPARED	0.243	<.0001	0.205	<.0001
PARED_MISSING	-0.008	<.0001	-0.019	<.0001
SPECED	-0.782	<.0001	-0.641	<.0001
YEAR99	0.011	0.0006	0.012	<.0001
YEAR00	0.019	<.0001	0.017	<.0001
YEAR01	-0.003	0.2072	-0.002	0.5452
YEAR02	0.037	<.0001	0.020	<.0001
YEAR03	0.055	<.0001	0.064	<.0001
YEAR04	0.008	0.0012	0.009	0.0007
GRADE10	-0.010	<.0001	0.001	0.4309
GRADE11	-0.006	<.0001	-0.017	<.0001
PCT_MEALS	-0.033	0.4647	-0.028	0.5506
IIUSP3	0.043	0.0202	0.038	0.0525
YEAR99*IIUSP3	0.005	0.4168	-0.005	0.4063
YEAR00*IIUSP3	-0.023	<.0001	-0.017	0.0010
YEAR01*IIUSP3	0.017	0.0002	0.003	0.4966
YEAR02*IIUSP3	-0.024	<.0001	0.001	0.9014
YEAR03*IIUSP3	0.040	<.0001	0.027	<.0001
YEAR04*IIUSP3	-0.001	0.8899	0.001	0.8065
EL*YEAR99	0.020	0.0032	0.002	0.7573
EL*YEAR00	0.020	0.0006	0.007	0.2281
EL*YEAR01	0.017	0.0018	0.003	0.6130
EL*YEAR02	-0.009	0.1088	-0.012	0.0275
EL*YEAR03	0.137	<.0001	-0.046	<.0001
EL*YEAR04	-0.014	0.0087	-0.024	<.0001
EL*IIUSP3	-0.035	<.0001	-0.007	0.4145
EL*YEAR99*IIUSP3	0.027	0.0203	0.033	0.0038
EL*YEAR00*IIUSP3	-0.012	0.2502	-0.014	0.1574
EL*YEAR01*IIUSP3	-0.022	0.0176	-0.003	0.7664
EL*YEAR02*IIUSP3	0.038	<.0001	0.014	0.1118
EL*YEAR03*IIUSP3	-0.012	0.1820	-0.010	0.2396
EL*YEAR04*IIUSP3	0.000	0.9847	-0.008	0.3583

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores) for ELs and EOs in Cohort 3 II/USP and Comparison High Schools



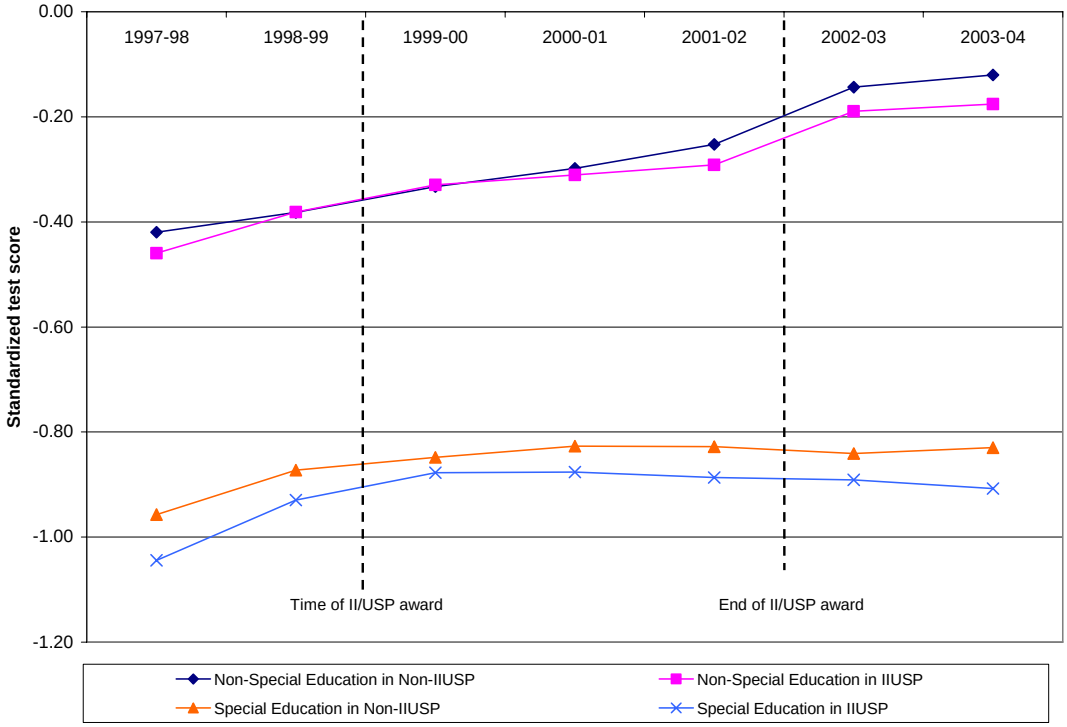
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores) for ELs and EOs in Cohort 3 II/USP and Comparison High Schools



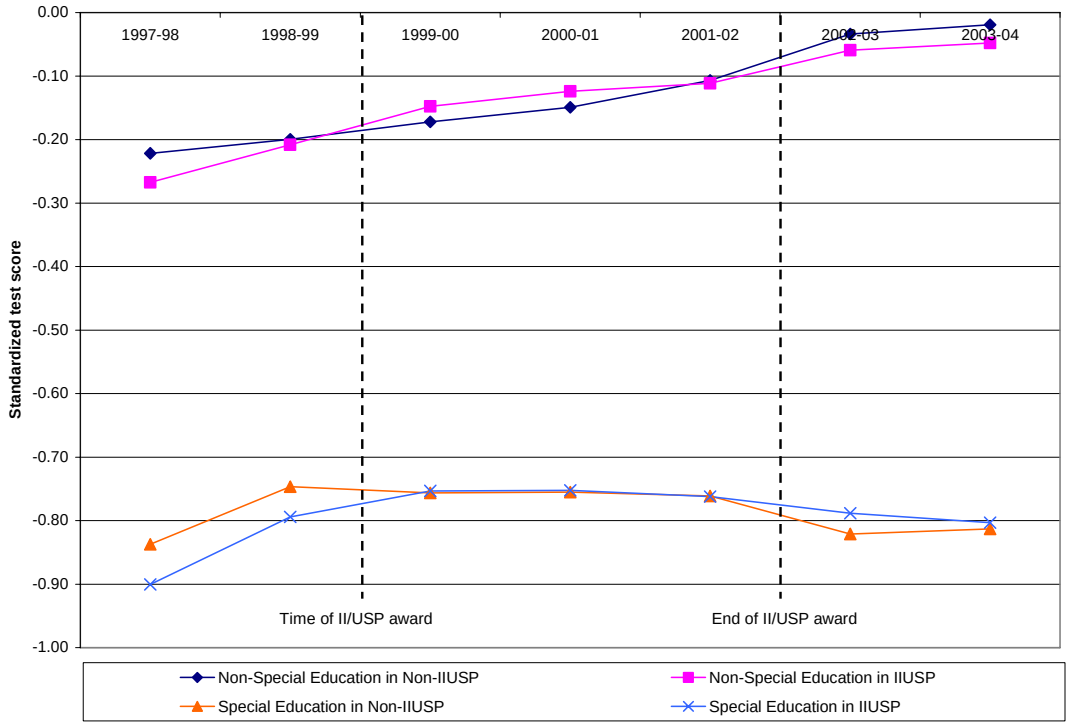
HLM Regression for Special Education and Non-Special Education Students in Cohort 1 Elementary Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.111	<.0001	0.058	0.0219
GIRL	0.108	<.0001	-0.015	<.0001
ASIAN	-0.007	0.0328	0.196	<.0001
HISPANIC	-0.259	<.0001	-0.235	<.0001
BLACK	-0.406	<.0001	-0.454	<.0001
OTHERS	-0.091	<.0001	-0.052	<.0001
ELMISN	-0.125	<.0001	-0.090	<.0001
EL	-0.374	<.0001	-0.174	<.0001
R_FEP	0.344	<.0001	0.483	<.0001
FEP	0.184	<.0001	0.230	<.0001
FLUNCH	-0.174	<.0001	-0.143	<.0001
HIGHPARED	0.222	<.0001	0.203	<.0001
PARED_MISSING	-0.013	<.0001	-0.034	<.0001
SPECED	-0.538	<.0001	-0.616	<.0001
YEAR99	0.038	<.0001	0.022	<.0001
YEAR00	0.049	<.0001	0.028	<.0001
YEAR01	0.035	<.0001	0.023	<.0001
YEAR02	0.046	<.0001	0.043	<.0001
YEAR03	0.109	<.0001	0.073	<.0001
YEAR04	0.023	<.0001	0.015	<.0001
GRADE3	-0.031	<.0001	-0.023	<.0001
GRADE4	-0.062	<.0001	-0.061	<.0001
GRADE5	-0.090	<.0001	-0.102	<.0001
PCT_MEALS	-0.191	<.0001	-0.097	0.0028
IIUSP1	-0.040	0.0003	-0.046	0.0002
YEAR99*IIUSP1	0.041	<.0001	0.037	<.0001
YEAR00*IIUSP1	0.003	0.5613	0.033	<.0001
YEAR01*IIUSP1	-0.016	<.0001	0.001	0.7761
YEAR02*IIUSP1	-0.026	<.0001	-0.030	<.0001
YEAR03*IIUSP1	-0.007	0.1117	-0.021	<.0001
YEAR04*IIUSP1	-0.010	0.0239	-0.004	0.4237
SPECED*YEAR99	0.047	<.0001	0.069	<.0001
SPECED*YEAR00	-0.025	0.0041	-0.037	<.0001
SPECED*YEAR01	-0.013	0.1009	-0.021	0.0087
SPECED*YEAR02	-0.047	<.0001	-0.049	<.0001
SPECED*YEAR03	-0.122	<.0001	-0.133	<.0001
SPECED*YEAR04	-0.012	0.1478	-0.007	0.4513
SPECED*IIUSP1	-0.047	<.0001	-0.018	0.1421
SPECED*YEAR99*IIUSP1	-0.010	0.5608	-0.022	0.2269
SPECED*YEAR00*IIUSP1	0.025	0.1335	0.018	0.2904
SPECED*YEAR01*IIUSP1	-0.005	0.7546	-0.002	0.9025
SPECED*YEAR02*IIUSP1	0.017	0.2704	0.027	0.0926
SPECED*YEAR03*IIUSP1	0.015	0.3291	0.054	0.0007
SPECED*YEAR04*IIUSP1	-0.017	0.2314	-0.019	0.2037

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Special Education and Non-Special Education Students in Cohort 1 II/USP and Comparison Elementary Schools



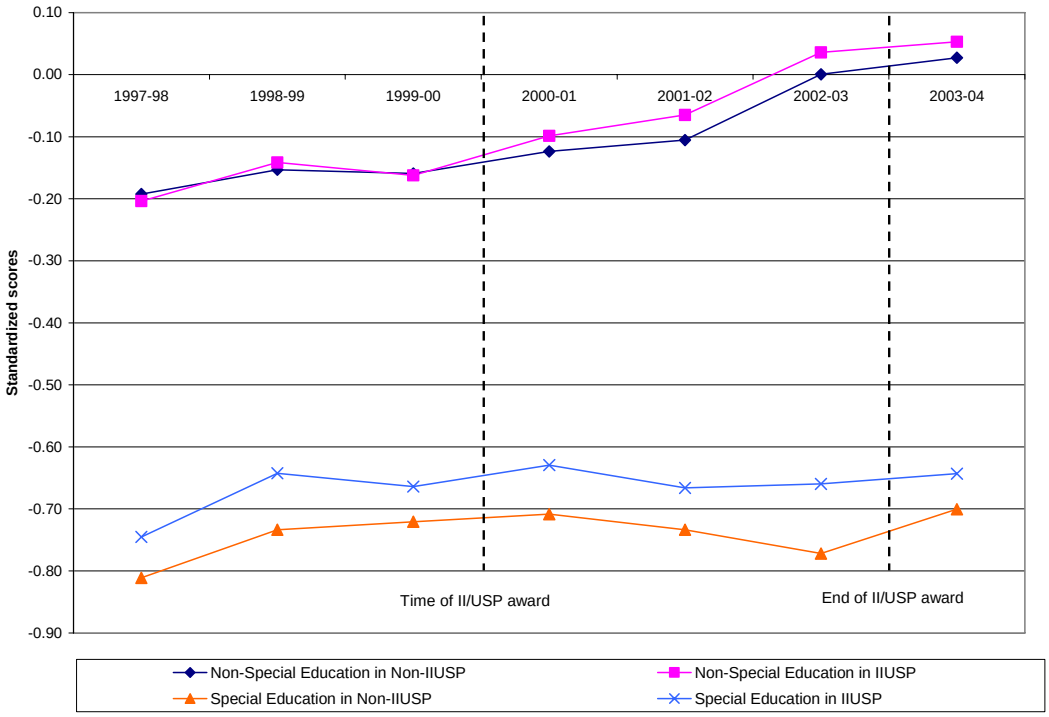
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Special Education and Non-Special Education Students in Cohort 1 II/USP and Comparison Elementary Schools



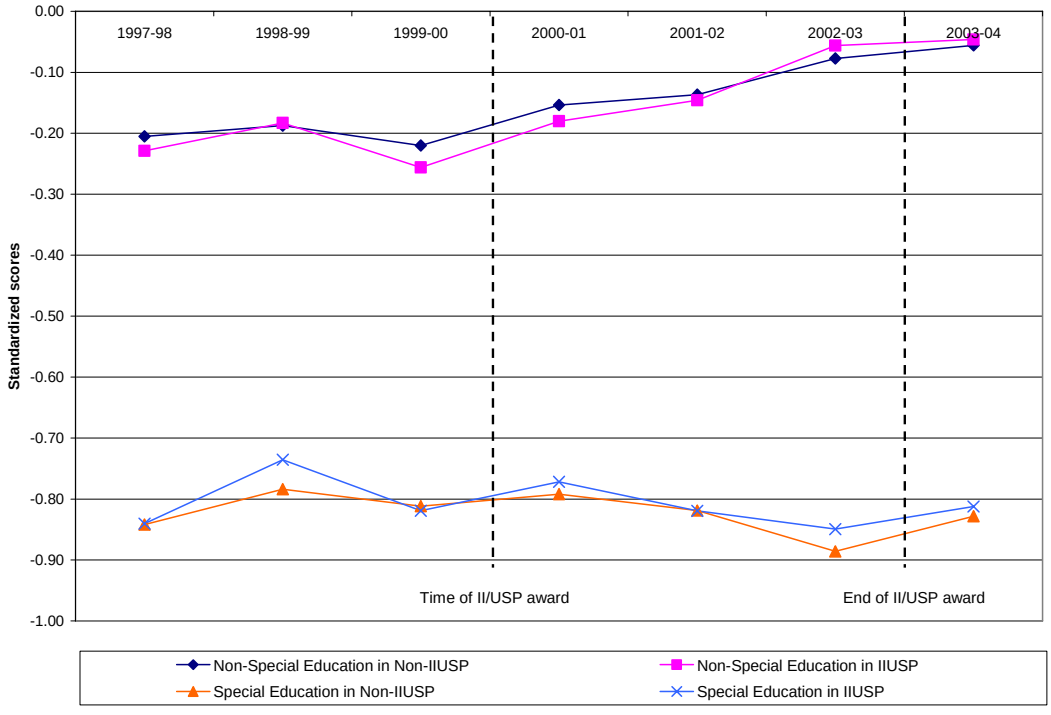
HLM Regression for Special Education and Non-Special Education Students in Cohort 2 Elementary Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.162	<.0001	0.116	0.0004
GIRL	0.114	<.0001	-0.011	<.0001
ASIAN	-0.030	<.0001	0.165	<.0001
HISPANIC	-0.268	<.0001	-0.240	<.0001
BLACK	-0.420	<.0001	-0.467	<.0001
OTHERS	-0.124	<.0001	-0.077	<.0001
ELMISN	-0.168	<.0001	-0.113	<.0001
EL	-0.371	<.0001	-0.180	<.0001
R_FEP	0.363	<.0001	0.498	<.0001
FEP	0.196	<.0001	0.240	<.0001
FLUNCH	-0.177	<.0001	-0.145	<.0001
HIGHPARED	0.237	<.0001	0.218	<.0001
PARED_MISSING	-0.016	<.0001	-0.036	<.0001
SPECED	-0.619	<.0001	-0.637	<.0001
YEAR99	0.039	<.0001	0.018	0.0002
YEAR00	-0.006	0.1585	-0.033	<.0001
YEAR01	0.036	<.0001	0.067	<.0001
YEAR02	0.019	<.0001	0.017	<.0001
YEAR03	0.106	<.0001	0.059	<.0001
YEAR04	0.027	<.0001	0.021	<.0001
GRADE3	-0.028	<.0001	-0.039	<.0001
GRADE4	-0.057	<.0001	-0.070	<.0001
GRADE5	-0.075	<.0001	-0.095	<.0001
PCT_MEALS	-0.268	<.0001	-0.160	0.0002
IIUSP2	-0.011	0.3922	-0.024	0.1169
YEAR99*IIUSP2	0.023	0.0003	0.028	<.0001
YEAR00*IIUSP2	-0.015	0.0118	-0.040	<.0001
YEAR01*IIUSP2	0.028	<.0001	0.009	0.0986
YEAR02*IIUSP2	0.015	0.0063	0.017	0.0021
YEAR03*IIUSP2	-0.005	0.3708	0.031	<.0001
YEAR04*IIUSP2	-0.010	0.0803	-0.011	0.0495
SPECED*YEAR99	0.038	0.0294	0.040	0.0221
SPECED*YEAR00	0.019	0.2608	0.005	0.7603
SPECED*YEAR01	-0.023	0.1013	-0.047	0.0011
SPECED*YEAR02	-0.044	0.0034	-0.044	0.0034
SPECED*YEAR03	-0.144	<.0001	-0.126	<.0001
SPECED*YEAR04	0.045	0.0029	0.036	0.0199
SPECED*IIUSP2	0.077	<.0001	0.025	0.1123
SPECED*YEAR99*IIUSP2	0.002	0.9377	0.019	0.4395
SPECED*YEAR00*IIUSP2	-0.019	0.3979	-0.016	0.4899
SPECED*YEAR01*IIUSP2	-0.006	0.7488	0.019	0.3364
SPECED*YEAR02*IIUSP2	-0.027	0.1784	-0.038	0.0603
SPECED*YEAR03*IIUSP2	0.050	0.0144	0.006	0.7635
SPECED*YEAR04*IIUSP2	-0.045	0.0208	-0.009	0.6419

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Special Education and Non-Special Education Students in Cohort 2 II/USP and Comparison Elementary Schools



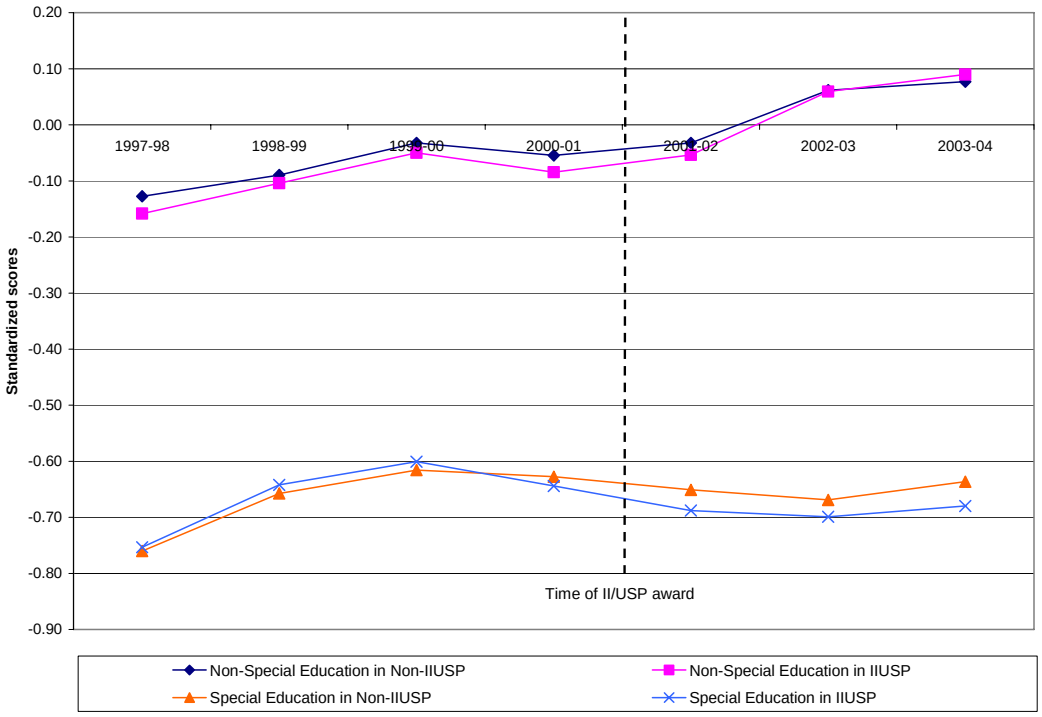
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Special Education and Non-Special Education Students in Cohort 2 II/USP and Comparison Elementary Schools



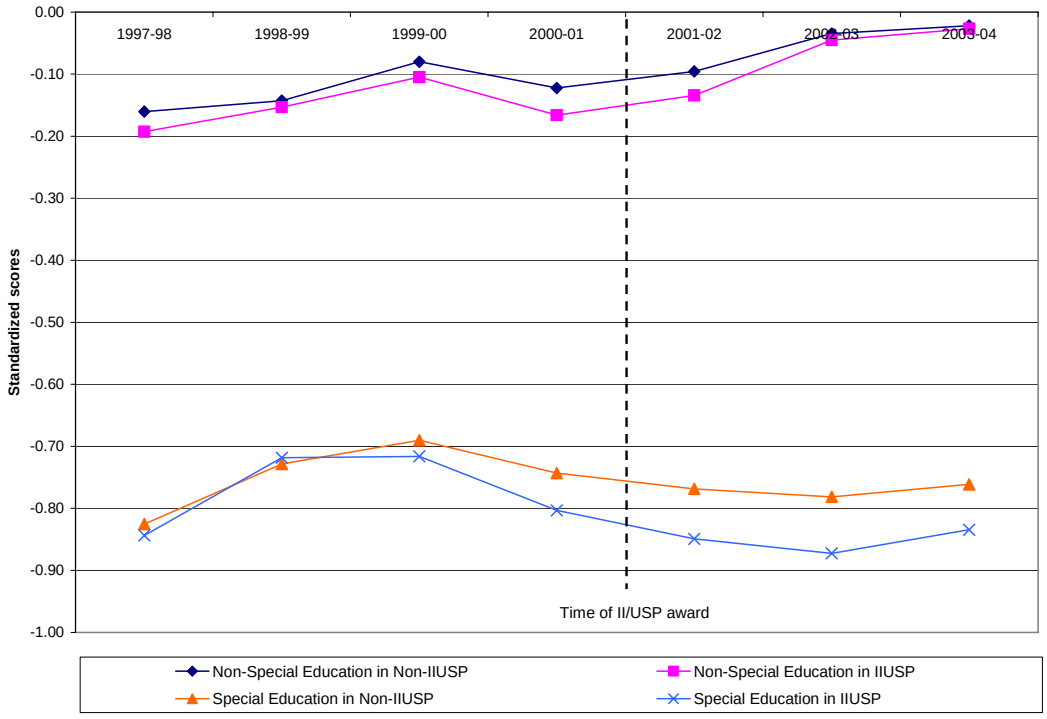
HLM Regression for Special Education and Non-Special Education Students in Cohort 3 Elementary Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.173	<.0001	0.106	<.0001
GIRL	0.118	<.0001	-0.014	<.0001
ASIAN	-0.009	0.0089	0.184	<.0001
HISPANIC	-0.254	<.0001	-0.230	<.0001
BLACK	-0.410	<.0001	-0.449	<.0001
OTHERS	-0.100	<.0001	-0.055	<.0001
ELMISN	-0.140	<.0001	-0.104	<.0001
EL	-0.391	<.0001	-0.195	<.0001
R_FEP	0.317	<.0001	0.450	<.0001
FEP	0.164	<.0001	0.216	<.0001
FLUNCH	-0.188	<.0001	-0.158	<.0001
HIGHPARED	0.239	<.0001	0.221	<.0001
PARED_MISSING	-0.003	0.052	-0.032	<.0001
SPECED	-0.633	<.0001	-0.665	<.0001
YEAR99	0.038	<.0001	0.017	<.0001
YEAR00	0.057	<.0001	0.063	<.0001
YEAR01	-0.022	<.0001	-0.042	<.0001
YEAR02	0.022	<.0001	0.027	<.0001
YEAR03	0.094	<.0001	0.061	<.0001
YEAR04	0.015	<.0001	0.013	<.0001
GRADE3	-0.018	<.0001	-0.024	<.0001
GRADE4	-0.044	<.0001	-0.052	<.0001
GRADE5	-0.061	<.0001	-0.079	<.0001
PCT_MEALS	-0.256	<.0001	-0.129	<.0001
IIUSP3	-0.031	0.0027	-0.033	0.0065
YEAR99*IIUSP3	0.016	0.0025	0.022	<.0001
YEAR00*IIUSP3	-0.003	0.5669	-0.014	0.0045
YEAR01*IIUSP3	-0.012	0.0067	-0.019	<.0001
YEAR02*IIUSP3	0.008	0.0621	0.005	0.2409
YEAR03*IIUSP3	0.019	<.0001	0.028	<.0001
YEAR04*IIUSP3	0.015	0.0008	0.005	0.2484
SPECED*YEAR99	0.065	<.0001	0.080	<.0001
SPECED*YEAR00	-0.016	0.1775	-0.025	0.0338
SPECED*YEAR01	0.010	0.2828	-0.011	0.2707
SPECED*YEAR02	-0.046	<.0001	-0.052	<.0001
SPECED*YEAR03	-0.112	<.0001	-0.074	<.0001
SPECED*YEAR04	0.017	0.0707	0.007	0.4532
SPECED*IIUSP3	0.038	0.0048	0.014	0.281
SPECED*YEAR99*IIUSP3	-0.008	0.7035	0.006	0.7535
SPECED*YEAR00*IIUSP3	0.002	0.8967	-0.022	0.2481
SPECED*YEAR01*IIUSP3	-0.019	0.2296	-0.015	0.349
SPECED*YEAR02*IIUSP3	-0.029	0.0691	-0.026	0.1056
SPECED*YEAR03*IIUSP3	-0.013	0.4164	-0.039	0.0167
SPECED*YEAR04*IIUSP3	-0.028	0.0663	0.013	0.422

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Special Education and Non-Special Education Students in Cohort 3 II/USP and Comparison Elementary Schools



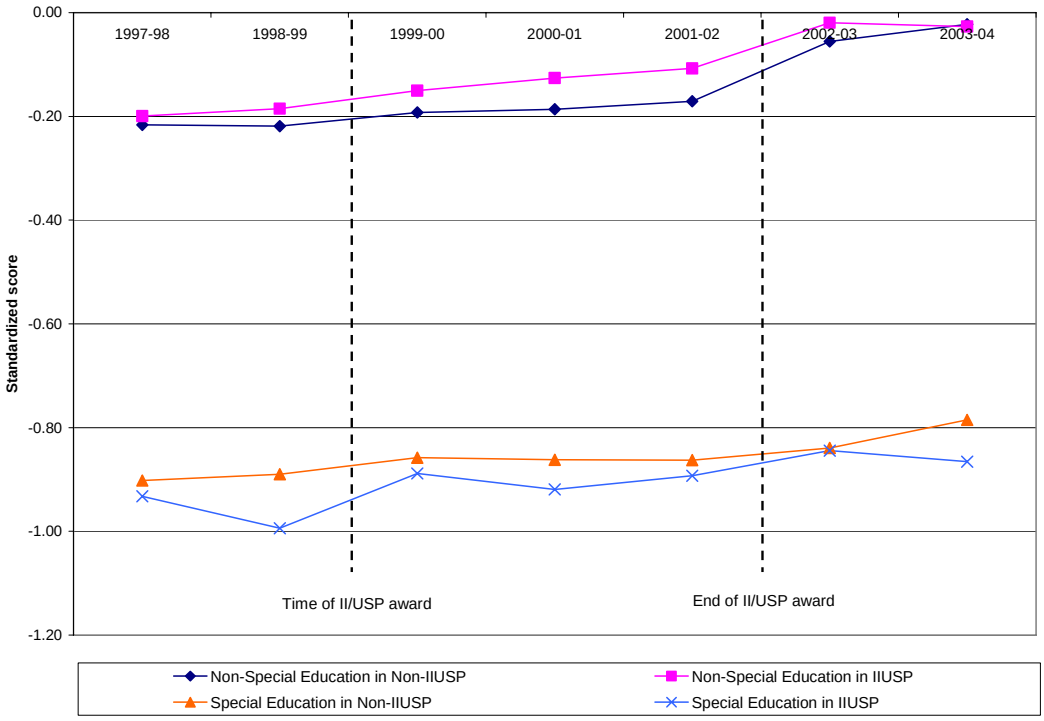
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Special Education and Non-Special Education Students in Cohort 3 II/USP and Comparison Elementary Schools



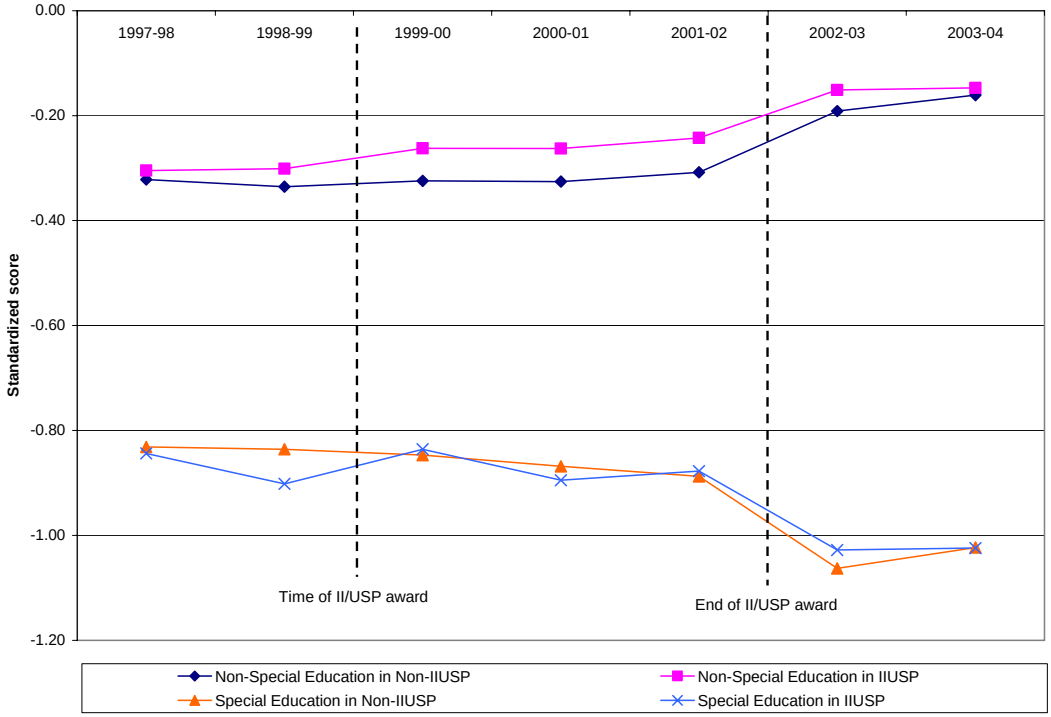
HLM Regression for Special Education and Non-Special Education Students in Cohort 1 Middle Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.213	<.0001	0.056	0.089
GIRL	0.103	<.0001	-0.007	<.0001
ASIAN	-0.065	<.0001	0.262	<.0001
HISPANIC	-0.339	<.0001	-0.316	<.0001
BLACK	-0.518	<.0001	-0.510	<.0001
OTHERS	-0.185	<.0001	-0.096	<.0001
ELMISN	-0.185	<.0001	-0.082	<.0001
EL	-0.564	<.0001	-0.310	<.0001
R_FEP	0.214	<.0001	0.281	<.0001
FEP	0.166	<.0001	0.183	<.0001
FLUNCH	-0.128	<.0001	-0.094	<.0001
HIGHPARED	0.217	<.0001	0.181	<.0001
PARED_MISSING	-0.037	<.0001	-0.037	<.0001
SPECED	-0.686	<.0001	-0.510	<.0001
YEAR99	-0.003	0.4308	-0.014	<.0001
YEAR00	0.026	<.0001	0.011	<.0001
YEAR01	0.006	0.0401	-0.001	0.6214
YEAR02	0.015	<.0001	0.018	<.0001
YEAR03	0.115	<.0001	0.117	<.0001
YEAR04	0.033	<.0001	0.031	<.0001
GRADE7	-0.010	<.0001	0.020	<.0001
GRADE8	-0.026	<.0001	-0.002	0.3531
PCT_MEALS	-0.321	<.0001	-0.206	<.0001
IIUSP1	0.017	0.3578	0.017	0.3544
YEAR99*IIUSP1	0.017	0.0109	0.017	0.0057
YEAR00*IIUSP1	0.009	0.1526	0.027	<.0001
YEAR01*IIUSP1	0.018	0.0014	0.001	0.8123
YEAR02*IIUSP1	0.003	0.5611	0.003	0.6022
YEAR03*IIUSP1	-0.027	<.0001	-0.025	<.0001
YEAR04*IIUSP1	-0.041	<.0001	-0.027	<.0001
SPECED*YEAR99	0.014	0.1646	0.009	0.3522
SPECED*YEAR00	0.006	0.556	-0.022	0.019
SPECED*YEAR01	-0.011	0.2656	-0.020	0.031
SPECED*YEAR02	-0.016	0.1562	-0.037	0.0007
SPECED*YEAR03	-0.092	<.0001	-0.292	<.0001
SPECED*YEAR04	0.021	0.0586	0.009	0.3865
SPECED*IIUSP1	-0.047	0.0005	-0.030	0.02
SPECED*YEAR99*IIUSP1	-0.090	<.0001	-0.071	0.0003
SPECED*YEAR00*IIUSP1	0.065	0.0011	0.050	0.0083
SPECED*YEAR01*IIUSP1	-0.044	0.0122	-0.039	0.0215
SPECED*YEAR02*IIUSP1	0.024	0.2081	0.034	0.0614
SPECED*YEAR03*IIUSP1	0.052	0.006	0.050	0.0053
SPECED*YEAR04*IIUSP1	-0.035	0.0559	-0.010	0.5765

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Special Education and Non-Special Education Students in Cohort 1 II/USP and Comparison Middle Schools



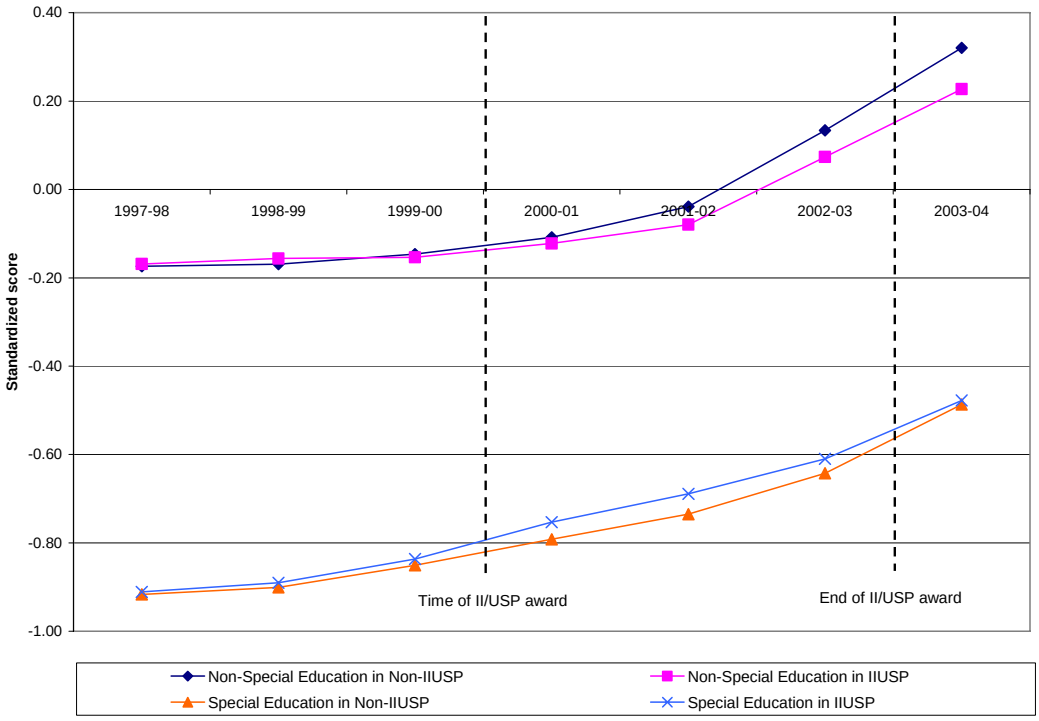
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Special Education and Non-Special Education Students in Cohort 1 II/USP and Comparison Middle Schools



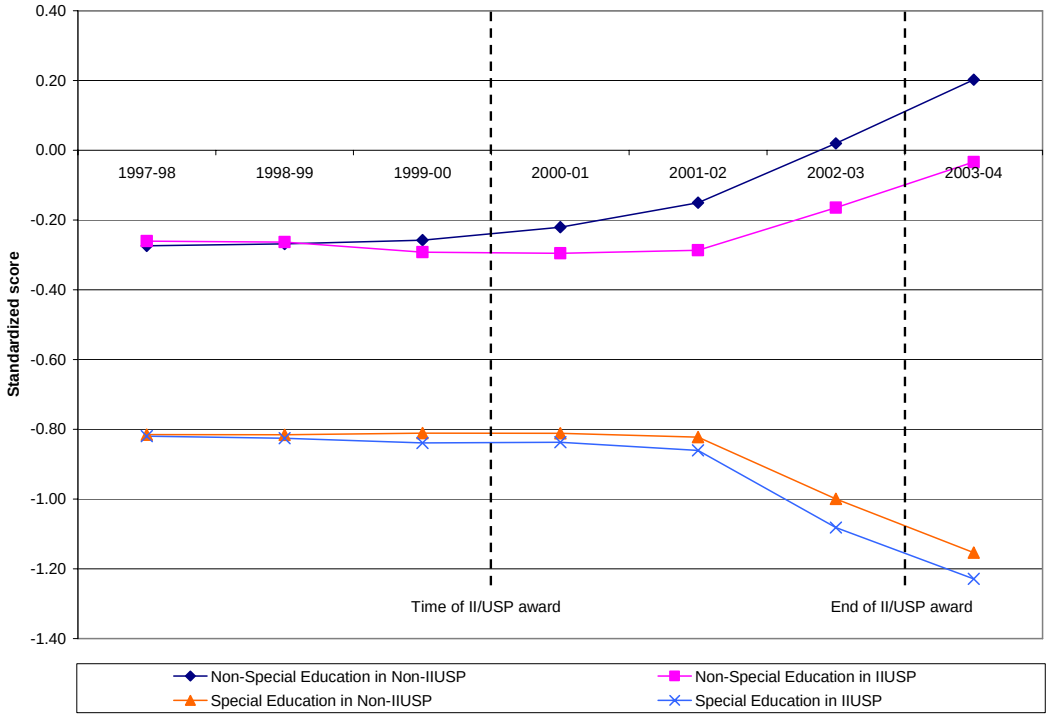
HLM Regression for Special Education and Non-Special Education Students in Cohort 2 Middle Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.200	<.0001	0.127	0.0002
GIRL	0.102	<.0001	-0.012	<.0001
ASIAN	-0.015	<.0001	0.270	<.0001
HISPANIC	-0.298	<.0001	-0.293	<.0001
BLACK	-0.489	<.0001	-0.491	<.0001
OTHERS	-0.151	<.0001	-0.076	<.0001
ELMISN	-0.165	<.0001	-0.092	<.0001
EL	-0.566	<.0001	-0.322	<.0001
R_FEP	0.218	<.0001	0.298	<.0001
FEP	0.182	<.0001	0.217	<.0001
FLUNCH	-0.139	<.0001	-0.104	<.0001
HIGHPARED	0.220	<.0001	0.189	<.0001
PARED_MISSING	-0.027	<.0001	-0.033	<.0001
SPECED	-0.743	<.0001	-0.541	<.0001
YEAR99	0.005	0.2093	0.005	0.1583
YEAR00	0.023	<.0001	0.011	0.0011
YEAR01	0.038	<.0001	0.037	<.0001
YEAR02	0.070	<.0001	0.070	<.0001
YEAR03	0.173	<.0001	0.170	<.0001
YEAR04	0.187	<.0001	0.182	<.0001
GRADE7	-0.007	<.0001	0.007	<.0001
GRADE8	-0.023	<.0001	-0.014	<.0001
PCT_MEALS	-0.302	<.0001	-0.293	<.0001
IIUSP2	0.005	0.7529	0.013	0.4506
YEAR99*IIUSP2	0.008	0.1561	-0.007	0.1707
YEAR00*IIUSP2	-0.020	0.0002	-0.040	<.0001
YEAR01*IIUSP2	-0.007	0.2132	-0.040	<.0001
YEAR02*IIUSP2	-0.027	<.0001	-0.062	<.0001
YEAR03*IIUSP2	-0.019	0.0002	-0.048	<.0001
YEAR04*IIUSP2	-0.033	<.0001	-0.052	<.0001
SPECED*YEAR99	0.011	0.367	-0.006	0.6229
SPECED*YEAR00	0.027	0.0129	-0.006	0.5625
SPECED*YEAR01	0.021	0.0585	-0.038	0.0003
SPECED*YEAR02	-0.013	0.2588	-0.081	<.0001
SPECED*YEAR03	-0.080	<.0001	-0.347	<.0001
SPECED*YEAR04	-0.031	0.0056	-0.337	<.0001
SPECED*IIUSP2	0.001	0.9632	-0.017	0.1346
SPECED*YEAR99*IIUSP2	-0.003	0.8599	0.002	0.9111
SPECED*YEAR00*IIUSP2	0.024	0.1435	0.021	0.1719
SPECED*YEAR01*IIUSP2	0.031	0.0554	0.043	0.0058
SPECED*YEAR02*IIUSP2	0.035	0.0362	0.049	0.002
SPECED*YEAR03*IIUSP2	0.006	0.7159	0.004	0.8195
SPECED*YEAR04*IIUSP2	0.009	0.564	0.059	0.0001

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Special Education and Non-Special Education Students in Cohort 2 II/USP and Comparison Middle Schools



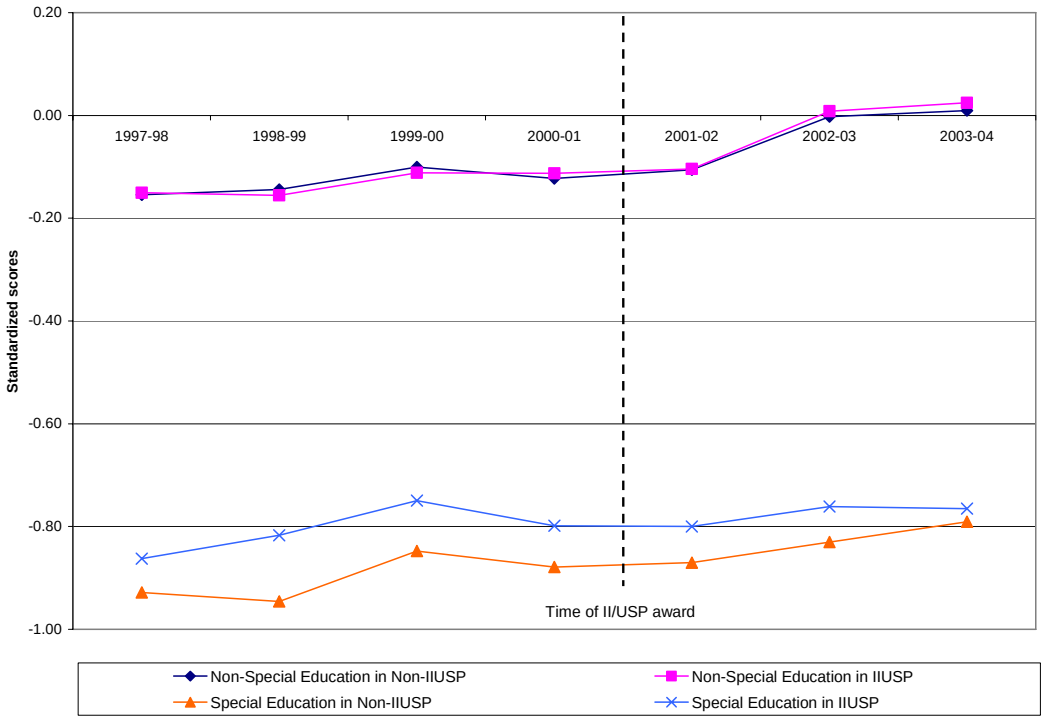
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Special Education and Non-Special Education Students in Cohort 2 II/USP and Comparison Middle Schools



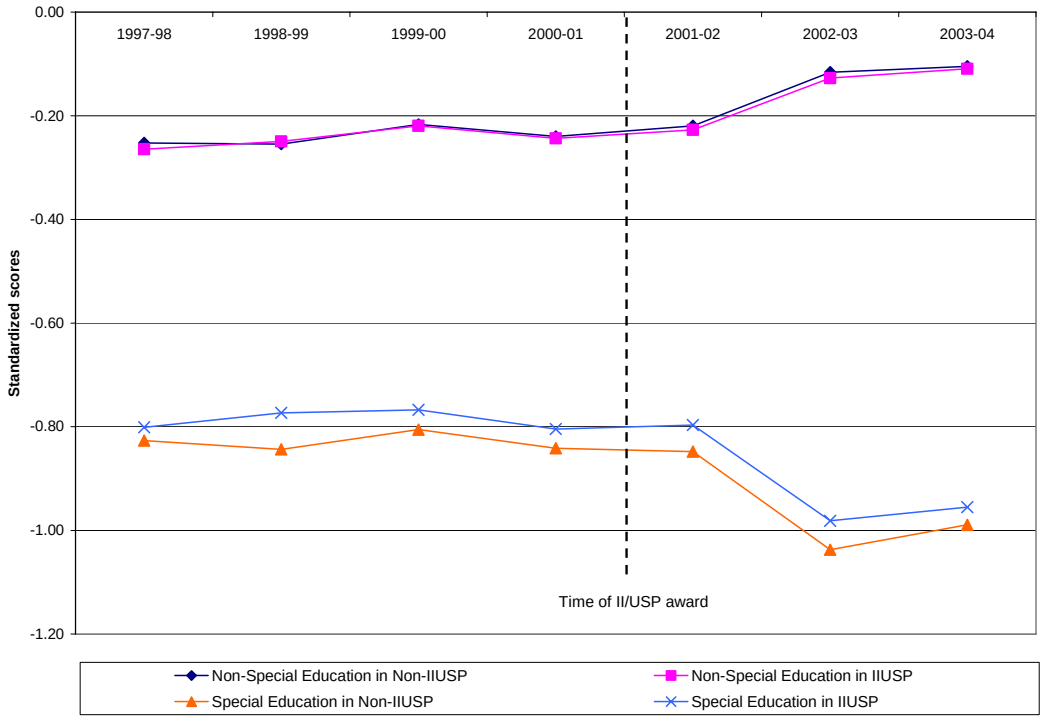
HLM Regression for Special Education and Non-Special Education Students in Cohort 3 Middle Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.248	<.0001	0.155	<.0001
GIRL	0.104	<.0001	-0.014	<.0001
ASIAN	-0.066	<.0001	0.230	<.0001
HISPANIC	-0.317	<.0001	-0.292	<.0001
BLACK	-0.482	<.0001	-0.485	<.0001
OTHERS	-0.165	<.0001	-0.071	<.0001
ELMISN	-0.159	<.0001	-0.128	<.0001
EL	-0.560	<.0001	-0.319	<.0001
R_FEP	0.225	<.0001	0.291	<.0001
FEP	0.166	<.0001	0.179	<.0001
FLUNCH	-0.153	<.0001	-0.119	<.0001
HIGHPARED	0.213	<.0001	0.183	<.0001
PARED_MISSING	-0.030	<.0001	-0.031	<.0001
SPECED	-0.774	<.0001	-0.574	<.0001
YEAR99	0.010	0.0016	-0.002	0.4584
YEAR00	0.043	<.0001	0.038	<.0001
YEAR01	-0.022	<.0001	-0.023	<.0001
YEAR02	0.017	<.0001	0.021	<.0001
YEAR03	0.103	<.0001	0.103	<.0001
YEAR04	0.012	<.0001	0.012	<.0001
GRADE7	-0.006	0.0003	0.007	<.0001
GRADE8	-0.023	<.0001	-0.016	<.0001
PCT_MEALS	-0.348	<.0001	-0.311	<.0001
IIUSP3	0.004	0.8158	-0.012	0.5195
YEAR99*IIUSP3	-0.016	0.022	0.017	0.009
YEAR00*IIUSP3	0.001	0.9264	-0.008	0.1936
YEAR01*IIUSP3	0.021	0.0002	-0.001	0.8471
YEAR02*IIUSP3	-0.008	0.1358	-0.004	0.4643
YEAR03*IIUSP3	0.009	0.0964	-0.003	0.5172
YEAR04*IIUSP3	0.004	0.4712	0.006	0.2305
SPECED*YEAR99	-0.028	0.0092	-0.014	0.158
SPECED*YEAR00	0.055	<.0001	0.000	0.9874
SPECED*YEAR01	-0.009	0.3206	-0.013	0.1382
SPECED*YEAR02	-0.009	0.3438	-0.027	0.0016
SPECED*YEAR03	-0.063	<.0001	-0.292	<.0001
SPECED*YEAR04	0.028	0.0008	0.036	<.0001
SPECED*IIUSP3	0.062	<.0001	0.038	0.0084
SPECED*YEAR99*IIUSP3	0.078	0.0004	0.027	0.1981
SPECED*YEAR00*IIUSP3	-0.031	0.1402	-0.025	0.2287
SPECED*YEAR01*IIUSP3	-0.039	0.0356	0.000	0.9979
SPECED*YEAR02*IIUSP3	-0.002	0.9324	0.019	0.2911
SPECED*YEAR03*IIUSP3	-0.010	0.58	0.008	0.6532
SPECED*YEAR04*IIUSP3	-0.048	0.0063	-0.028	0.0969

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Special Education and Non-Special Education Students in Cohort 3 II/USP and Comparison Middle Schools



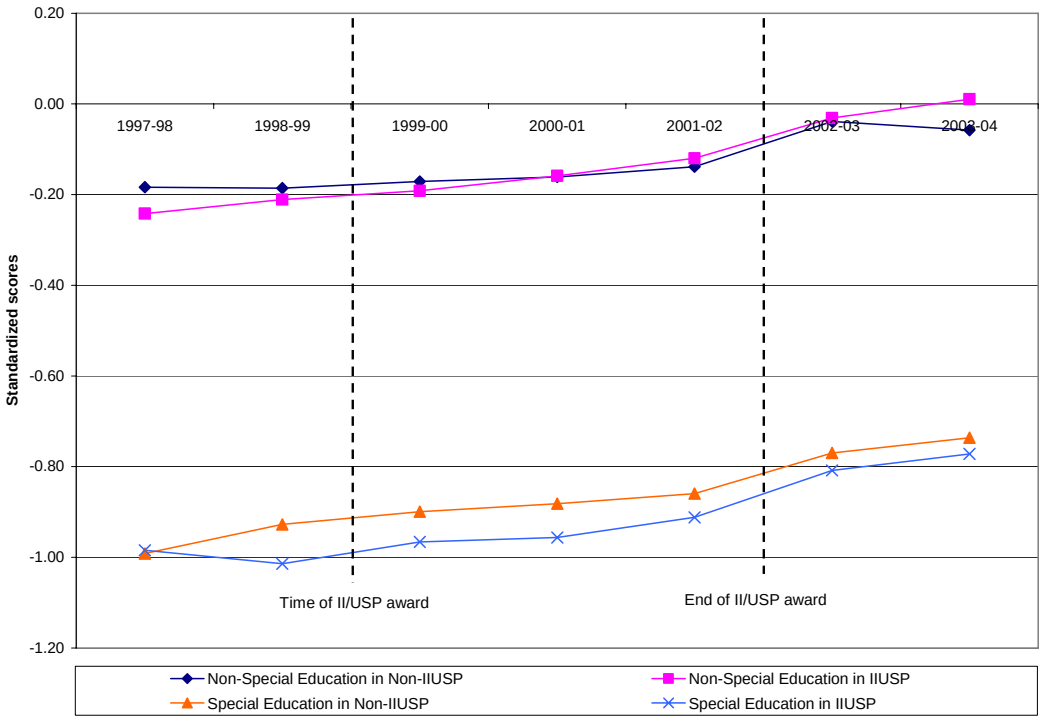
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Special Education and Non-Special Education Students in Cohort 3 II/USP and Comparison Middle Schools



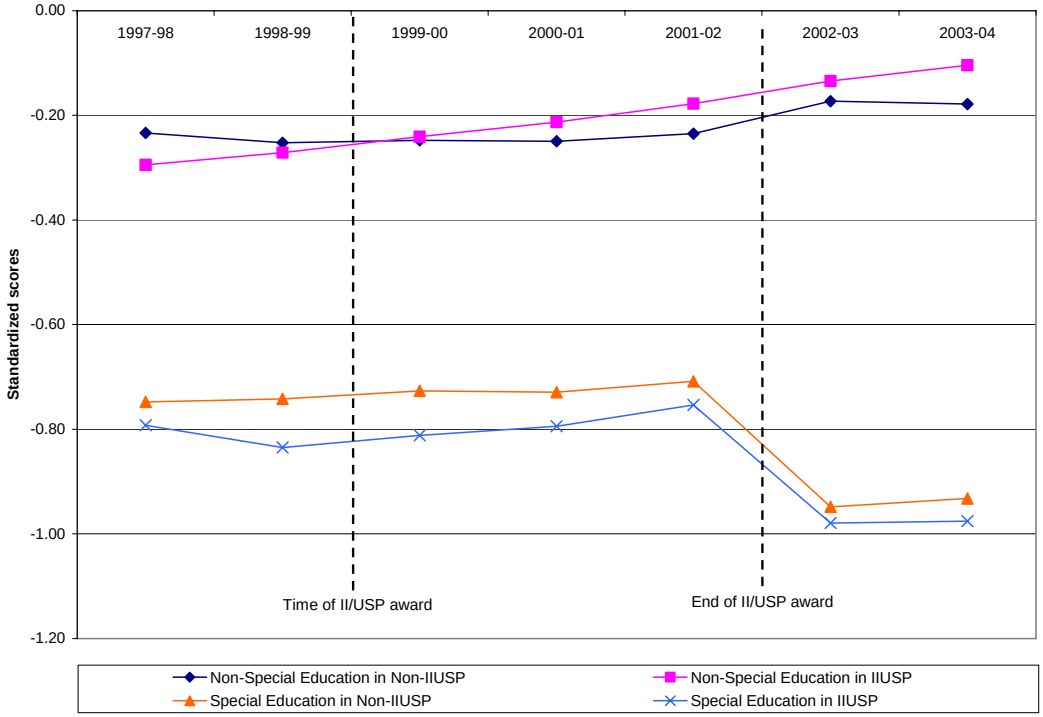
HLM Regression for Special Education and Non-Special Education Students in Cohort 1 High Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.086	0.0148	0.110	0.0005
GIRL	0.110	<.0001	-0.066	<.0001
ASIAN	-0.121	<.0001	0.256	<.0001
HISPANIC	-0.328	<.0001	-0.326	<.0001
BLACK	-0.536	<.0001	-0.530	<.0001
OTHERS	-0.213	<.0001	-0.120	<.0001
ELMISN	-0.028	0.0013	0.032	0.0001
EL	-0.616	<.0001	-0.335	<.0001
R_FEP	0.069	<.0001	0.111	<.0001
FEP	0.099	<.0001	0.106	<.0001
FLUNCH	-0.078	<.0001	-0.040	<.0001
HIGHPARED	0.245	<.0001	0.194	<.0001
PARED_MISSING	-0.017	<.0001	-0.026	<.0001
SPECED	-0.808	<.0001	-0.514	<.0001
YEAR99	-0.002	0.6978	-0.019	<.0001
YEAR00	0.015	0.0004	0.005	0.2604
YEAR01	0.010	0.0132	-0.002	0.6305
YEAR02	0.023	<.0001	0.014	0.0001
YEAR03	0.100	<.0001	0.062	<.0001
YEAR04	-0.020	<.0001	-0.005	0.1516
GRADE10	0.004	0.0223	0.016	<.0001
GRADE11	0.017	<.0001	0.003	0.1568
PCT_MEALS	-0.119	0.0546	-0.176	0.0015
IIUSP1	-0.058	0.0212	-0.061	0.0071
YEAR99*IIUSP1	0.033	<.0001	0.042	<.0001
YEAR00*IIUSP1	0.005	0.4896	0.026	<.0001
YEAR01*IIUSP1	0.023	0.0003	0.030	<.0001
YEAR02*IIUSP1	0.015	0.015	0.021	0.0008
YEAR03*IIUSP1	-0.011	0.0941	-0.019	0.0023
YEAR04*IIUSP1	0.061	<.0001	0.035	<.0001
SPECED*YEAR99	0.066	0.0001	0.024	0.1402
SPECED*YEAR00	0.013	0.3812	0.011	0.4644
SPECED*YEAR01	0.008	0.5628	-0.001	0.962
SPECED*YEAR02	-0.001	0.9685	0.006	0.6382
SPECED*YEAR03	-0.010	0.4375	-0.302	<.0001
SPECED*YEAR04	0.053	<.0001	0.021	0.0875
SPECED*IIUSP1	0.065	0.0013	0.016	0.3964
SPECED*YEAR99*IIUSP1	-0.127	<.0001	-0.091	0.0009
SPECED*YEAR00*IIUSP1	0.016	0.5385	-0.018	0.4709
SPECED*YEAR01*IIUSP1	-0.031	0.1672	-0.010	0.6463
SPECED*YEAR02*IIUSP1	0.006	0.7752	0.000	0.9852
SPECED*YEAR03*IIUSP1	0.025	0.2729	0.032	0.138
SPECED*YEAR04*IIUSP1	-0.057	0.0062	-0.047	0.0219

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Special Education and Non-Special Education Students in Cohort 1 II/USP and Comparison High Schools



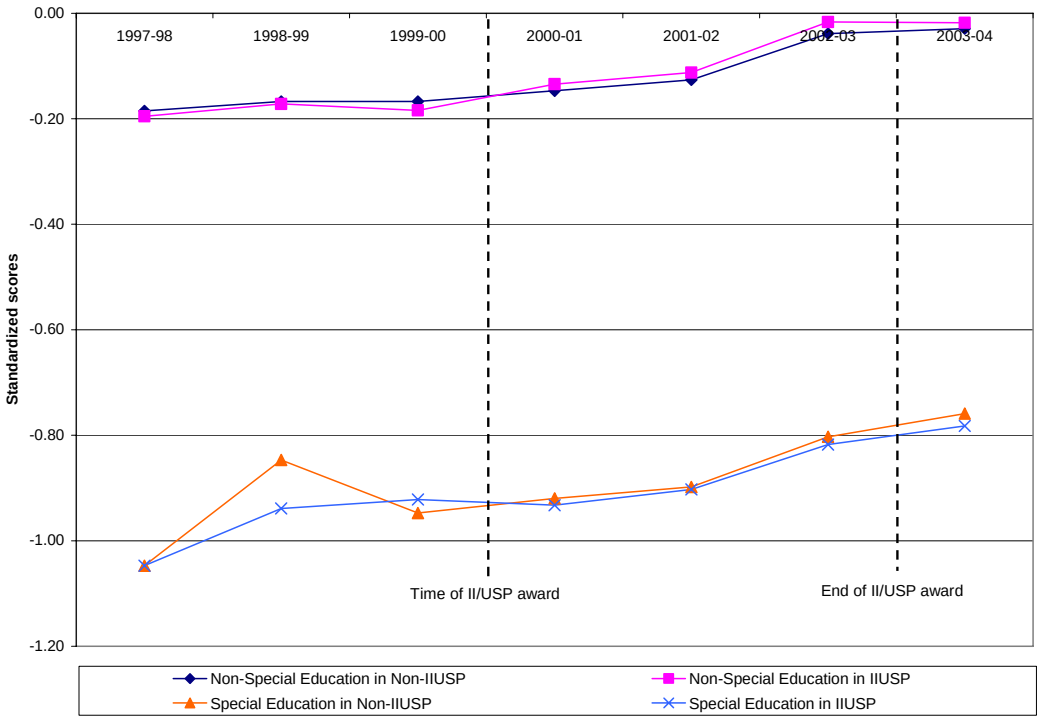
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Special Education and Non-Special Education Students in Cohort 1 II/USP and Comparison High Schools



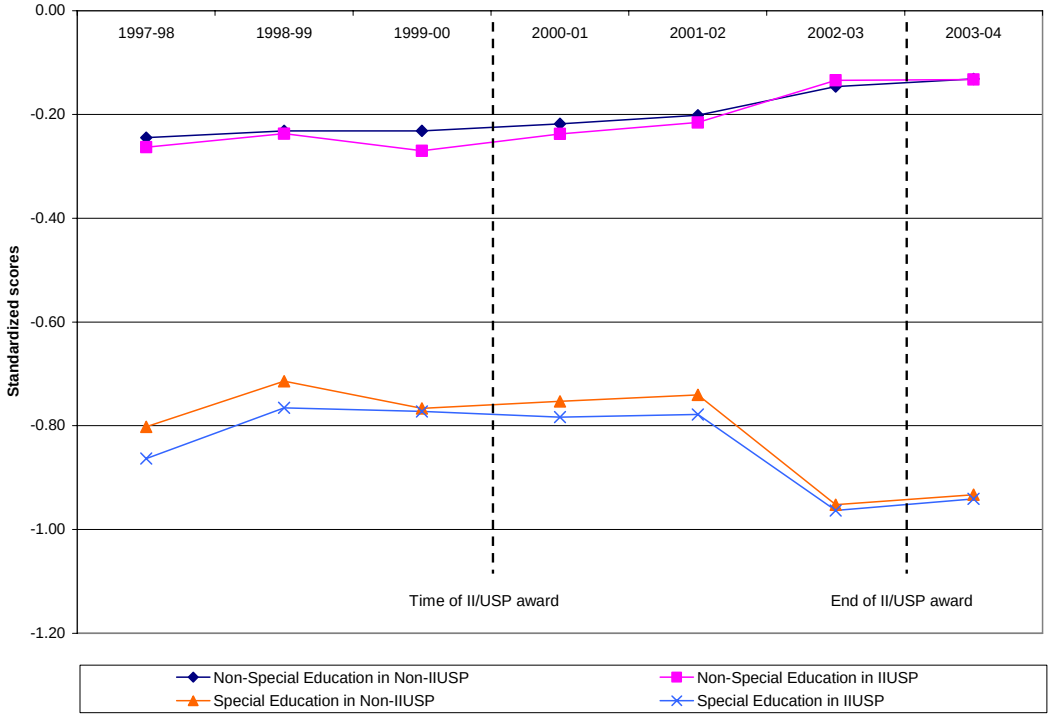
HLM Regression for Special Education and Non-Special Education Students in Cohort 2 High Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.067	0.0049	0.086	0.0002
GIRL	0.116	<.0001	-0.067	<.0001
ASIAN	-0.083	<.0001	0.253	<.0001
HISPANIC	-0.312	<.0001	-0.323	<.0001
BLACK	-0.515	<.0001	-0.521	<.0001
OTHERS	-0.183	<.0001	-0.109	<.0001
ELMISN	0.004	0.6332	0.018	0.0171
EL	-0.626	<.0001	-0.348	<.0001
R_FEP	0.077	<.0001	0.136	<.0001
FEP	0.129	<.0001	0.147	<.0001
FLUNCH	-0.095	<.0001	-0.059	<.0001
HIGHPARED	0.246	<.0001	0.204	<.0001
PARED_MISSING	-0.008	<.0001	-0.020	<.0001
SPECED	-0.862	<.0001	-0.557	<.0001
YEAR99	0.018	<.0001	0.013	<.0001
YEAR00	0.000	0.9788	0.000	0.9749
YEAR01	0.020	<.0001	0.014	<.0001
YEAR02	0.021	<.0001	0.016	<.0001
YEAR03	0.088	<.0001	0.055	<.0001
YEAR04	0.009	0.0007	0.015	<.0001
GRADE10	-0.007	<.0001	0.006	<.0001
GRADE11	-0.001	0.6066	-0.010	<.0001
PCT_MEALS	-0.132	0.0053	-0.144	0.0018
IIUSP2	-0.010	0.5829	-0.018	0.2975
YEAR99*IIUSP2	0.005	0.3552	0.013	0.0235
YEAR00*IIUSP2	-0.012	0.0147	-0.033	<.0001
YEAR01*IIUSP2	0.030	<.0001	0.019	<.0001
YEAR02*IIUSP2	0.001	0.8418	0.006	0.2195
YEAR03*IIUSP2	0.008	0.0711	0.026	<.0001
YEAR04*IIUSP2	-0.011	0.0173	-0.013	0.0037
SPECED*YEAR99	0.183	<.0001	0.075	<.0001
SPECED*YEAR00	-0.101	<.0001	-0.053	<.0001
SPECED*YEAR01	0.008	0.4267	0.000	0.985
SPECED*YEAR02	0.000	0.9873	-0.004	0.6945
SPECED*YEAR03	0.008	0.4293	-0.267	<.0001
SPECED*YEAR04	0.035	0.0001	0.005	0.6162
SPECED*IIUSP2	0.010	0.4958	-0.043	0.0029
SPECED*YEAR99*IIUSP2	-0.098	<.0001	-0.003	0.8807
SPECED*YEAR00*IIUSP2	0.130	<.0001	0.079	<.0001
SPECED*YEAR01*IIUSP2	-0.068	<.0001	-0.044	0.0064
SPECED*YEAR02*IIUSP2	0.008	0.6346	-0.013	0.4075
SPECED*YEAR03*IIUSP2	-0.018	0.2488	0.001	0.9657
SPECED*YEAR04*IIUSP2	0.002	0.9044	0.016	0.2833

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Special Education and Non-Special Education Students in Cohort 2 II/USP and Comparison High Schools



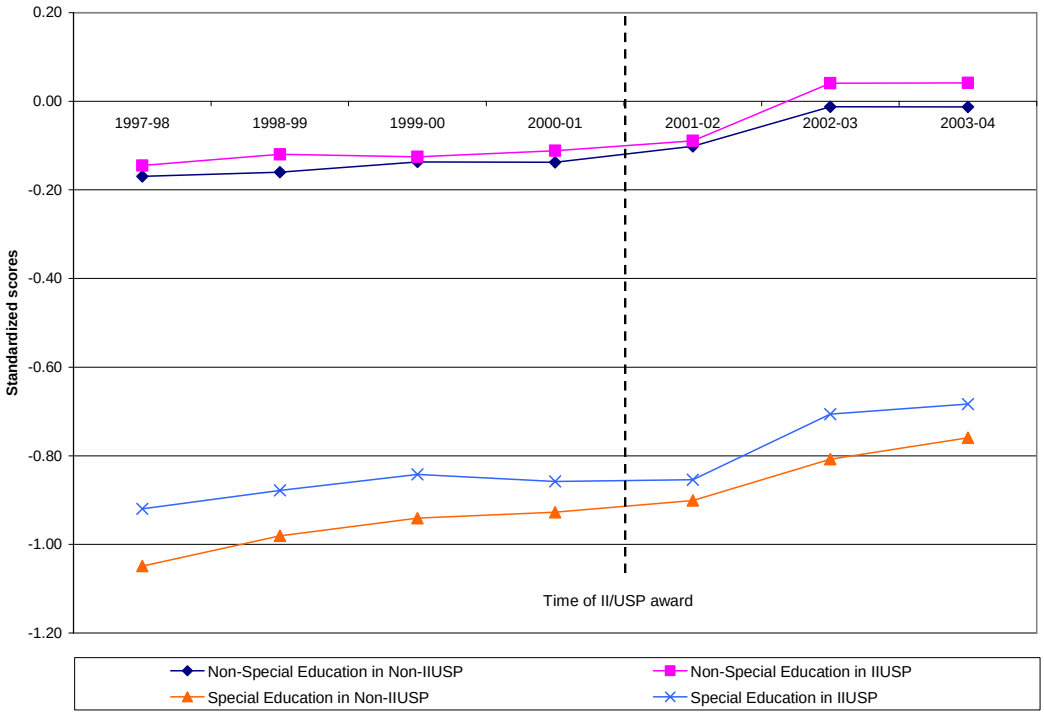
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Special Education and Non-Special Education Students in Cohort 2 II/USP and Comparison High Schools



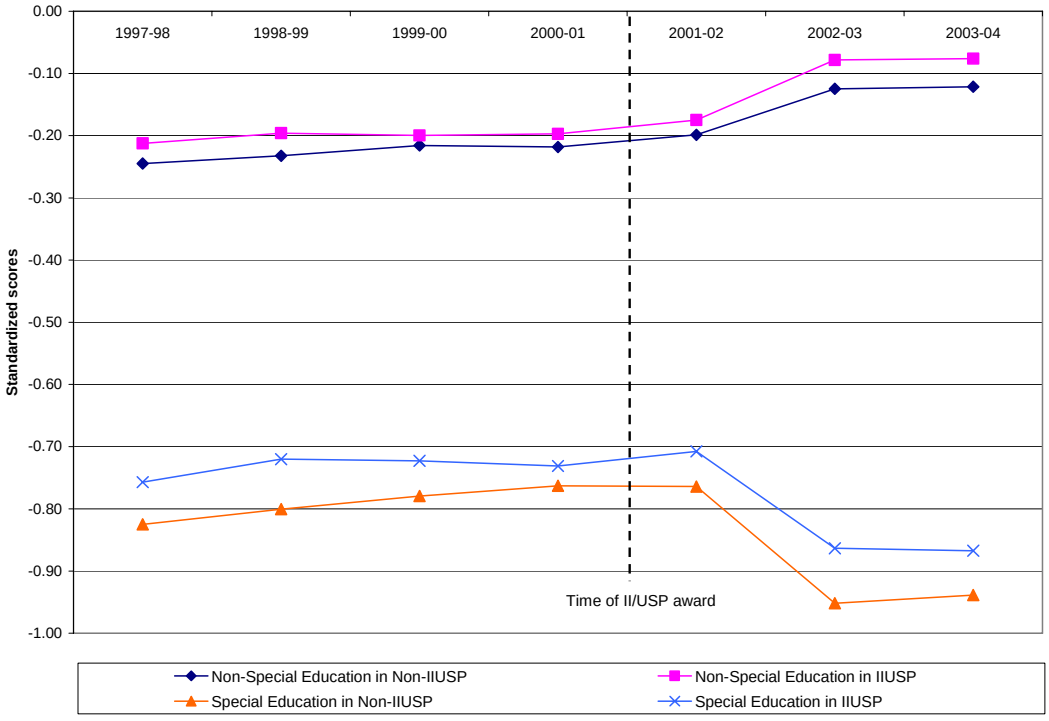
HLM Regression for Special Education and Non-Special Education Students in Cohort 3 High Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.040	0.064	0.030	0.1944
GIRL	0.114	<.0001	-0.073	<.0001
ASIAN	-0.097	<.0001	0.250	<.0001
HISPANIC	-0.319	<.0001	-0.324	<.0001
BLACK	-0.520	<.0001	-0.523	<.0001
OTHERS	-0.195	<.0001	-0.112	<.0001
ELMISN	-0.109	<.0001	-0.038	<.0001
EL	-0.636	<.0001	-0.358	<.0001
R_FEP	0.076	<.0001	0.131	<.0001
FEP	0.102	<.0001	0.122	<.0001
FLUNCH	-0.100	<.0001	-0.061	<.0001
HIGHPARED	0.245	<.0001	0.205	<.0001
PARED_MISSING	-0.006	<.0001	-0.019	<.0001
SPECED	-0.880	<.0001	-0.580	<.0001
YEAR99	0.009	0.0018	0.013	<.0001
YEAR00	0.023	<.0001	0.016	<.0001
YEAR01	-0.001	0.8343	-0.002	0.3016
YEAR02	0.036	<.0001	0.020	<.0001
YEAR03	0.090	<.0001	0.074	<.0001
YEAR04	0.000	0.8843	0.003	0.1395
GRADE10	-0.010	<.0001	0.001	0.331
GRADE11	-0.005	<.0001	-0.017	<.0001
PCT_MEALS	-0.032	0.472	-0.029	0.5465
IIUSP3	0.024	0.1839	0.033	0.0927
YEAR99*IIUSP3	0.016	0.0021	0.004	0.4492
YEAR00*IIUSP3	-0.029	<.0001	-0.020	<.0001
YEAR01*IIUSP3	0.014	0.0009	0.005	0.2538
YEAR02*IIUSP3	-0.013	0.0014	0.003	0.5019
YEAR03*IIUSP3	0.040	<.0001	0.023	<.0001
YEAR04*IIUSP3	0.001	0.7784	-0.001	0.7195
SPECED*YEAR99	0.059	<.0001	0.012	0.2495
SPECED*YEAR00	0.017	0.0802	0.005	0.6166
SPECED*YEAR01	0.014	0.1023	0.019	0.0212
SPECED*YEAR02	-0.009	0.2692	-0.021	0.0108
SPECED*YEAR03	0.004	0.6405	-0.262	<.0001
SPECED*YEAR04	0.049	<.0001	0.010	0.2131
SPECED*IIUSP3	0.105	<.0001	0.035	0.006
SPECED*YEAR99*IIUSP3	-0.043	0.0213	0.008	0.6398
SPECED*YEAR00*IIUSP3	0.025	0.1421	-0.004	0.8235
SPECED*YEAR01*IIUSP3	-0.043	0.0034	-0.029	0.0411
SPECED*YEAR02*IIUSP3	-0.009	0.5402	0.022	0.1201
SPECED*YEAR03*IIUSP3	0.014	0.3275	0.009	0.5173
SPECED*YEAR04*IIUSP3	-0.027	0.0523	-0.016	0.2436

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Special Education and Non-Special Education Students in Cohort 3 II/USP and Comparison High Schools



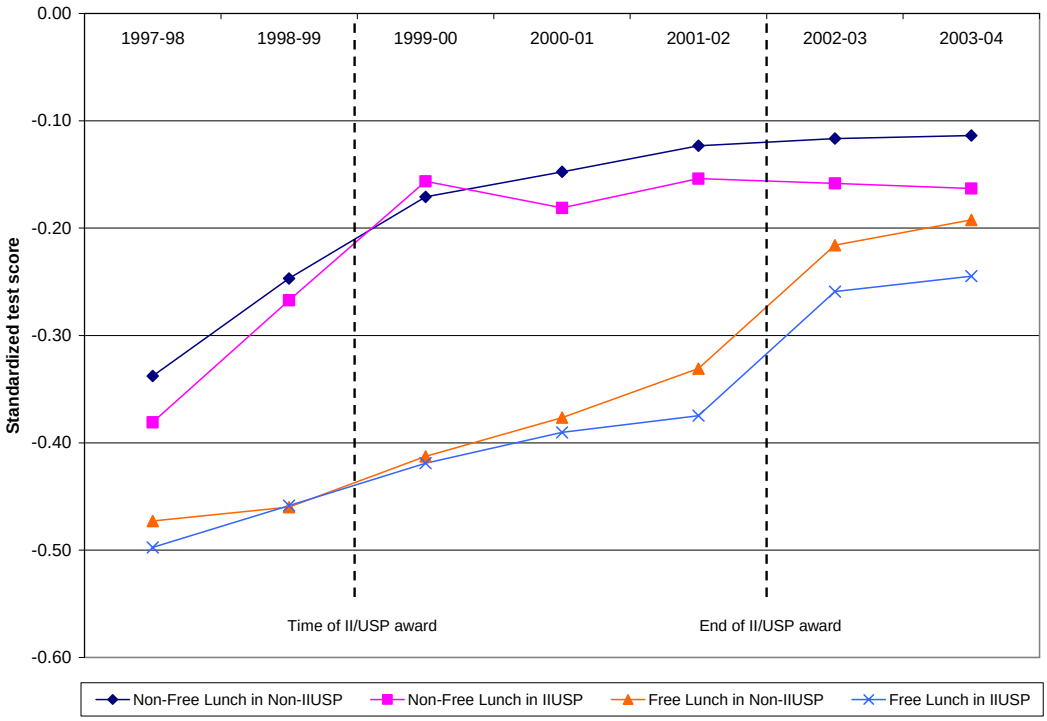
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Special Education and Non-Special Education Students in Cohort 3 II/USP and Comparison High Schools



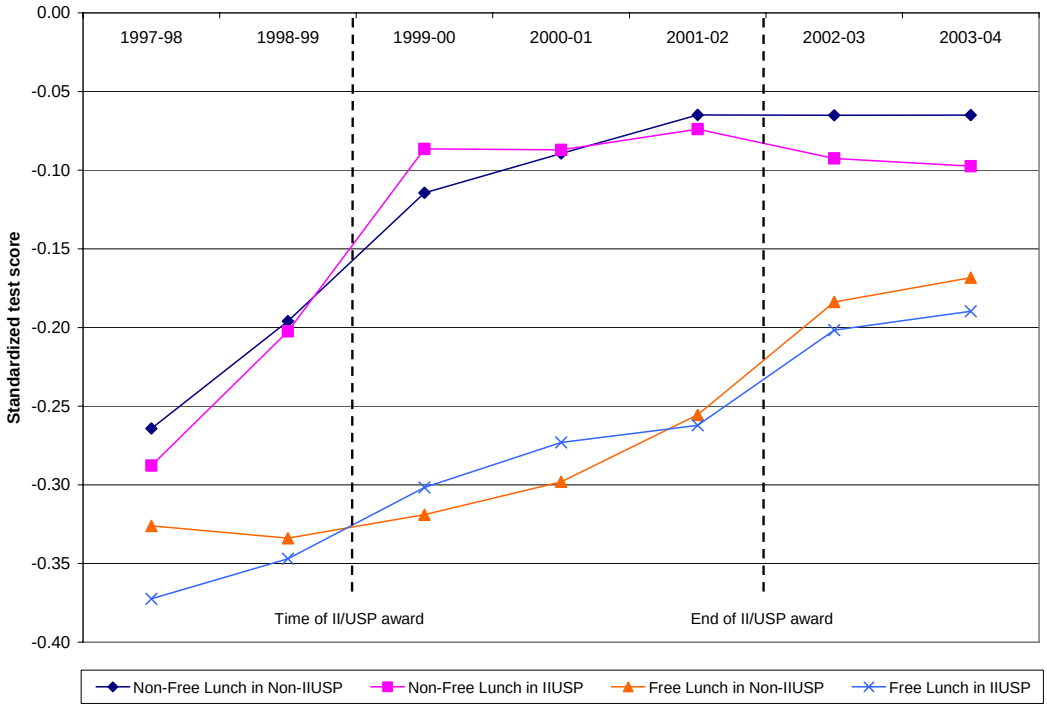
HLM Regression for Students Eligible for Free and Reduced Price Lunch and "Regular" Students in Cohort 1 Elementary Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.092	<.0001	0.017	0.5084
GIRL	0.108	<.0001	-0.015	<.0001
ASIAN	-0.005	0.0803	0.197	<.0001
HISPANIC	-0.257	<.0001	-0.233	<.0001
BLACK	-0.404	<.0001	-0.451	<.0001
OTHERS	-0.090	<.0001	-0.051	<.0001
ELMISN	-0.120	<.0001	-0.077	<.0001
EL	-0.374	<.0001	-0.173	<.0001
R_FEP	0.346	<.0001	0.487	<.0001
FEP	0.185	<.0001	0.232	<.0001
FLUNCH	-0.135	<.0001	-0.062	<.0001
HIGHPARED	0.222	<.0001	0.202	<.0001
PARED_MISSING	-0.012	<.0001	-0.034	<.0001
SPECED	-0.592	<.0001	-0.657	<.0001
YEAR99	0.091	<.0001	0.068	<.0001
YEAR00	0.076	<.0001	0.081	<.0001
YEAR01	0.023	<.0001	0.025	<.0001
YEAR02	0.024	<.0001	0.025	<.0001
YEAR03	0.007	0.302	0.000	0.9751
YEAR04	0.003	0.6645	0.000	0.9838
GRADE3	-0.031	<.0001	-0.024	<.0001
GRADE4	-0.063	<.0001	-0.061	<.0001
GRADE5	-0.090	<.0001	-0.102	<.0001
PCT_MEALS	-0.191	<.0001	-0.101	0.0016
IIUSP1	-0.043	0.0001	-0.023	0.0614
YEAR99*IIUSP1	0.023	0.002	0.017	0.0246
YEAR00*IIUSP1	0.035	<.0001	0.034	0.0001
YEAR01*IIUSP1	-0.048	<.0001	-0.026	0.0065
YEAR02*IIUSP1	0.003	0.7655	-0.011	0.2562
YEAR03*IIUSP1	-0.011	0.274	-0.018	0.0779
YEAR04*IIUSP1	-0.008	0.4406	-0.005	0.6204
FLUNCH*YEAR99	-0.078	<.0001	-0.076	<.0001
FLUNCH*YEAR00	-0.029	<.0001	-0.067	<.0001
FLUNCH*YEAR01	0.013	0.0286	-0.004	0.4956
FLUNCH*YEAR02	0.021	0.0012	0.018	0.0084
FLUNCH*YEAR03	0.108	<.0001	0.072	<.0001
FLUNCH*YEAR04	0.021	0.0027	0.015	0.0367
FLUNCH*IIUSP1	0.019	0.0089	-0.023	0.0015
FLUNCH*YEAR99*IIUSP1	0.003	0.7538	0.016	0.1035
FLUNCH*YEAR00*IIUSP1	-0.043	<.0001	-0.004	0.689
FLUNCH*YEAR01*IIUSP1	0.041	<.0001	0.033	0.0014
FLUNCH*YEAR02*IIUSP1	-0.033	0.0018	-0.020	0.065
FLUNCH*YEAR03*IIUSP1	0.012	0.2842	0.007	0.5367
FLUNCH*YEAR04*IIUSP1	-0.002	0.8786	0.002	0.8753

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Students Eligible for Free and Reduced Price Lunch and "Regular" Students in Cohort 1 II/USP and Comparison Elementary Schools,



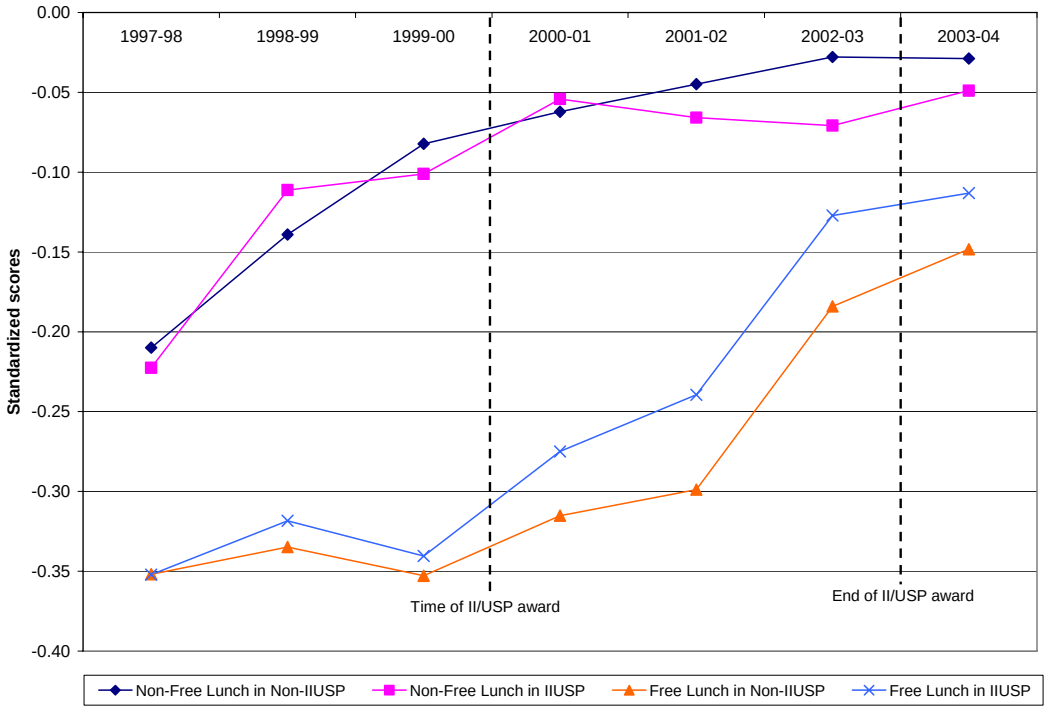
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Students Eligible for Free and Reduced Price Lunch and "Regular" Students in Cohort 1 II/USP and Comparison Elementary Schools



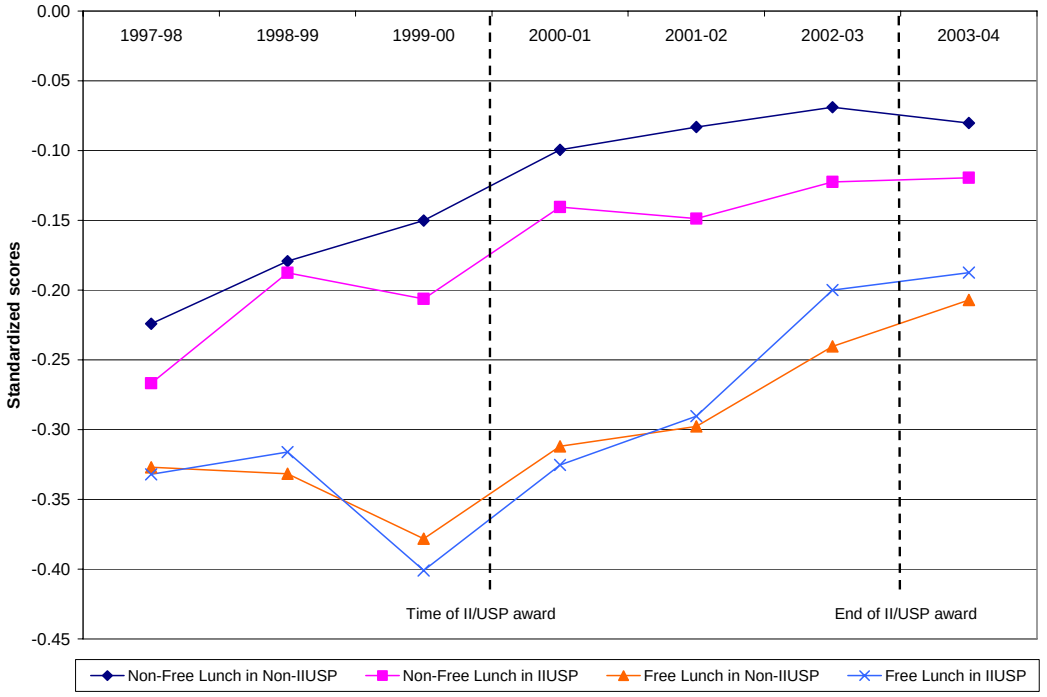
HLM Regression for Students Eligible for Free and Reduced Price Lunch and "Regular" Students in Cohort 2 Elementary Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.145	<.0001	0.099	0.0024
GIRL	0.114	<.0001	-0.011	<.0001
ASIAN	-0.028	<.0001	0.166	<.0001
HISPANIC	-0.265	<.0001	-0.237	<.0001
BLACK	-0.415	<.0001	-0.463	<.0001
OTHERS	-0.122	<.0001	-0.075	<.0001
ELMISN	-0.165	<.0001	-0.104	<.0001
EL	-0.370	<.0001	-0.179	<.0001
R_FEP	0.365	<.0001	0.501	<.0001
FEP	0.198	<.0001	0.242	<.0001
FLUNCH	-0.142	<.0001	-0.103	<.0001
HIGHPARED	0.236	<.0001	0.216	<.0001
PARED_MISSING	-0.015	<.0001	-0.034	<.0001
SPECED	-0.614	<.0001	-0.667	<.0001
YEAR99	0.071	<.0001	0.045	<.0001
YEAR00	0.057	<.0001	0.029	0.0003
YEAR01	0.020	0.0139	0.051	<.0001
YEAR02	0.017	0.045	0.016	0.0654
YEAR03	0.017	0.0557	0.014	0.1199
YEAR04	-0.001	0.9131	-0.011	0.2265
GRADE3	-0.028	<.0001	-0.039	<.0001
GRADE4	-0.057	<.0001	-0.071	<.0001
GRADE5	-0.075	<.0001	-0.095	<.0001
PCT_MEALS	-0.273	<.0001	-0.167	0.0001
IIUSP2	-0.013	0.3623	-0.043	0.0071
YEAR99*IIUSP2	0.040	<.0001	0.034	0.0007
YEAR00*IIUSP2	-0.047	<.0001	-0.048	<.0001
YEAR01*IIUSP2	0.027	0.023	0.015	0.2087
YEAR02*IIUSP2	-0.029	0.0183	-0.025	0.0504
YEAR03*IIUSP2	-0.022	0.0759	0.012	0.3513
YEAR04*IIUSP2	0.023	0.0692	0.014	0.2727
FLUNCH*YEAR99	-0.054	<.0001	-0.050	<.0001
FLUNCH*YEAR00	-0.075	<.0001	-0.076	<.0001
FLUNCH*YEAR01	0.018	0.0591	0.016	0.1015
FLUNCH*YEAR02	-0.001	0.9156	-0.002	0.8387
FLUNCH*YEAR03	0.098	<.0001	0.043	<.0001
FLUNCH*YEAR04	0.037	0.0003	0.045	<.0001
FLUNCH*IIUSP2	0.012	0.1709	0.037	<.0001
FLUNCH*YEAR99*IIUSP2	-0.024	0.066	-0.014	0.2953
FLUNCH*YEAR00*IIUSP2	0.043	0.0013	0.010	0.4805
FLUNCH*YEAR01*IIUSP2	0.001	0.9366	-0.006	0.668
FLUNCH*YEAR02*IIUSP2	0.048	0.0004	0.045	0.0011
FLUNCH*YEAR03*IIUSP2	0.019	0.1581	0.021	0.1383
FLUNCH*YEAR04*IIUSP2	-0.045	0.0014	-0.035	0.0156

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Students Eligible for Free and Reduced Price Lunch and "Regular" Students in Cohort 2 II/USP and Comparison Elementary Schools



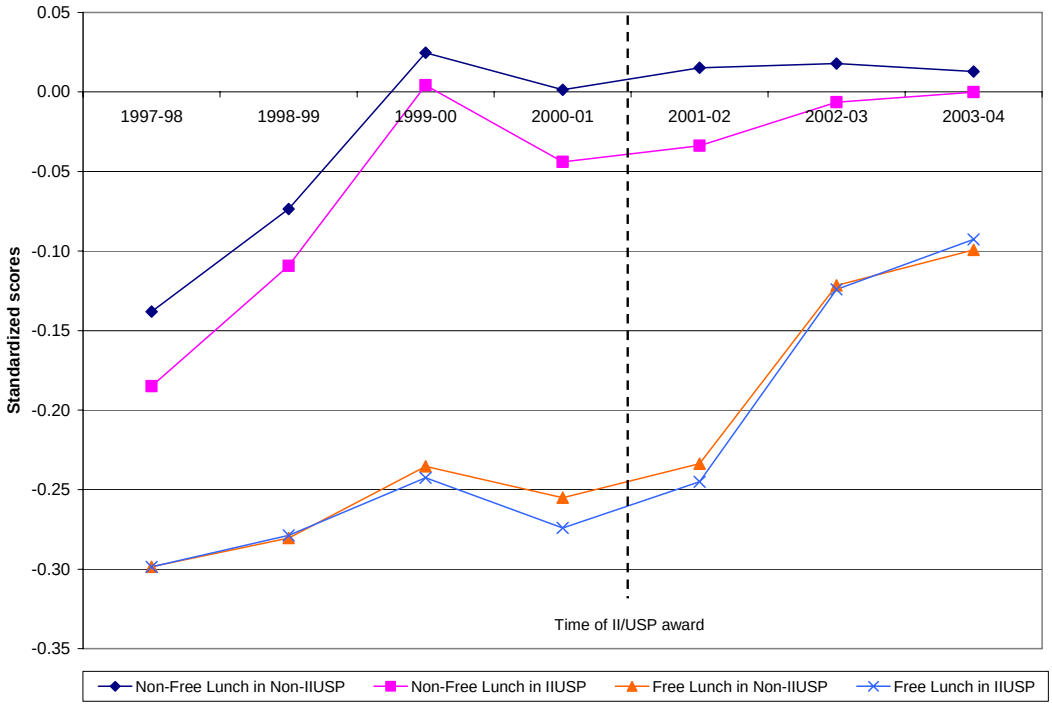
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Students Eligible for Free and Reduced Price Lunch and "Regular" Students in Cohort 2 II/USP and Comparison Elementary Schools



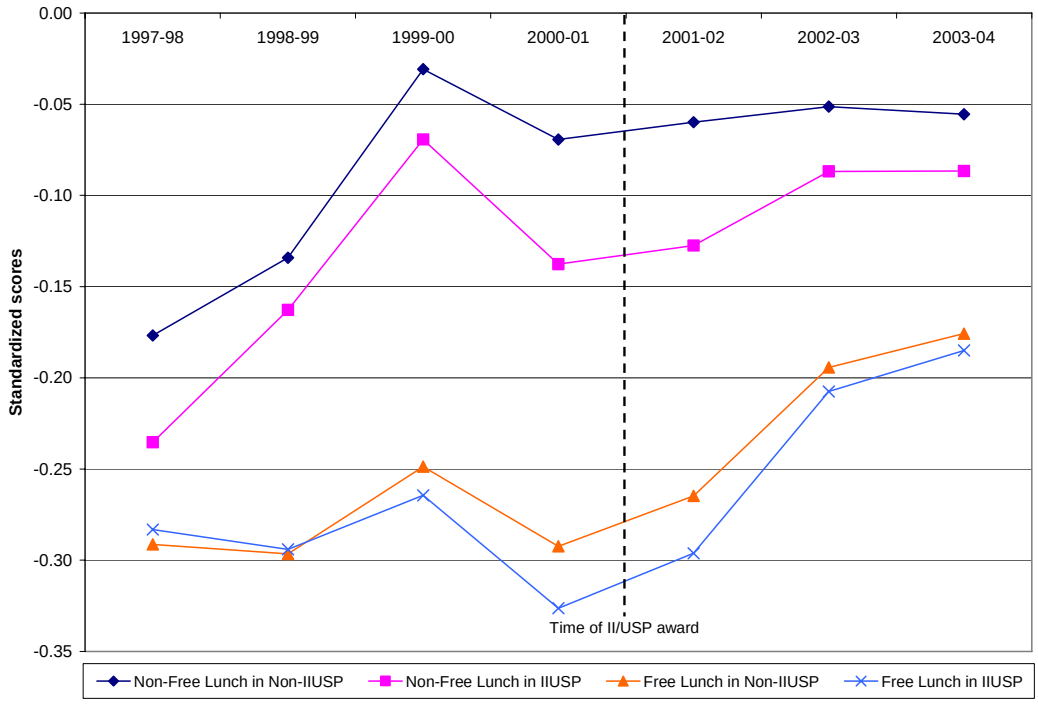
HLM Regression for Students Eligible for Free and Reduced Price Lunch and "Regular" Students in Cohort 3 Elementary Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.161	<.0001	0.089	<.0001
GIRL	0.118	<.0001	-0.014	<.0001
ASIAN	-0.007	0.0353	0.185	<.0001
HISPANIC	-0.252	<.0001	-0.228	<.0001
BLACK	-0.408	<.0001	-0.447	<.0001
OTHERS	-0.099	<.0001	-0.054	<.0001
ELMISN	-0.135	<.0001	-0.096	<.0001
EL	-0.390	<.0001	-0.195	<.0001
R_FEP	0.318	<.0001	0.451	<.0001
FEP	0.164	<.0001	0.217	<.0001
FLUNCH	-0.160	<.0001	-0.115	<.0001
HIGHPARED	0.238	<.0001	0.219	<.0001
PARED_MISSING	-0.003	0.0867	-0.032	<.0001
SPECED	-0.639	<.0001	-0.680	<.0001
YEAR99	0.065	<.0001	0.043	<.0001
YEAR00	0.098	<.0001	0.103	<.0001
YEAR01	-0.023	<.0001	-0.039	<.0001
YEAR02	0.014	0.008	0.009	0.075
YEAR03	0.003	0.6071	0.009	0.1114
YEAR04	-0.005	0.3422	-0.004	0.4371
GRADE3	-0.018	<.0001	-0.024	<.0001
GRADE4	-0.044	<.0001	-0.052	<.0001
GRADE5	-0.061	<.0001	-0.079	<.0001
PCT_MEALS	-0.256	<.0001	-0.130	<.0001
IIUSP3	-0.047	<.0001	-0.059	<.0001
YEAR99*IIUSP3	0.011	0.1496	0.030	0.0001
YEAR00*IIUSP3	0.015	0.0811	-0.010	0.2715
YEAR01*IIUSP3	-0.025	0.006	-0.030	0.0013
YEAR02*IIUSP3	-0.004	0.6962	0.001	0.9303
YEAR03*IIUSP3	0.025	0.0082	0.032	0.0008
YEAR04*IIUSP3	0.011	0.2277	0.004	0.6458
FLUNCH*YEAR99	-0.047	<.0001	-0.048	<.0001
FLUNCH*YEAR00	-0.053	<.0001	-0.056	<.0001
FLUNCH*YEAR01	0.004	0.555	-0.005	0.3917
FLUNCH*YEAR02	0.008	0.2071	0.018	0.0032
FLUNCH*YEAR03	0.109	<.0001	0.062	<.0001
FLUNCH*YEAR04	0.027	<.0001	0.023	0.0003
FLUNCH*IIUSP3	0.047	<.0001	0.067	<.0001
FLUNCH*YEAR99*IIUSP3	-0.009	0.3889	-0.036	0.0009
FLUNCH*YEAR00*IIUSP3	-0.024	0.0217	-0.008	0.4351
FLUNCH*YEAR01*IIUSP3	0.013	0.2058	0.012	0.2647
FLUNCH*YEAR02*IIUSP3	0.011	0.2807	0.002	0.8759
FLUNCH*YEAR03*IIUSP3	-0.016	0.1323	-0.014	0.1975
FLUNCH*YEAR04*IIUSP3	-0.002	0.8266	0.000	0.9651

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Students Eligible for Free and Reduced Price Lunch and "Regular" Students in Cohort 3 II/USP and Comparison Elementary Schools



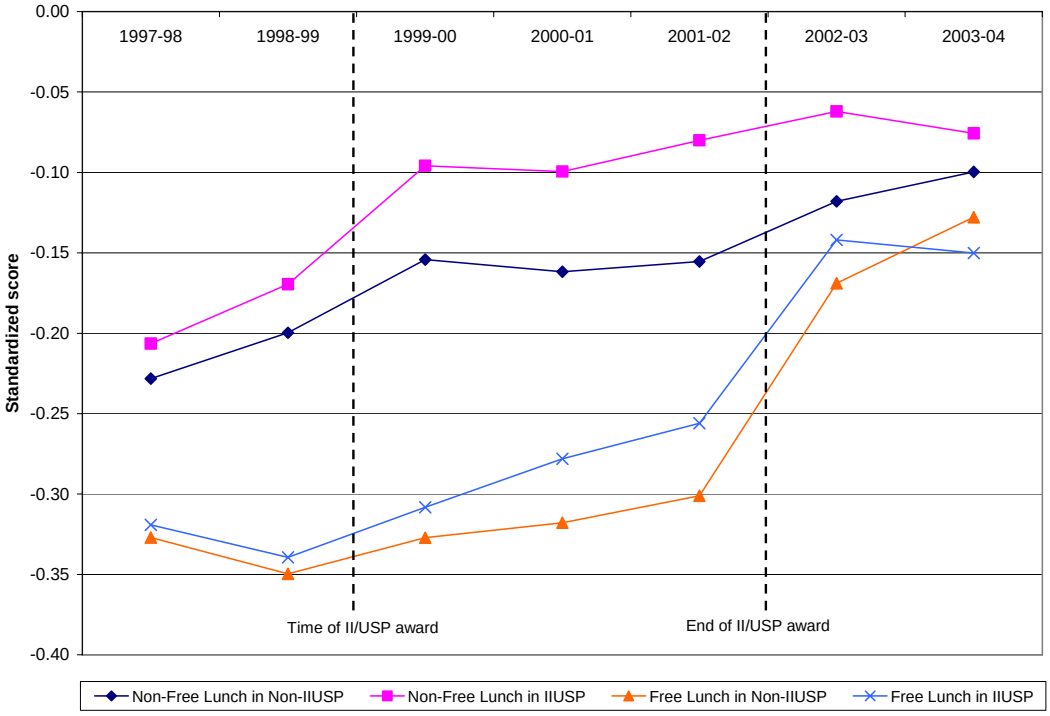
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Students Eligible for Free and Reduced Price Lunch and "Regular" Students in Cohort 3 II/USP and Comparison Elementary Schools



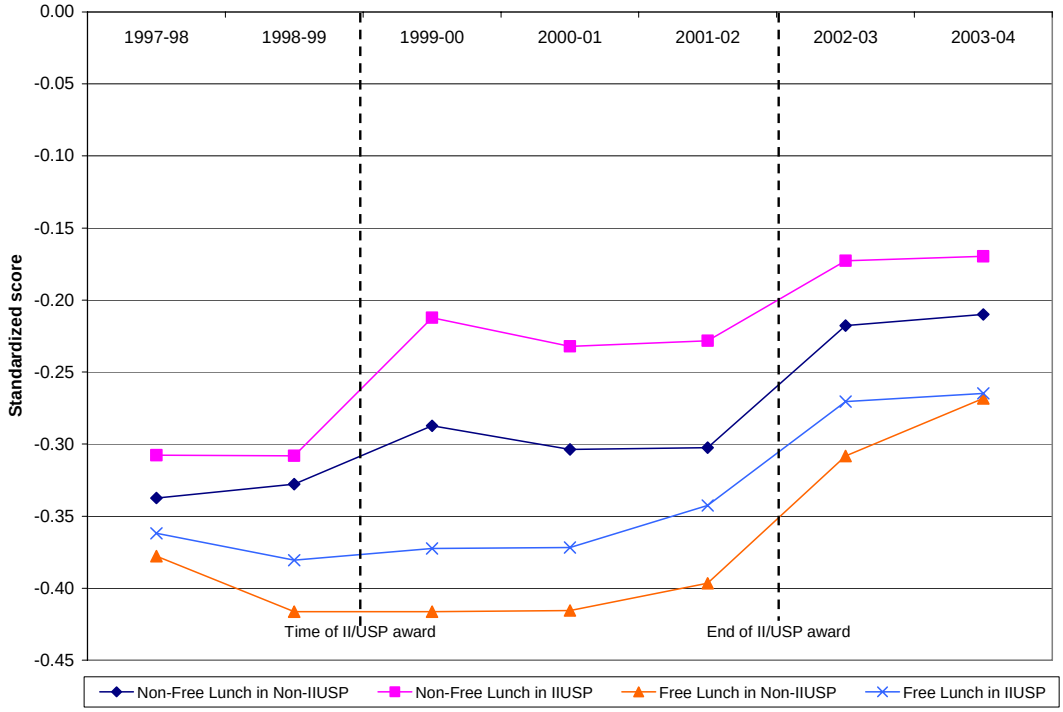
HLM Regression for Students Eligible for Free and Reduced Price Lunch and "Regular" Students in Cohort 1 Middle Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.198	<.0001	0.039	0.2378
GIRL	0.104	<.0001	-0.007	<.0001
ASIAN	-0.063	<.0001	0.263	<.0001
HISPANIC	-0.337	<.0001	-0.314	<.0001
BLACK	-0.516	<.0001	-0.508	<.0001
OTHERS	-0.184	<.0001	-0.095	<.0001
ELMISN	-0.180	<.0001	-0.057	<.0001
EL	-0.563	<.0001	-0.310	<.0001
R_FEP	0.215	<.0001	0.284	<.0001
FEP	0.167	<.0001	0.184	<.0001
FLUNCH	-0.099	<.0001	-0.040	<.0001
HIGHPARED	0.216	<.0001	0.181	<.0001
PARED_MISSING	-0.036	<.0001	-0.036	<.0001
SPECED	-0.730	<.0001	-0.636	<.0001
YEAR99	0.028	<.0001	0.010	0.0432
YEAR00	0.046	<.0001	0.040	<.0001
YEAR01	-0.008	0.1565	-0.016	0.0015
YEAR02	0.006	0.2937	0.001	0.8277
YEAR03	0.037	<.0001	0.085	<.0001
YEAR04	0.018	0.0053	0.008	0.2232
GRADE7	-0.010	<.0001	0.020	<.0001
GRADE8	-0.026	<.0001	-0.003	0.1359
PCT_MEALS	-0.318	<.0001	-0.205	<.0001
IIUSP1	0.022	0.2411	0.030	0.1174
YEAR99*IIUSP1	0.009	0.3468	-0.010	0.2417
YEAR00*IIUSP1	0.028	0.0025	0.055	<.0001
YEAR01*IIUSP1	0.004	0.6666	-0.003	0.7003
YEAR02*IIUSP1	0.013	0.1804	0.003	0.7861
YEAR03*IIUSP1	-0.019	0.0525	-0.029	0.0025
YEAR04*IIUSP1	-0.032	0.0016	-0.005	0.637
FLUNCH*YEAR99	-0.051	<.0001	-0.048	<.0001
FLUNCH*YEAR00	-0.023	0.0002	-0.040	<.0001
FLUNCH*YEAR01	0.017	0.0075	0.017	0.0047
FLUNCH*YEAR02	0.010	0.151	0.018	0.0108
FLUNCH*YEAR03	0.095	<.0001	0.004	0.6103
FLUNCH*YEAR04	0.023	0.0031	0.032	<.0001
FLUNCH*IIUSP1	-0.014	0.1399	-0.014	0.1217
FLUNCH*YEAR99*IIUSP1	-0.006	0.6184	0.030	0.0135
FLUNCH*YEAR00*IIUSP1	-0.019	0.1053	-0.047	<.0001
FLUNCH*YEAR01*IIUSP1	0.017	0.1405	0.003	0.7582
FLUNCH*YEAR02*IIUSP1	-0.008	0.5119	0.008	0.4973
FLUNCH*YEAR03*IIUSP1	0.001	0.9168	0.013	0.2678
FLUNCH*YEAR04*IIUSP1	-0.017	0.1579	-0.030	0.0112

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Students Eligible for Free and Reduced Price Lunch and "Regular" Students in Cohort 1 II/USP and Comparison Middle Schools



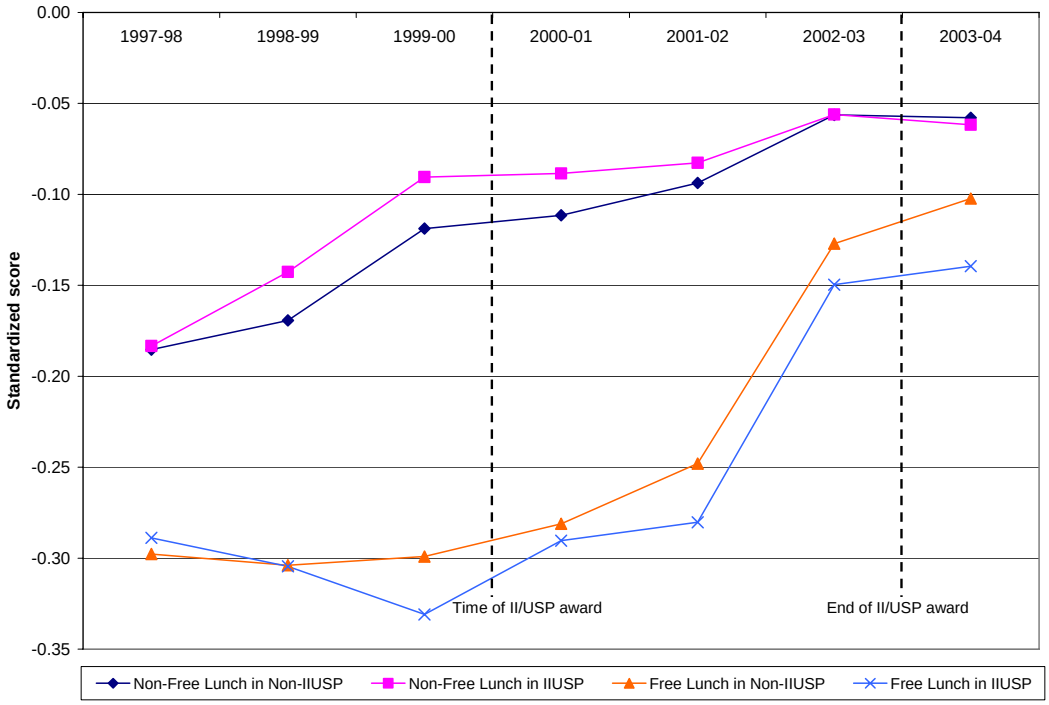
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Students Eligible for Free and Reduced Price Lunch and "Regular" Students in Cohort 1 II/USP and Comparison Middle Schools



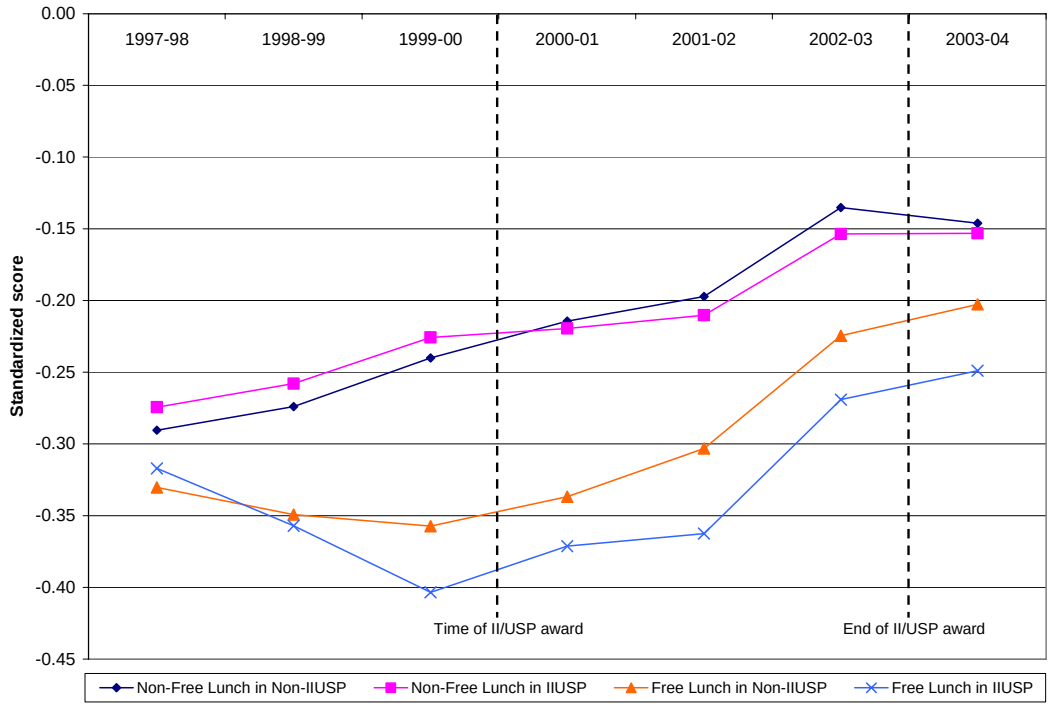
HLM Regression for Students Eligible for Free and Reduced Price Lunch and "Regular" Students in Cohort 2 Middle Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.187	<.0001	0.112	0.0012
GIRL	0.102	<.0001	-0.013	<.0001
ASIAN	-0.014	<.0001	0.269	<.0001
HISPANIC	-0.296	<.0001	-0.291	<.0001
BLACK	-0.487	<.0001	-0.489	<.0001
OTHERS	-0.150	<.0001	-0.075	<.0001
ELMISN	-0.162	<.0001	-0.068	<.0001
EL	-0.566	<.0001	-0.322	<.0001
R_FEP	0.218	<.0001	0.301	<.0001
FEP	0.182	<.0001	0.218	<.0001
FLUNCH	-0.112	<.0001	-0.040	<.0001
HIGHPARED	0.220	<.0001	0.188	<.0001
PARED_MISSING	-0.027	<.0001	-0.032	<.0001
SPECED	-0.747	<.0001	-0.667	<.0001
YEAR99	0.016	0.0026	0.016	0.0014
YEAR00	0.050	<.0001	0.034	<.0001
YEAR01	0.007	0.1796	0.026	<.0001
YEAR02	0.018	0.0015	0.017	0.0017
YEAR03	0.037	<.0001	0.062	<.0001
YEAR04	-0.002	0.796	-0.011	0.0584
GRADE7	-0.007	<.0001	0.006	0.0001
GRADE8	-0.023	<.0001	-0.015	<.0001
PCT_MEALS	-0.304	<.0001	-0.297	<.0001
IIUSP2	0.002	0.9035	0.016	0.374
YEAR99*IIUSP2	0.025	0.0015	0.000	0.9964
YEAR00*IIUSP2	0.002	0.8361	-0.002	0.841
YEAR01*IIUSP2	-0.005	0.5134	-0.020	0.013
YEAR02*IIUSP2	-0.012	0.1463	-0.008	0.3293
YEAR03*IIUSP2	-0.011	0.1917	-0.005	0.496
YEAR04*IIUSP2	-0.004	0.6207	0.011	0.1631
FLUNCH*YEAR99	-0.022	0.0022	-0.035	<.0001
FLUNCH*YEAR00	-0.046	<.0001	-0.042	<.0001
FLUNCH*YEAR01	0.011	0.1081	-0.005	0.422
FLUNCH*YEAR02	0.015	0.0214	0.017	0.0109
FLUNCH*YEAR03	0.083	<.0001	0.016	0.0136
FLUNCH*YEAR04	0.026	0.0002	0.033	<.0001
FLUNCH*IIUSP2	0.007	0.3841	-0.003	0.7227
FLUNCH*YEAR99*IIUSP2	-0.034	0.0015	-0.021	0.0435
FLUNCH*YEAR00*IIUSP2	-0.033	0.0014	-0.037	0.0002
FLUNCH*YEAR01*IIUSP2	0.028	0.0045	0.031	0.001
FLUNCH*YEAR02*IIUSP2	-0.011	0.2533	-0.017	0.0714
FLUNCH*YEAR03*IIUSP2	0.021	0.0367	0.020	0.034
FLUNCH*YEAR04*IIUSP2	-0.010	0.2915	-0.013	0.1722

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Students Eligible for Free and Reduced Price Lunch and "Regular" Students in Cohort 2 II/USP and Comparison Middle Schools



Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Students Eligible for Free and Reduced Price Lunch and "Regular" Students in Cohort 2 II/USP and Comparison Middle Schools

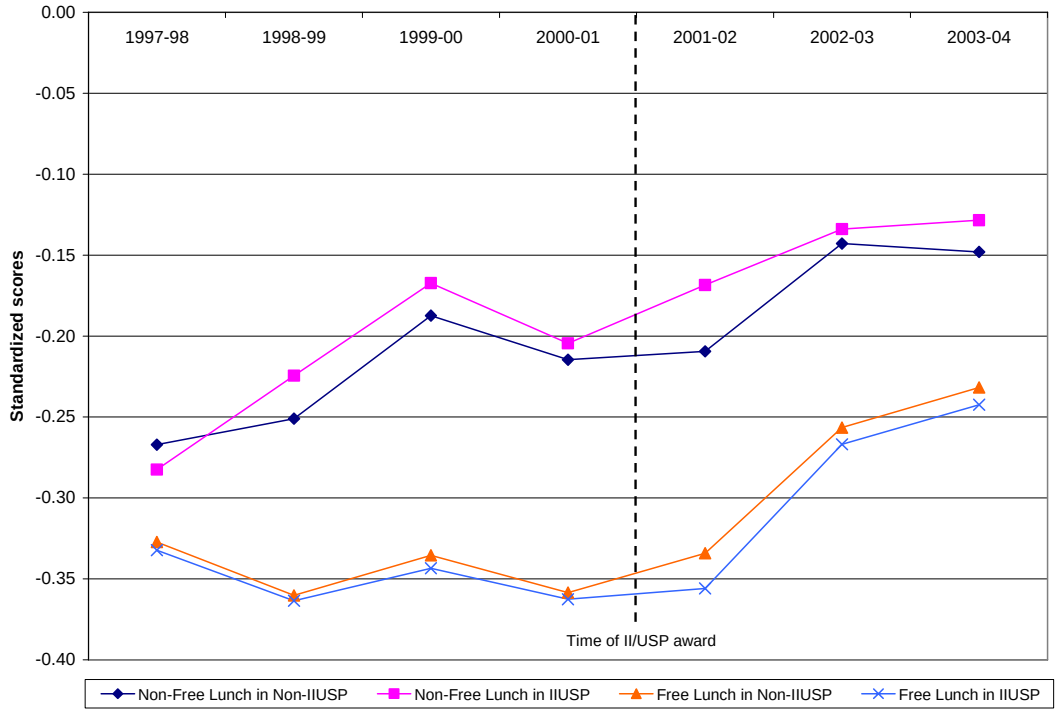


HLM Regression for Students Eligible for Free and Reduced Price Lunch and "Regular" Students in Cohort 3 Middle Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.240	<.0001	0.141	<.0001
GIRL	0.104	<.0001	-0.014	<.0001
ASIAN	-0.064	<.0001	0.229	<.0001
HISPANIC	-0.316	<.0001	-0.291	<.0001
BLACK	-0.481	<.0001	-0.483	<.0001
OTHERS	-0.163	<.0001	-0.070	<.0001
ELMISN	-0.157	<.0001	-0.102	<.0001
EL	-0.560	<.0001	-0.319	<.0001
R_FEP	0.225	<.0001	0.294	<.0001
FEP	0.166	<.0001	0.180	<.0001
FLUNCH	-0.140	<.0001	-0.060	<.0001
HIGHPARED	0.213	<.0001	0.182	<.0001
PARED_MISSING	-0.031	<.0001	-0.030	<.0001
SPECED	-0.767	<.0001	-0.691	<.0001
YEAR99	0.030	<.0001	0.016	0.0003
YEAR00	0.067	<.0001	0.064	<.0001
YEAR01	-0.032	<.0001	-0.027	<.0001
YEAR02	0.006	0.1794	0.005	0.228
YEAR03	0.036	<.0001	0.067	<.0001
YEAR04	-0.001	0.7945	-0.005	0.2348
GRADE7	-0.006	0.0004	0.007	<.0001
GRADE8	-0.023	<.0001	-0.017	<.0001
PCT_MEALS	-0.347	<.0001	-0.315	<.0001
IIUSP3	0.015	0.3973	-0.015	0.4245
YEAR99*IIUSP3	-0.004	0.6907	0.042	<.0001
YEAR00*IIUSP3	0.005	0.6039	-0.007	0.4928
YEAR01*IIUSP3	0.016	0.1006	-0.010	0.2775
YEAR02*IIUSP3	0.011	0.2538	0.031	0.0007
YEAR03*IIUSP3	-0.022	0.0232	-0.032	0.0006
YEAR04*IIUSP3	0.022	0.0261	0.011	0.2671
FLUNCH*YEAR99	-0.040	<.0001	-0.049	<.0001
FLUNCH*YEAR00	-0.024	<.0001	-0.039	<.0001
FLUNCH*YEAR01	0.013	0.0249	0.004	0.4308
FLUNCH*YEAR02	0.016	0.0034	0.019	0.0004
FLUNCH*YEAR03	0.088	<.0001	0.011	0.0362
FLUNCH*YEAR04	0.023	<.0001	0.030	<.0001
FLUNCH*IIUSP3	-0.012	0.1995	0.010	0.2649
FLUNCH*YEAR99*IIUSP3	-0.010	0.4564	-0.040	0.0017
FLUNCH*YEAR00*IIUSP3	-0.004	0.7734	0.002	0.8783
FLUNCH*YEAR01*IIUSP3	0.002	0.841	0.014	0.2204
FLUNCH*YEAR02*IIUSP3	-0.027	0.0195	-0.048	<.0001
FLUNCH*YEAR03*IIUSP3	0.042	0.0003	0.043	0.0001
FLUNCH*YEAR04*IIUSP3	-0.035	0.0032	-0.011	0.3314

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Students Eligible for Free and Reduced Price Lunch and "Regular" Students in Cohort 3 II/USP and Comparison Middle Schools

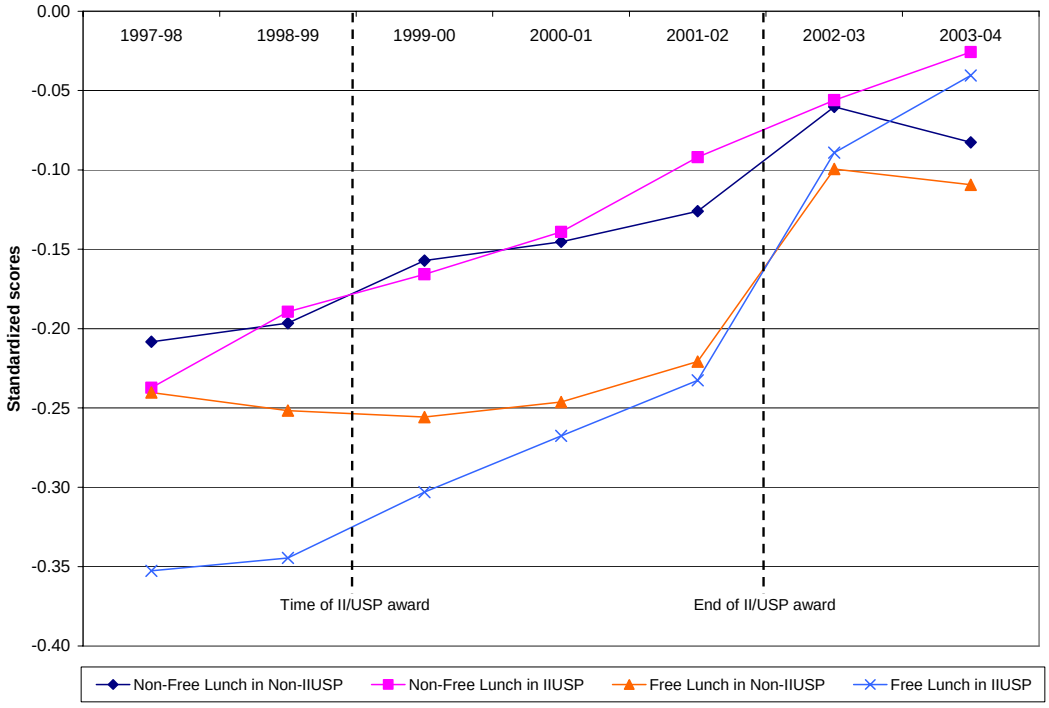
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Students Eligible for Free and Reduced Price Lunch and "Regular" Students in Cohort 3 II/USP and Comparison Middle Schools



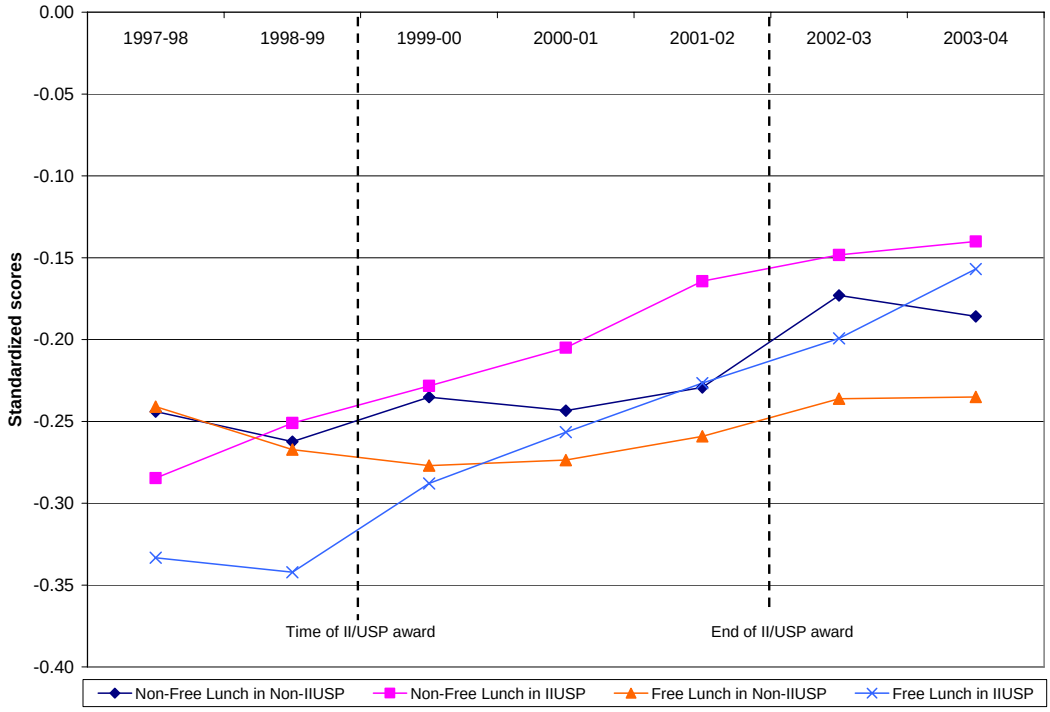
HLM Regression for Students Eligible for Free and Reduced Price Lunch and "Regular" Students in Cohort 1 High Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.059	0.0922	0.097	0.0021
GIRL	0.110	<.0001	-0.065	<.0001
ASIAN	-0.119	<.0001	0.257	<.0001
HISPANIC	-0.327	<.0001	-0.326	<.0001
BLACK	-0.534	<.0001	-0.529	<.0001
OTHERS	-0.213	<.0001	-0.119	<.0001
ELMISN	-0.030	0.0006	0.053	<.0001
EL	-0.615	<.0001	-0.335	<.0001
R_FEP	0.069	<.0001	0.112	<.0001
FEP	0.101	<.0001	0.107	<.0001
FLUNCH	-0.032	<.0001	0.003	0.674
HIGHPARED	0.244	<.0001	0.194	<.0001
PARED_MISSING	-0.018	<.0001	-0.025	<.0001
SPECED	-0.746	<.0001	-0.611	<.0001
YEAR99	0.012	0.0615	-0.018	0.0025
YEAR00	0.039	<.0001	0.027	<.0001
YEAR01	0.012	0.0398	-0.008	0.1427
YEAR02	0.019	0.0008	0.014	0.0107
YEAR03	0.066	<.0001	0.056	<.0001
YEAR04	-0.022	0.0001	-0.013	0.0275
GRADE10	0.004	0.0221	0.016	<.0001
GRADE11	0.017	<.0001	0.002	0.2978
PCT_MEALS	-0.117	0.0605	-0.171	0.0021
IIUSP1	-0.029	0.254	-0.041	0.0732
YEAR99*IIUSP1	0.036	0.0001	0.052	<.0001
YEAR00*IIUSP1	-0.016	0.0789	-0.005	0.6051
YEAR01*IIUSP1	0.015	0.0917	0.032	0.0002
YEAR02*IIUSP1	0.028	0.0014	0.027	0.0018
YEAR03*IIUSP1	-0.030	0.0008	-0.040	<.0001
YEAR04*IIUSP1	0.053	<.0001	0.021	0.0173
FLUNCH*YEAR99	-0.023	0.0119	-0.008	0.3822
FLUNCH*YEAR00	-0.043	<.0001	-0.037	<.0001
FLUNCH*YEAR01	-0.003	0.7449	0.012	0.1182
FLUNCH*YEAR02	0.006	0.4072	0.000	0.9644
FLUNCH*YEAR03	0.056	<.0001	-0.033	<.0001
FLUNCH*YEAR04	0.012	0.099	0.014	0.0607
FLUNCH*IIUSP1	-0.083	<.0001	-0.052	0.0004
FLUNCH*YEAR99*IIUSP1	-0.017	0.361	-0.035	0.0484
FLUNCH*YEAR00*IIUSP1	0.061	<.0001	0.068	<.0001
FLUNCH*YEAR01*IIUSP1	0.011	0.3609	-0.004	0.7677
FLUNCH*YEAR02*IIUSP1	-0.018	0.1315	-0.011	0.3534
FLUNCH*YEAR03*IIUSP1	0.052	<.0001	0.045	0.0002
FLUNCH*YEAR04*IIUSP1	0.006	0.6198	0.020	0.0875

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Students Eligible for Free and Reduced Price Lunch and "Regular" Students in Cohort 1 II/USP and Comparison High Schools



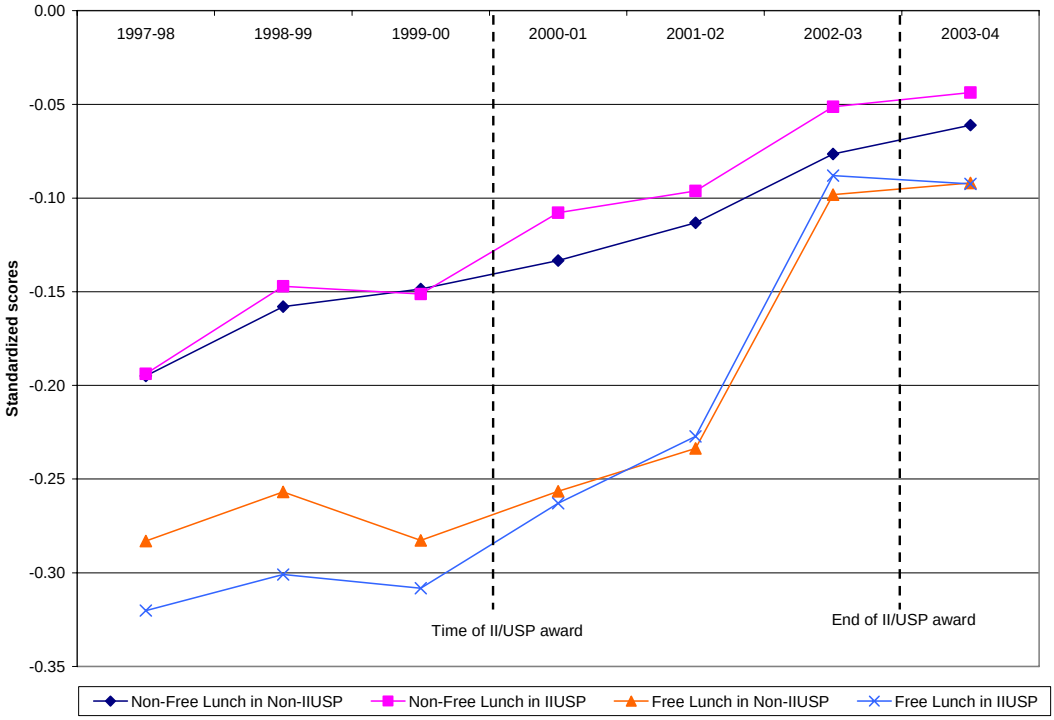
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Students Eligible for Free and Reduced Price Lunch and "Regular" Students in Cohort 1 II/USP and Comparison High Schools



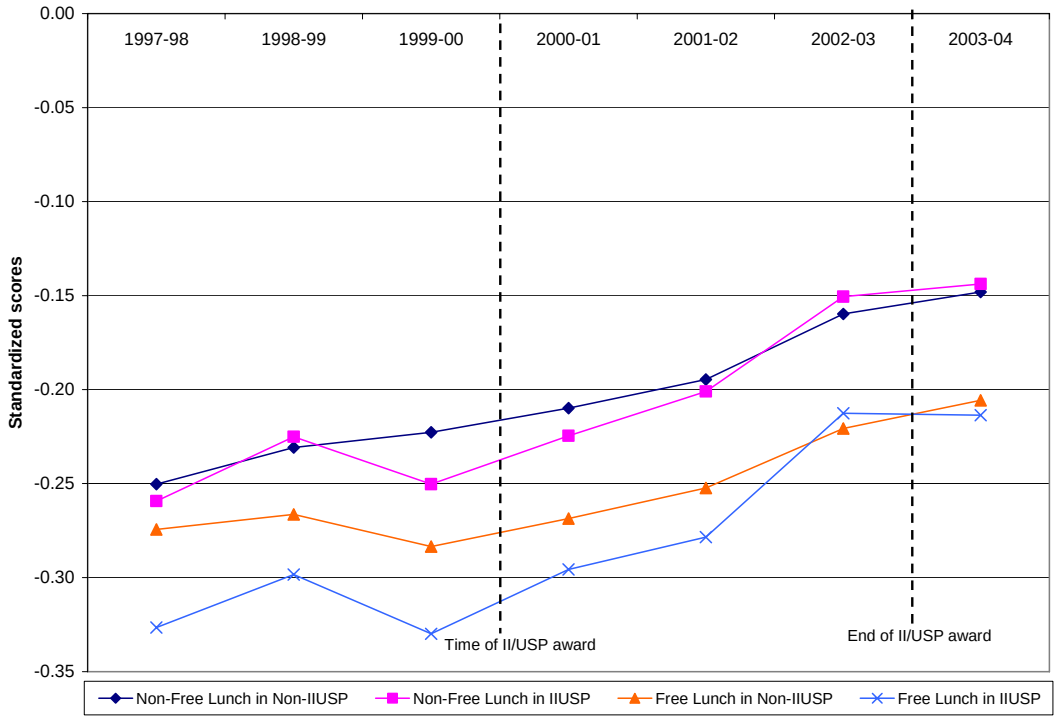
HLM Regression for Students Eligible for Free and Reduced Price Lunch and "Regular" Students in Cohort 2 High Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.056	0.0188	0.081	0.0006
GIRL	0.116	<.0001	-0.067	<.0001
ASIAN	-0.082	<.0001	0.252	<.0001
HISPANIC	-0.311	<.0001	-0.324	<.0001
BLACK	-0.515	<.0001	-0.520	<.0001
OTHERS	-0.181	<.0001	-0.106	<.0001
ELMISN	0.002	0.806	0.052	<.0001
EL	-0.626	<.0001	-0.348	<.0001
R_FEP	0.077	<.0001	0.139	<.0001
FEP	0.129	<.0001	0.147	<.0001
FLUNCH	-0.088	<.0001	-0.024	<.0001
HIGHPARED	0.245	<.0001	0.204	<.0001
PARED_MISSING	-0.009	<.0001	-0.020	<.0001
SPECED	-0.771	<.0001	-0.629	<.0001
YEAR99	0.037	<.0001	0.020	<.0001
YEAR00	0.009	0.0177	0.008	0.0354
YEAR01	0.015	<.0001	0.013	0.0005
YEAR02	0.020	<.0001	0.015	<.0001
YEAR03	0.037	<.0001	0.035	<.0001
YEAR04	0.015	<.0001	0.012	0.0018
GRADE10	-0.007	<.0001	0.006	<.0001
GRADE11	-0.001	0.7023	-0.011	<.0001
PCT_MEALS	-0.129	0.0062	-0.145	0.0017
IIUSP2	0.001	0.9493	-0.009	0.6115
YEAR99*IIUSP2	0.010	0.1723	0.015	0.0321
YEAR00*IIUSP2	-0.013	0.046	-0.033	<.0001
YEAR01*IIUSP2	0.028	<.0001	0.013	0.038
YEAR02*IIUSP2	-0.009	0.1789	0.008	0.1776
YEAR03*IIUSP2	0.008	0.1977	0.016	0.0127
YEAR04*IIUSP2	-0.008	0.2267	-0.005	0.4347
FLUNCH*YEAR99	-0.011	0.1051	-0.012	0.0761
FLUNCH*YEAR00	-0.035	<.0001	-0.025	<.0001
FLUNCH*YEAR01	0.011	0.0408	0.002	0.69
FLUNCH*YEAR02	0.003	0.6058	0.001	0.8454
FLUNCH*YEAR03	0.099	<.0001	-0.003	0.5453
FLUNCH*YEAR04	-0.009	0.0796	0.003	0.536
FLUNCH*IIUSP2	-0.038	<.0001	-0.043	<.0001
FLUNCH*YEAR99*IIUSP2	-0.017	0.152	0.006	0.6259
FLUNCH*YEAR00*IIUSP2	0.032	0.0012	0.019	0.0516
FLUNCH*YEAR01*IIUSP2	-0.009	0.3243	0.006	0.4661
FLUNCH*YEAR02*IIUSP2	0.021	0.0182	-0.007	0.4023
FLUNCH*YEAR03*IIUSP2	-0.004	0.6199	0.019	0.033
FLUNCH*YEAR04*IIUSP2	-0.003	0.7484	-0.011	0.2055

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Students Eligible for Free and Reduced Price Lunch and "Regular" Students in Cohort 2 II/USP and Comparison High Schools



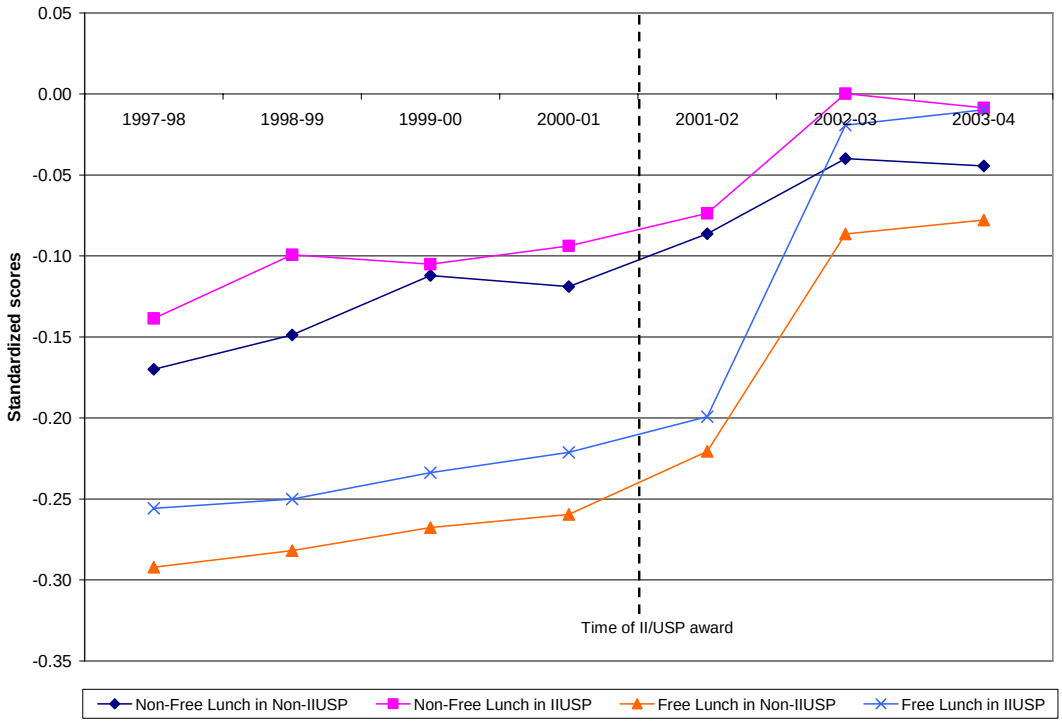
Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Students Eligible for Free and Reduced Price Lunch and "Regular" Students in Cohort 2 II/USP and Comparison High Schools



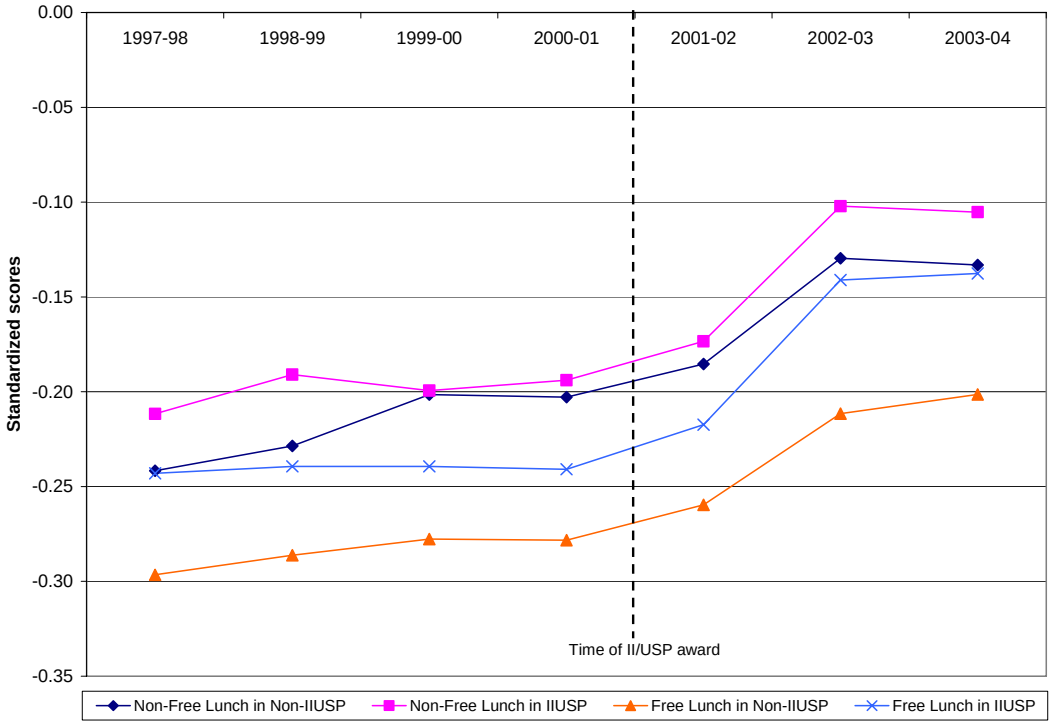
HLM Regression for Students Eligible for Free and Reduced Price Lunch and "Regular" Students in Cohort 3 High Schools, Standardized SAT-9 and CAT/6 Reading and Math Scale Scores

	Reading		Math	
Variables	**Estimate**	**P value**	**Estimate**	**P value**
Intercept	0.039	0.0747	0.033	0.1513
GIRL	0.113	<.0001	-0.073	<.0001
ASIAN	-0.094	<.0001	0.250	<.0001
HISPANIC	-0.318	<.0001	-0.324	<.0001
BLACK	-0.520	<.0001	-0.522	<.0001
OTHERS	-0.195	<.0001	-0.111	<.0001
ELMISN	-0.119	<.0001	-0.017	0.0065
EL	-0.636	<.0001	-0.359	<.0001
R_FEP	0.075	<.0001	0.132	<.0001
FEP	0.102	<.0001	0.122	<.0001
FLUNCH	-0.122	<.0001	-0.055	<.0001
HIGHPARED	0.244	<.0001	0.204	<.0001
PARED_MISSING	-0.007	<.0001	-0.019	<.0001
SPECED	-0.781	<.0001	-0.641	<.0001
YEAR99	0.021	<.0001	0.013	0.0001
YEAR00	0.037	<.0001	0.027	<.0001
YEAR01	-0.007	0.0331	-0.001	0.6833
YEAR02	0.033	<.0001	0.017	<.0001
YEAR03	0.046	<.0001	0.056	<.0001
YEAR04	-0.005	0.1606	-0.004	0.2697
GRADE10	-0.010	<.0001	0.001	0.5074
GRADE11	-0.005	<.0001	-0.017	<.0001
PCT_MEALS	-0.032	0.4715	-0.029	0.5404
IIUSP3	0.031	0.0879	0.030	0.1226
YEAR99*IIUSP3	0.018	0.0038	0.008	0.2093
YEAR00*IIUSP3	-0.042	<.0001	-0.036	<.0001
YEAR01*IIUSP3	0.018	0.0016	0.007	0.227
YEAR02*IIUSP3	-0.013	0.0274	0.003	0.5902
YEAR03*IIUSP3	0.028	<.0001	0.016	0.0064
YEAR04*IIUSP3	-0.004	0.4647	0.000	0.9542
FLUNCH*YEAR99	-0.011	0.0613	-0.003	0.6435
FLUNCH*YEAR00	-0.022	<.0001	-0.019	0.0002
FLUNCH*YEAR01	0.015	0.0014	0.001	0.8728
FLUNCH*YEAR02	0.006	0.1758	0.001	0.7816
FLUNCH*YEAR03	0.088	<.0001	-0.008	0.0847
FLUNCH*YEAR04	0.013	0.0037	0.014	0.0021
FLUNCH*IIUSP3	0.005	0.5584	0.024	0.0036
FLUNCH*YEAR99*IIUSP3	-0.023	0.0362	-0.014	0.171
FLUNCH*YEAR00*IIUSP3	0.045	<.0001	0.027	0.0025
FLUNCH*YEAR01*IIUSP3	-0.014	0.0937	-0.008	0.3305
FLUNCH*YEAR02*IIUSP3	-0.004	0.5974	0.002	0.8112
FLUNCH*YEAR03*IIUSP3	0.018	0.0238	0.013	0.11
FLUNCH*YEAR04*IIUSP3	0.005	0.5138	-0.007	0.3659

Relative Reading Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Students Eligible for Free and Reduced Price Lunch and "Regular" Students in Cohort 3 II/USP and Comparison High Schools



Relative Math Performance (Standardized SAT-9 and CAT/6 Scale Scores) for Students Eligible for Free and Reduced Price Lunch and "Regular" Students in Cohort 3 II/USP and Comparison High Schools



Appendix A-10:
Analysis of the CDE Growth Target Requirements

Elementary Schools

Cohort 1

II/USP

Light Status	Yearly Number of Schools Achieving Green for the First Time after Planning Year								
	99 - 00	00 - 01		01 - 02		02 - 03		03 - 04	
	P	**I1**		**I2**					
	N	N	%	N	%	N	%	N	%
Yellow	54	31	57	8	15	10	19		
Red	13	7	54	2	15	2	15		

Comparison

	99 - 00	00 - 01		01 - 02		02 - 03		03 - 04	
	P	**I1**		**I2**					
	N	N	%	N	%	N	%	N	%
Yellow	114	64	56	25	22	23	20	1	1
Red	45	31	69	7	16	4	9	1	2

Cohort 2

II/USP

	99 - 00	00 - 01		01 - 02		02 - 03		03 - 04	
	P - 1	**P**		**I1**		**I2**			
	N	N		N	%	N	%	N	%
Yellow	186	47		25	53	19	40		
Red	67	24		12	50	9	38	2	8

Comparison

	99 - 00	00 - 01		01 - 02		02 - 03		03 - 04	
	P - 1	**P**		**I1**		**I2**			
	N	N		N	%	N	%	N	%
Yellow	165	55		28	51	22	40	2	4
Red	49	31		16	52	11	35	2	6

Cohort 3

II/USP

	99 - 00	00 - 01	01 - 02	02 - 03		03 - 04	
	P - 2	**P - 1**	**P**	**I1**		**I2**	
	N	N	N	N	%	N	%
Yellow	31	165	67	54	81	4	6
Red	8	116	46	38	83	4	9

Comparison

	99 - 00	00 - 01	01 - 02	02 - 03		03 - 04	
	P - 2	**P - 1**	**P**	**I1**		**I2**	
	N	N	N	N	%	N	%
Yellow	67	285	120	96	80	5	4
Red	25	188	85	74	87	2	2

Middle Schools

		Light Status	Yearly Number of Schools Achieving Green for the First Time after Planning Year							
		99 - 00	00 - 01		01 - 02		02 - 03		03 - 04	
		P	**I1**		**I2**					
		N	N	%	N	%	N	%	N	%
II/USP	Yellow	23	11	48	2	9	7	30	2	9
	Red	9	3	33	4	44	2	22		

Cohort 1

		99 - 00	00 - 01		01 - 02		02 - 03		03 - 04	
		P	**I1**		**I2**					
		N	N	%	N	%	N	%	N	%
Comparison	Yellow	57	25	44	5	9	19	33	3	5
	Red	29	12	41	1	3	10	34	2	7

		99 - 00	00 - 01		01 - 02		02 - 03		03 - 04	
		P - 1	**P**		**I1**		**I2**			
		N	N		N	%	N	%	N	%
II/USP	Yellow	66	35		9	26	15	43	4	11
	Red	36	18		4	22	10	56	3	17

Cohort 2

		99 - 00	00 - 01		01 - 02		02 - 03		03 - 04	
		P - 1	**P**		**I1**		**I2**			
		N	N		N	%	N	%	N	%
Comparison	Yellow	87	42		12	29	20	48	5	12
	Red	34	23		12	52	9	39	2	9

		99 - 00	00 - 01	01 - 02		02 - 03		03 - 04	
		P - 2	**P - 1**	**P**		**I1**		**I2**	
		N	N	N		N	%	N	%
II/USP	Yellow	12	30	19		15	79	2	11
	Red	4	21	19		13	68	4	21

Cohort 3

		99 - 00	00 - 01	01 - 02		02 - 03		03 - 04	
		P - 2	**P - 1**	**P**		**I1**		**I2**	
		N	N	N		N	%	N	%
Comparison	Yellow	48	97	70		46	66	16	23
	Red	20	81	50		34	68	10	20

High Schools

		Light Status	Yearly Number of Schools Achieving Green for the First Time after Planning Year							
		99 - 00	00 - 01		01 - 02		02 - 03		03 - 04	
		P	**I1**		**I2**					
		N	N	%	N	%	N	%	N	%
II/USP	Yellow	16	4	25	2	13	4	25	2	13
	Red	14	4	29	4	29	1	7	2	14

Cohort 1

		99 - 00	00 - 01		01 - 02		02 - 03		03 - 04	
		P	**I1**		**I2**					
		N	N	%	N	%	N	%	N	%
Comparison	Yellow	57	4	7	9	16	24	42	6	11
	Red	33	6	18	5	15	8	24		

		99 - 00	00 - 01		01 - 02		02 - 03		03 - 04	
		P - 1	**P**		**I1**		**I2**			
		N	N		N	%	N	%	N	%
II/USP	Yellow	25	32		5	16	15	47	3	9
	Red	45	18		6	33	6	33	1	6

Cohort 2

		99 - 00	00 - 01		01 - 02		02 - 03		03 - 04	
		P - 1	**P**		**I1**		**I2**			
		N	N		N	%	N	%	N	%
Comparison	Yellow	88	60		8	13	32	53	10	17
	Red	58	56		17	30	17	30	9	16

		99 - 00	00 - 01	01 - 02	02 - 03		03 - 04	
		P - 2	**P - 1**	**P**	**I1**		**I2**	
		N	N	N	N	%	N	%
II/USP	Yellow	17	34	25	17	68	3	12
	Red	26	42	30	19	63	3	10

Cohort 3

		99 - 00	00 - 01	01 - 02	02 - 03		03 - 04	
		P - 2	**P - 1**	**P**	**I1**		**I2**	
		N	N	N	N	%	N	%
Comparison	Yellow	63	72	59	30	51	12	20
	Red	31	87	49	31	63	8	16

Appendix A-11:
District Influences on School Performance Set of Regression Results

Within-District Regression for Low Performing Schools

Dependent Variable: API

	Elementary	
Variables	**Estimate**	**P value**
Intercept	672.050	<.0001
Percent African American	-2.430	<.0001
Percent Asian	0.778	0.0228
Percent Hispanic	-1.603	<.0001
Percent Eligible for Free or Reduced Price Lunch	-0.429	0.0108
Percent English Learners	-1.631	<.0001
Percentage of Students Who First Attended this School in the Present Year	-0.377	0.0137
Average Parental Education	-0.020	0.9925
Percent Teachers at This School with Full Credentials	0.760	<.0001
II/USP Cohort 1	-0.182	0.9887
CSR1	-20.586	0.1323
YEAR99	34.484	<.0001
YEAR00	46.117	<.0001
YEAR01	42.752	<.0001
YEAR02	48.319	<.0001
YEAR03	46.027	<.0001
YEAR04	6.758	0.0234
IIUSP1*YEAR99	6.141	0.4689
IIUSP1*YEAR00	-5.748	0.4978
IIUSP1*YEAR01	-0.894	0.9181
IIUSP1*YEAR02	-2.644	0.7633
IIUSP1*YEAR03	3.588	0.6749
IIUSP1*YEAR04	7.919	0.3482
District X	-1.568	0.9095
District Z	-45.425	0.0007
District W	-35.163	0.0008
IIUSP1*District X	11.195	0.6226
IIUSP1*District Z	10.730	0.6388
IIUSP1*District W	-3.521	0.8727
YEAR99*District X	23.237	0.0338
YEAR99*District Z	-2.207	0.8212
YEAR99*District W	26.302	0.0001
YEAR00*District X	-3.753	0.7316
YEAR00*District Z	0.169	0.9862
YEAR00*District W	-29.858	<.0001
YEAR01*District X	-29.221	0.0135
YEAR01*District Z	-34.996	0.002
YEAR01*District W	-26.646	0.0065
YEAR02*District X	-23.883	0.1043
YEAR02*District Z	-36.517	0.0076

Variables	Elementary	
	Estimate	P value
YEAR02*District W	-5.164	0.6741
YEAR03*District X	-7.054	0.6623
YEAR03*District Z	11.917	0.4216
YEAR03*District W	-3.729	0.7731
YEAR04*District X	7.792	0.6283
YEAR04*District Z	-1.009	0.9458
YEAR04*District W	-29.755	0.0265
IIUSP1*YEAR99*District X	10.780	0.5408
IIUSP1*YEAR99*District Z	8.773	0.6225
IIUSP1*YEAR99*District W	-12.510	0.4632
IIUSP1*YEAR00*District X	12.384	0.4822
IIUSP1*YEAR00*District Z	5.087	0.7752
IIUSP1*YEAR00*District W	6.633	0.6973
IIUSP1*YEAR01*District X	-9.735	0.593
IIUSP1*YEAR01*District Z	-6.821	0.7133
IIUSP1*YEAR01*District W	1.983	0.917
IIUSP1*YEAR02*District X	0.027	0.9989
IIUSP1*YEAR02*District Z	-13.361	0.5096
IIUSP1*YEAR02*District W	-9.328	0.6514
IIUSP1*YEAR03*District X	-9.212	0.6667
IIUSP1*YEAR03*District Z	-23.046	0.2713
IIUSP1*YEAR03*District W	-6.118	0.7701
IIUSP1*YEAR04*District X	-9.481	0.6587
IIUSP1*YEAR04*District Z	-8.948	0.6699
IIUSP1*YEAR04*District W	36.050	0.0891

Within-District Regression for Low Performing Schools

Dependent Variable: Standardized SAT-9 and CAT/6 Reading Scale Scores

	Elementary	
Effect	**Estimate**	**P value**
Intercept	0.786	<.0001
Female	0.094	<.0001
Asian	0.117	<.0001
Hispanic	-0.228	<.0001
African American	-0.388	<.0001
Other Ethnicity	-0.054	<.0001
English Proficiency Missing	-0.265	<.0001
English Learner	-0.364	<.0001
Redesignated Fluent English Proficient	0.373	<.0001
Fluent English Proficient	0.222	<.0001
Eligible for Free or Reduced Price Lunch	-0.171	<.0001
College Parental Education or Higher	0.162	<.0001
Parental Education Missing	-0.028	<.0001
Receives Special Education Services	-0.543	<.0001
YEAR99	-0.003	0.4647
YEAR00	0.050	<.0001
YEAR01	0.065	<.0001
YEAR02	0.062	<.0001
YEAR03	0.111	<.0001
YEAR04	0.023	<.0001
GRADE3	-0.047	<.0001
GRADE4	-0.100	<.0001
GRADE5	-0.142	<.0001
Percent of Students in the School Eligible for Free or Reduced Price Lunch	-0.817	<.0001
II/USP Cohort 1	-0.058	0.0444
YEAR99*II/USP Cohort 1	0.015	0.1721
YEAR00*II/USP Cohort 1	0.006	0.5282
YEAR01*II/USP Cohort 1	-0.004	0.6948
YEAR02*II/USP Cohort 1	-0.035	0.0002
YEAR03*II/USP Cohort 1	0.026	0.0065
YEAR04*II/USP Cohort 1	0.004	0.6394
District X	-0.298	<.0001
District Z	0.048	0.4438
District W	-0.335	<.0001
IIUSP1*District X	0.054	0.3384
IIUSP1*District Z	0.048	0.662
IIUSP1*District W	0.039	0.5024
YEAR99*District X	0.241	<.0001
YEAR99*District Z	-0.152	0.0045
YEAR99*District W	0.237	<.0001
YEAR00*District X	-0.035	0.0257

Effect	Elementary	
	Estimate	P value
YEAR00*District Z	0.041	0.0501
YEAR00*District W	-0.164	<.0001
YEAR01*District X	0.025	0.116
YEAR01*District Z	-0.129	<.0001
YEAR01*District W	-0.016	0.3764
YEAR02*District X	-0.072	<.0001
YEAR02*District Z	-0.106	0.0002
YEAR02*District W	0.039	0.0875
YEAR03*District X	0.008	0.6857
YEAR03*District Z	-0.015	0.638
YEAR03*District W	-0.085	0.0006
YEAR04*District X	-0.030	0.1331
YEAR04*District Z	-0.033	0.2989
YEAR04*District W	-0.060	0.0198
IIUSP1*YEAR99*District X	0.003	0.9004
IIUSP1*YEAR99*District Z	-0.003	0.9754
IIUSP1*YEAR99*District W	-0.040	0.1368
IIUSP1*YEAR00*District X	0.037	0.1176
IIUSP1*YEAR00*District Z	-0.012	0.7407
IIUSP1*YEAR00*District W	-0.003	0.8908
IIUSP1*YEAR01*District X	-0.058	0.0119
IIUSP1*YEAR01*District Z	0.009	0.8071
IIUSP1*YEAR01*District W	-0.049	0.0727
IIUSP1*YEAR02*District X	0.073	0.0023
IIUSP1*YEAR02*District Z	0.021	0.5954
IIUSP1*YEAR02*District W	0.018	0.5625
IIUSP1*YEAR03*District X	0.002	0.9377
IIUSP1*YEAR03*District Z	-0.112	0.0085
IIUSP1*YEAR03*District W	0.017	0.5935
IIUSP1*YEAR04*District X	0.022	0.3824
IIUSP1*YEAR04*District Z	0.005	0.9059
IIUSP1*YEAR04*District W	0.063	0.0564

Within-District Regression for Low Performing Schools

Dependent Variable: Standardized SAT-9 and CAT/6 Math Scale Scores

	Elementary	
Effect	**Estimate**	**P value**
Intercept	0.693	<.0001
Female	-0.019	<.0001
Asian	0.343	<.0001
Hispanic	-0.240	<.0001
African American	-0.466	<.0001
Other Ethnicity	-0.018	0.0068
English Proficiency Missing	-0.117	<.0001
English Learner	-0.147	<.0001
Redesignated Fluent English Proficient	0.539	<.0001
Fluent English Proficient	0.260	<.0001
Eligible for Free or Reduced Price Lunch	-0.144	<.0001
College Parental Education or Higher	0.149	<.0001
Parental Education Missing	-0.043	<.0001
Receives Special Education Services	-0.639	<.0001
YEAR99	-0.015	<.0001
YEAR00	-0.015	<.0001
YEAR01	0.021	<.0001
YEAR02	0.061	<.0001
YEAR03	0.099	<.0001
YEAR04	0.017	<.0001
GRADE3	-0.018	<.0001
GRADE4	-0.091	<.0001
GRADE5	-0.155	<.0001
Percent of Students in the School Eligible for Free or Reduced Price Lunch	-0.649	<.0001
II/USP Cohort 1	-0.047	0.159
YEAR99*II/USP Cohort 1	0.003	0.7892
YEAR00*II/USP Cohort 1	-0.006	0.5951
YEAR01*II/USP Cohort 1	0.004	0.6781
YEAR02*II/USP Cohort 1	0.001	0.8847
YEAR03*II/USP Cohort 1	0.021	0.0345
YEAR04*II/USP Cohort 1	0.008	0.4023
District X	-0.243	<.0001
District Z	-0.088	0.1901
District W	-0.390	<.0001
IIUSP1*District X	0.052	0.4205
IIUSP1*District Z	0.252	0.0335
IIUSP1*District W	0.132	0.0497
YEAR99*District X	0.159	<.0001
YEAR99*District Z	-0.076	0.1667
YEAR99*District W	0.225	<.0001
YEAR00*District X	0.028	0.0952

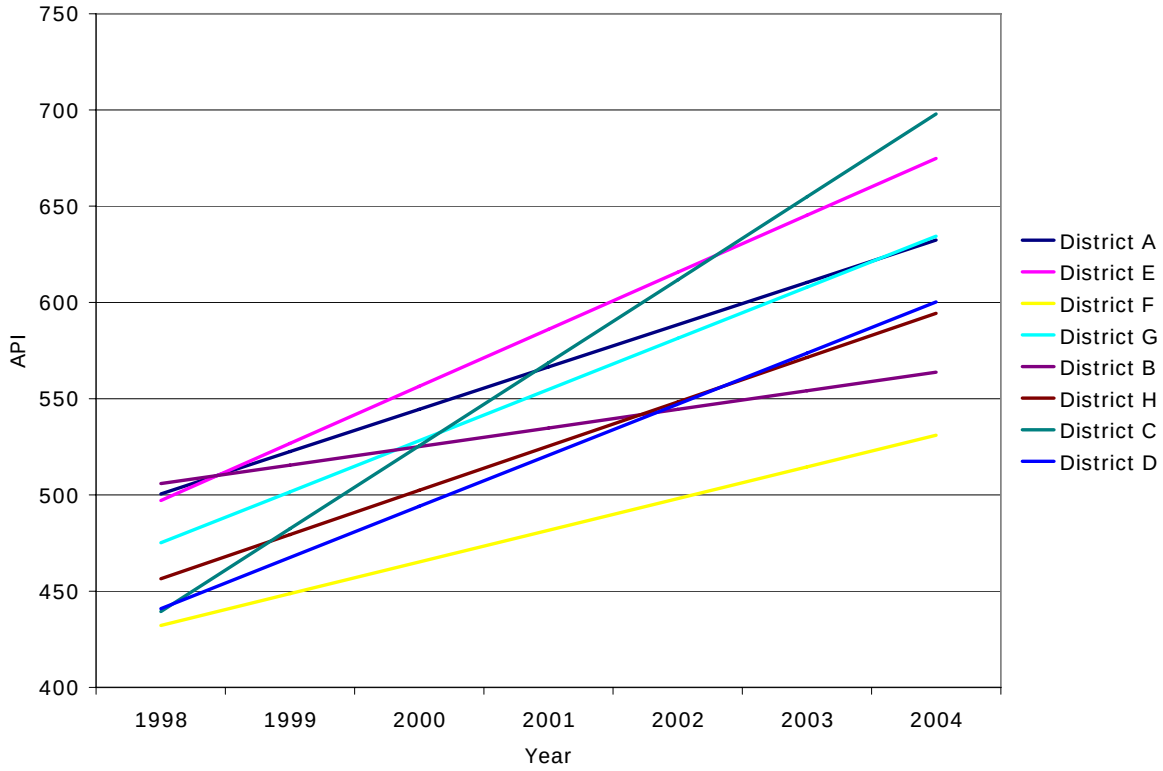
Effect	Elementary	
	Estimate	P value
YEAR00*District Z	0.081	0.0003
YEAR00*District W	-0.094	<.0001
YEAR01*District X	-0.002	0.8834
YEAR01*District Z	-0.125	<.0001
YEAR01*District W	-0.026	0.1813
YEAR02*District X	-0.131	<.0001
YEAR02*District Z	-0.099	0.0012
YEAR02*District W	-0.013	0.5966
YEAR03*District X	0.035	0.0882
YEAR03*District Z	0.031	0.3621
YEAR03*District W	-0.013	0.6192
YEAR04*District X	-0.004	0.8342
YEAR04*District Z	-0.032	0.3502
YEAR04*District W	0.000	0.9939
IIUSP1*YEAR99*District X	0.048	0.0619
IIUSP1*YEAR99*District Z	-0.031	0.7564
IIUSP1*YEAR99*District W	0.007	0.8121
IIUSP1*YEAR00*District X	0.070	0.0049
IIUSP1*YEAR00*District Z	-0.040	0.2842
IIUSP1*YEAR00*District W	-0.004	0.8651
IIUSP1*YEAR01*District X	-0.081	0.0009
IIUSP1*YEAR01*District Z	-0.012	0.755
IIUSP1*YEAR01*District W	-0.081	0.0054
IIUSP1*YEAR02*District X	0.016	0.5334
IIUSP1*YEAR02*District Z	-0.038	0.3675
IIUSP1*YEAR02*District W	-0.053	0.1062
IIUSP1*YEAR03*District X	-0.053	0.0477
IIUSP1*YEAR03*District Z	-0.111	0.0141
IIUSP1*YEAR03*District W	0.037	0.2853
IIUSP1*YEAR04*District X	0.021	0.4507
IIUSP1*YEAR04*District Z	-0.052	0.2653
IIUSP1*YEAR04*District W	0.023	0.5118

Across-District Regression for Low Performing Schools

Dependent Variable: API

	Elementary	
Variables	**Estimate**	**P value**
Intercept	535.50	<.0001
Percent African American	-1.49	<.0001
Percent Asian	-0.23	0.0001
Percent Hispanic	-1.05	<.0001
Percent Eligible for Free or Reduced Price Lunch	-0.19	<.0001
Percent English Learners	-0.89	<.0001
Percentage of Students Who First Attended this School in the Present Year	0.08	0.0056
Average Parental Education	9.88	<.0001
Percent Teachers at This School with Full Credentials	0.74	<.0001
Year	26.56	<.0001
District A	25.46	0.0005
District B	22.09	0.0102
District C	3.50	0.7166
District D	-18.70	0.0146
District E	30.81	<.0001
District F	-35.67	0.0003
District G	-34.17	<.0001
District H	-42.90	<.0001
Year * District A	-4.59	0.0007
Year * District B	3.06	0.0529
Year * District C	-0.01	0.9955
Year * District D	-3.56	0.0067
Year * District E	-16.92	<.0001
Year * District F	16.53	<.0001
Year * District G	1.88	0.2314
Year * District H	-10.08	<.0001

Average Test Score Trajectories of II/USP and Comparison Elementary Schools in Nine School Districts in California



Appendix B:
Supplementary Exhibits

Appendix B-1:
II/USP Legislative Changes

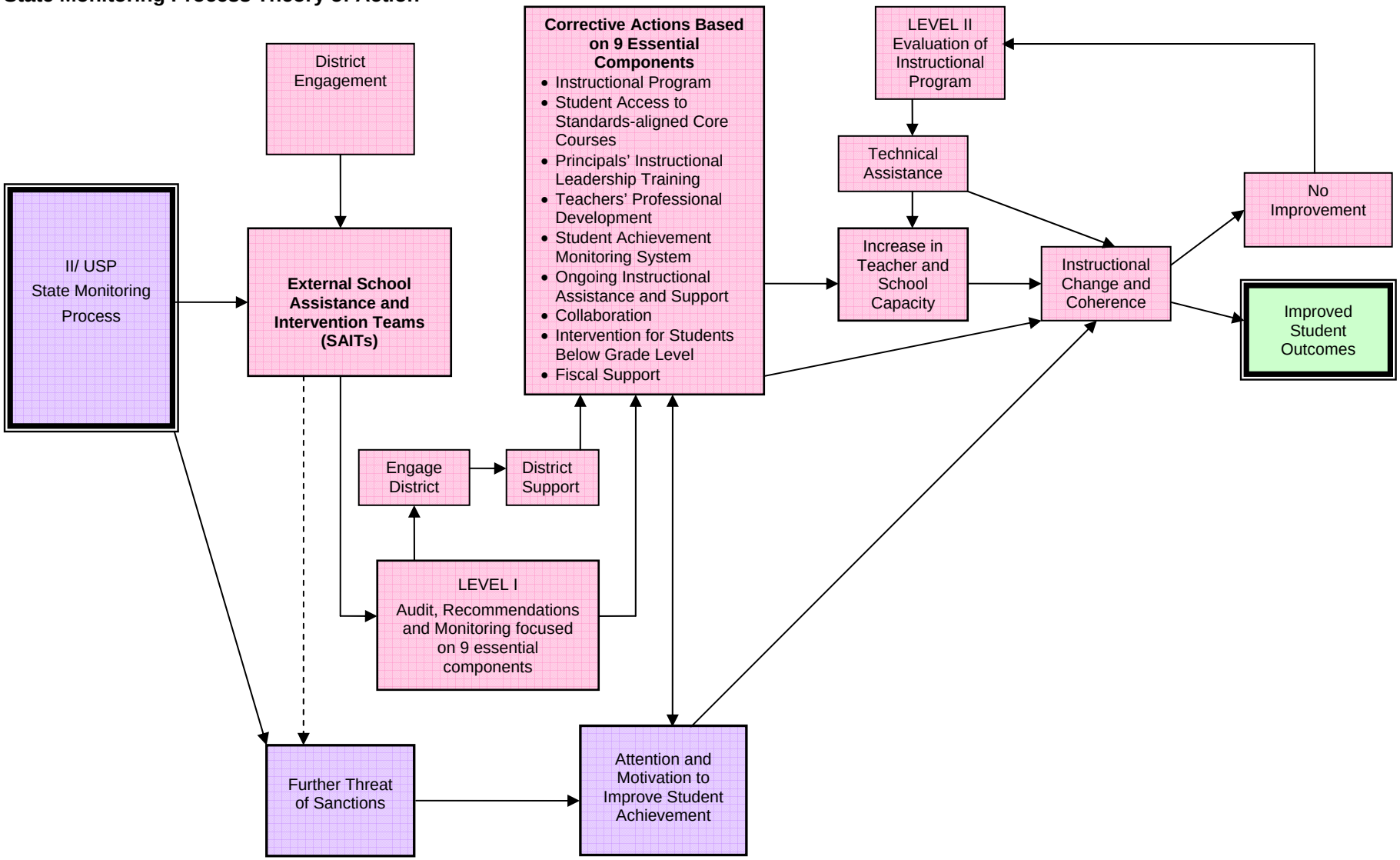
Appendix B-1: Summary of legislative and administrative modifications for each II/USP cohort

Group	Funding Source	II/USP Eligibility	Application Process	Selection	External Evaluator (E.E.)	Action Plan
Cohort 1	State-funded	Scored in the bottom half of the statewide distribution of STAR testing in spring 1998 and spring 1999	Districts provided with list of eligible schools in Sept.1999. Districts provided state with names of schools volunteering for program in Oct. 1999.	Schools selected Oct. 1999. Random selection by grade level, considering balance across decile ranks.***	Schools could select E.E. from list of approved E.E.'s. E.E. approval process involved a written application and interviews with E.E. organization representatives and individual applicants.	Action Plan due to the SBE for approval April 15, 2000.
	CSRD	Eligible for CSRD program*	Schools applied for CSRD program in summer 1999.	Schools selected Sept. 1999 for CSRD and automatically placed into II/USP program.	No E.E. required.	No Action Plan required. Joint district and school CSRD applications submitted.
Cohort 2	State-funded	API in the bottom five deciles in spring 2000 and did not meet growth targets in the 1999-2000 school year	Districts provided with list of eligible schools in Sept. 2000. Districts provided state with names of schools volunteering for program in Oct. 2000.	Schools selected Oct. 2000. Random selection by grade level, considering balance across decile ranks.***	Schools could select E.E. from list of approved E.E.'s. Approved list was the same as list for Cohort 1, with additional groups and individual applicants.	Action Plan with stricter guidelines than Cohort 1 due May 15, 2001.
	CSRD		II/USP schools could apply for CSRD program in spring 2001, after 1 year of planning.	Schools selected for II/USP Oct. 2000, selected for CSRD summer 2001.		Joint district and school CSRD applications due May 15, 2001.
Cohort 3	State-funded	API in the bottom five deciles in spring 2001 and did not meet growth targets in the 2000-01 school year	Districts provided with list of eligible schools in Sept. 2001. District provided state with names of schools volunteering for program in Oct. 2001.	Schools selected Oct. 2001. Random selection by grade level, considering balance across decile ranks.***	Schools could select E.E. from new list of approved E.E.'s. E.E. approval process involved a written application and demonstrated success with underperforming schools. All individual E.E.'s had to be approved. Due to a temporary change in law, during Oct-Nov 2001 schools could select E.E. outside of the list. Use of approved E.E. list reinstated January 1, 2002.	Narrative Summary of Key Elements of the Action Plan due May 15, 2002.
	CSRD		II/USP schools could apply for CSRD program in spring 2002, after one year of planning.	Schools selected for II/USP Oct. 2001, selected for CSRD summer 2002.**		Joint district and school CSRD applications due May 15, 2002.

Note: This chart is from the 2003 PSAA Evaluation Report.
* Criteria included: identified by Program Improvement and on the certified Program Improvement list due to performance on locally-determined measures; scored in bottom half of the statewide distribution of STAR testing in spring of 1998 and spring of 1999. CSRD is now called the "Comprehensive School Reform" (CSR) Program.
** See http://www.cde.ca.gov/iasa/csrd/rfa.html for a scoring rubric used in CSRD selection for 2002.
*** Schools were first randomly selected from applications submitted by a state-designated date. In one year fewer applications were received, and therefore additional schools were selected from applications arriving after the initial date.

Appendix B-2:
State-Monitoring Process Theory of Action

State Monitoring Process Theory of Action



Appendix B-3:
Table of “Growth” School Staff Disaggregated Responses

Appendix B-3: "Growth" School Staff Disaggregated Responses to the Question: "What factors contributed to your school's growth?" by detailed Factors for Growth

		School Administrators		Teachers	
Detailed Factors for Growth		**N**	**%**	**N**	**%**
School and Staff Capacity					
A.	Leadership	2	4%	7	14%
B.	Instructional coaches/support	2	4%	3	6%
C.	Professional development	5	9%	3	6%
D.	Professional community/teacher collaboration	7	13%	9	18%
E.	School culture/climate	8	15%	3	6%
F.	Experience, qualifications and characteristics of instructional staff	3	5%	3	6%
G.	Resources: funding, etc.	0	0%	0	0%
Curriculum/Instruction					
H.	Coherent curriculum/curriculum & instruction tied to standards*	10	18%	8	16%
I.	Vision and common goals for instruction	3	5%	1	2%
J.	Additional instructional time for students	1	2%	0	0%
K.	Whole school reform model	0	0%	3	6%
Systematic Assessment and Data Based Decision Making					
L.	Use of data and monitoring of student learning	9	16%	3	6%
M.	Attention to student outcomes	2	4%	3	6%
District					
N.	District support	0	0%	1	2%
Other					
O.	Parent involvement	1	2%	0	0%
P.	Other	2	4%	2	4%
Q.	No Response	5		5	
Total		**60**	**100**	**54**	**100**

Appendix B-4:
Essential Program Components of the State Monitoring Process by Grade Level

Essential Program Components (EPCs) of the State Monitoring Process by Grade Level

Essential Program Components for Grades K-8:

1. Use of State Board-adopted core instructional programs
2. Instructional time: protected ELA and mathematics periods.
3. District provides principals and vice principals with appropriate trainings. (e.g. AB 75)
4. Fully credentialed teachers in all grade levels and plan for training the remaining teachers.
5. District/school implements assessment system including monitoring and reporting of student progress. Assessments based on core programs.
6. District/school provides instructional assistance and support to teachers (e.g. coaches/content experts)
7. District/school facilitates and supports teacher collaboration across grade levels and programs.
8. District/schoolwide lesson pacing schedule for each grade level or program.
9. District/school uses funds to support ELA and math goals.

Essential Program Components for Grades 9-12:

1. Use of standards-aligned textbooks and SBE-adopted intervention programs in 9th and 10th ELA, Algebra 1 and remedial mathematics classes.
2. School's master schedule reflects effective use of instructional time and student access to required ELA and math instruction.
3. District provides principals and vice principals with appropriate trainings. (e.g. AB 75)
4. Substantial number of fully credentialed 9th and 10th ELA intervention teachers, Algebra 1, and remedial math teachers and appropriate professional development.
5. District/school implements assessment system including monitoring and reporting of student progress. Assessments based on standards-aligned programs.
6. District/school provides instructional assistance and support to teachers of ELA and math. (e.g. coaches/content experts)
7. Teacher collaboration within departments.
8. District/school provides SBE-approved intervention programs for all students working two or more grade levels behind in ELA and Algebra 1 as assessed on the CAHSEE.
9. District/school uses funds to support ELA and math goals.

Appendix C:
Interview Protocols

Appendix C-1:
School Administrator Interview Protocol for "Growth" Schools

School Administrator Interview Protocol for Growth Schools[1]

Respondent Background

1. I'd like to start by asking you to tell me a bit about your background in schools. How long have you been a principal? What was your prior experience? Prior to becoming a principal how many years did you teach?
2. How many years in total have you worked at this school?

Assessment of success

3. Based our analyses, we have found that your school has made growth over the past few years. Do you see yourself as a school that has been successful in improving student outcomes?

Factors for Growth

4. We are interested in learning what you feel are the crucial elements to your school's growth. We recognize that there are likely multiple factors. But if you had to limit it, what would you list as the three most important factors that contributed to your school's improvement?

   Which of the three factors would you say was the most critical for your school's improvement?

   Can you tell me more about this? How did you see this factor as contributing to your schools' improvement? What specifically about this was important to your success? Are you or your teachers doing something differently now than before?

Challenges

5. We are also interested in learning what you feel are the greatest challenges to your school's growth. We recognize there are likely multiple challenges. But if you had to limit it, what would you list as the three greatest challenges to your school's improvement?

   Which of the three factors you mentioned would you say has been the greatest challenge to your school's improvement? I'm interested in learning more about this challenge. How has this been a challenge? Are you addressing this challenge? How?

6. I'd now like to ask you about factors that may have contributed to your school's improvement over the past few years. On a scale of 0-10, with 0 meaning not at all important and 10 meaning the most important factor,

   How important have each of the following factors been to your growth? (additional probes on each)

   - The use of standards-based instructional materials
   - Increased professional development opportunities for teachers
   - Teacher collaboration
   - The use of data for instructional planning

[1] Note: this is an abbreviated protocol. Additional probes were asked for each question.

7. I'd also like to ask you a little bit about leadership.
   a. Besides yourself, who are the primary leaders in regard to your improvement efforts at your school?
   b. How has this leadership distribution affected your improvement efforts, if at all?
   c. Overall, on the same scale of 0-10, how important has leadership at your school been to your growth?
8. Has there been a high level of turnover at your school in the past few years? How has the level of turnover among your staff affected your improvement efforts, if at all?
9. Over the past few years, has the amount of funding available to your school been a *constraint* towards your improvement efforts, or has it *created opportunities* for improvement?

Influence of II/USP[2]

10. We're interested in learning about II/USP's influence on your growth. On a scale of 0-10, with 0 meaning not at all important and 10 meaning the most important factor, how important has II/USP been to your school's growth? Why? How?
11. How did II/USP affect your most critical factor to growth, if at all?
12. Are there any additional strategies you implemented through II/USP that you feel contributed to your growth? If yes, please describe.
13. I'd like to ask you about the funding you received through II/USP. What, if anything, did this funding allow you to do that you would not have been able to do otherwise?

Sustainability

14. Our analysis shows that your school has grown over the past few years. Do you think you will continue to meet your API growth targets over the next three years? Why / Why not?

    Do you think you will continue to meet your AYP targets over the next three years? Why / Why not?

District supports

15. We're interested in learning about your district's influence on your growth. On a scale of 0-10, with 0 meaning not at all important and 10 meaning the most important factor, how important was your district's support and assistance to your school's growth? Why? How?
16. How did the district affect the most critical factor to your success, if at all?
17. What other things could the district have done to assist your improvement efforts?

[2] The questions in this section were only asked of growth II/USP schools.

Balancing Multiple Accountability Systems

18. How is your school balancing the CA accountability requirements with the federal NCLB accountability requirements?

Wrap-up

19. Are there any aspects of the II/USP program that you feel should have been different to make it more effective? What do you see as the strengths and weakness of the II/USP policy?

Appendix C-2:
Teacher Interview Protocol for “Growth” Schools

Teacher Interview Protocol for Growth Schools[3]

Respondent Background

1. I'd like to start by asking you to tell me a bit about your teaching background. How long have you been teaching? How long have you been teaching at this school? What grade levels have you taught?
2. What is your current role at this school, i.e., what grades/courses do you teach? Do you have other activities or responsibilities in the school?

Assessment of success

3. Based our analyses, we have found that your school has made growth over the past few years. Do you see yourself as a school that has been successful in improving student outcomes?

Factors for Growth

4. We are interested in learning what you feel are the crucial elements to your school's growth. We realize there are likely multiple factors, but if you had to limit it, what would you list as the three most important factors that contributed to your school's improvement?

   Which of the three factors would you say was the most critical for your school's improvement?

   Can you talk to me more about this. What specifically about this was important to your success?

Challenges

5. We are also interested in learning what you feel are the greatest challenges to your school's growth. We recognize there are likely multiple challenges. But if you had to limit it, what would you list as the three greatest challenges to your school's improvement?

   Which of the three factors you mentioned would you say has been the greatest challenge to your school's improvement? I'm interested in learning more about this challenge. How has this been a challenge? Is your school addressing this challenge? How?

6. I'd now like to ask you about factors that may have contributed to your school's improvement over the past few years. On a scale of 0-10, with 0 meaning not at all important and 10 meaning the most important factor, how important have each of the following factors been to your growth? (additional probes on each)

   The use of standards-based instructional materials

   Increased professional development opportunities for teachers

   Leadership at your school

   Teacher collaboration

   The use of data for instructional planning

7. I'd also like to ask you a little bit about leadership.
   a. Who are the primary leaders in regard to your improvement efforts at your school?
   b. How has this leadership distribution affected your improvement efforts, if at all?

[3] Note: this is an abbreviated protocol. Additional probes were asked for each question.

8. Has there been a high level of turnover at your school in the past few years? How has the level of turnover among your staff affected your improvement efforts, if at all?
9. Over the past few years, has the amount of funding available to your school been a *constraint* towards your improvement efforts, or has it *created opportunities* for improvement?

Influence of II/USP[4]

10. We're interested in learning about II/USP's influence on your growth. On a scale of 0-10, with 0 meaning not at all important and 10 meaning the most important factor, how important was II/USP to your school's growth? Why? How?
11. How did II/USP affect your most critical factor to growth, if at all?
12. Are there any additional strategies you implemented through II/USP that you feel contributed to your growth? If yes, please describe.
13. I'd like to ask you about the funding you received through II/USP. What, if anything, did this funding allow you to do that you would not have been able to do otherwise?

Sustainability

14. Our analysis shows that on average your school has grown by X amount over the past few years and has continued to meet the growth targets. Do you think you will continue to meet your API growth targets over the next three years? Why / Why not?

    Do you think you will continue to meet your AYP targets over the next three years? Why / Why not?

District supports

15. We're interested in learning about your district's influence on your growth. On a scale of 0-10, with 0 meaning not at all important and 10 meaning the most important factor, how important was your district's support and assistance to your school's growth? Why? How?
16. How did the district affect the most critical factor to your growth, if at all?
17. What other things could the district have done to assist your improvement efforts?

Wrap-up

18. Are there any aspects of the II/USP program that you feel should have been different to make it more effective? Are there any lessons learned about the program that you would like to share?

[4] The questions in this section were only asked of growth II/USP schools.

Appendix C-3:
School Administrator Interview Protocol
for State Monitored Schools

School Administrator Interview Protocol for State Monitored Schools[5]

Respondent Background

1. I'd like to start by asking you to tell me a bit about your background in schools. How long have you been a principal? What was your prior experience? How many years did you teach?
2. How many years in total have you worked at this school?

Assessment of success

3. Do you see yourself as a school that has made progress in improving student outcomes?

Factors for Progress

4. (If progress made) We are interested in learning what you feel are the crucial elements to your school's progress. We recognize that there are likely multiple factors. But if you had to limit it, what would you list as the three most important factors that contributed to your school's progress?

   Which of the three factors would you say was the most critical for your school's progress?

   Can you talk to me more about this. What specifically about this was important to your progress? Are you doing something differently now as a result, that has led to your progress?

Challenges

5. We are also interested in learning what you feel are the greatest challenges to your school's growth. We recognize there are likely multiple challenges. But if you had to limit it, what would you list as the three greatest challenges to your school's growth?

   Which of the three factors you mentioned would you say has been the greatest challenge to your school's growth? I'm interested in learning more about this challenge. How has this been a challenge? Are you addressing this challenge? How?

[5] Note: this is an abbreviated protocol. Additional probes were asked for each question.

6. I'd now like to ask you about factors that may have contributed to your school's progress over the past few years. On a scale of 0-10, with 0 meaning not at all important and 10 meaning the most important factor, how important have each of the following been to your improvement efforts? (additional probes on each)
    - The use of standards-based instructional materials
    - Increased professional development opportunities for teachers
    - Teacher collaboration
    - The use of data for instructional planning
7. I'd also like to ask you a little bit about leadership.
    a. Besides yourself, who are the primary leaders in regard to your improvement efforts at your school?
    b. How has this leadership distribution affected your improvement efforts, if at all?
8. Has there been a high level of turnover at your school in the past few years? How has the level of turnover among your staff affected your improvement efforts, if at all?
9. Over the past few years, has the amount of funding available to your school been a *constraint* towards your improvement efforts, or has it *created opportunities* for improvement?

Influence of II/USP

10. We're interested in learning about II/USP's influence on your progress. On a scale of 0-10, with 0 meaning not at all important and 10 meaning the most important factor, how important has II/USP been to your school's progress? Why? How?
11. How did II/USP affect the most critical factor to your progress, if at all?
12. Are there any additional strategies you implemented through II/USP that you feel contributed to your progress? If yes, please describe.
13. I'd like to ask you about the funding you received through II/USP. What, if anything, did this funding allow you to do that you would not have been able to do otherwise?

State-Monitoring Process

14. I would like to learn about the team that is assisting your school through the state-monitoring process. What has been the role of the SAIT team?
15. On a scale of 0-10, with 0 as not at all involved and 10 as extremely involved: to what extent were you involved with the audit conducted by the SAIT? In what ways were you involved?
16. Are you familiar with the 9 essential components associated with the state-monitoring process? If yes, what components were identified as areas of improvement for your school? Did these areas align with what you think should be done to improve student learning at your school?
17. What strategies has your school implemented as a result of the SAIT process?
18. Since the initial audit, on average how often has the SAIT team visited your school? What did they do during the visits?
19. How has working with the SAIT been different, or not, from working with the External Evaluator?

20. Overall, do you think the state-monitoring process will help your school to improve student outcomes and meet your growth targets? What do you think will happen to your school if you do **not** meet your growth targets in the next few years?

District supports

21. We're interested in learning about your district's influence on your improvement efforts. On a scale of 0-10, with 0 meaning not at all important and 10 meaning the most important factor, how important was your district's support and assistance to your school's improvement efforts? Why? How?
22. How did the district affect the most critical factor to your progress, if at all?
23. What other things could the district have done to assist your improvement efforts?

Balancing Multiple Accountability Systems

24. How is your school balancing the CA accountability requirements with the federal NCLB accountability requirements?

Wrap-up

25. Are there any aspects of the II/USP program that you feel should have been different to make it more effective? What do you see as the strengths and weakness of the II/USP policy?

Appendix C-4:
Teacher Interview Protocol for State Monitored Schools

Teacher Interview Protocol for State Monitored Schools[6]

Respondent Background

1. I'd like to start by asking you to tell me a bit about your teaching background. How long have you been teaching? How long have you been teaching at this school? What grade levels have you taught?
2. What is your current role at this school, i.e., what grades/courses do you teach? Do you have other activities or responsibilities in the school?

Assessment of success

3. Do you see yourself as a school that has made progress in improving student outcomes?

Factors for progress

4. (If progress made) We are interested in learning what you feel are the crucial elements to your school's progress in improving student outcomes. What would you list as the three most important factors that contributed to your school's progress?

   Which of the three factors would you say was the most critical for your school's progress?

   Can you talk to me more about this factor. What specifically about this was important to your progress?

Challenges

5. We are also interested in learning what you feel are the greatest challenges to your school's growth. We recognize there are likely multiple challenges. But if you had to limit it, what would you list as the three greatest challenges to your school's improvement?

   Which of the three factors you mentioned would you say has been the greatest challenge to your school's improvement? I'm interested in learning more about this challenge. How has this been a challenge? Is your school addressing this challenge? How?

6. I'd now like to ask you about factors that may have contributed to your school's progress in improving student outcomes over the past few years. On a scale of 0-10, with 0 meaning not important at all and 10 meaning the most important factor:

   How important have each of the following been to your improvement efforts? (additional probes on each)

   - The use of standards-based instructional materials
   - Increased professional development opportunities for teachers
   - Leadership at your school
   - Teacher collaboration
   - The use of data for instructional planning

7. I'd also like to ask you a little bit about leadership.
   a. Who are the primary leaders in regard to your improvement efforts at your school?
   b. How has this leadership distribution affected your improvement efforts, if at all?

[6] Note: this is an abbreviated protocol. Additional probes were asked for each question.

8. Has there been a high level of turnover at your school in the past few years? How has the level of turnover among your staff affected your improvement efforts, if at all?
9. Over the past few years, has the amount of funding available to your school been a *constraint* towards your improvement efforts, or has it *created opportunities* for improvement?

Influence of II/USP

10. We're interested in learning about II/USP's influence on your progress. On a scale of 0-10, with 0 meaning not at all important and 10 meaning the most important factor, how important was II/USP to your school's progress?
11. How did II/USP affect the most critical factor to your progress, if at all?
12. Are there any additional strategies you implemented through II/USP that you feel contributed to your progress? If yes, please describe.
13. I'd like to ask you about the funding you received through II/USP. What, if anything, did this funding allow your school to do that you would not have been able to do otherwise?

State-Monitoring Process

14. What was your reaction when you first heard that your school would be state-monitored?
15. I would like to learn about the team that is assisting your school through the state-monitoring process. Have you met the SAIT team? How much contact have you had with the team? What is the nature of the contact? What are the strengths and weaknesses of the SAIT team?
16. What strategies did the SAIT team recommend for your school? Did these areas align with what you think should be done to improve student learning at your school?
17. What strategies has your school implemented as a result of the SAIT process? Since the initial audit, on average, how often has the SAIT team visited your school?What did they do during the visits?
18. How has working with the SAIT been different, or not, from working with the External Evaluator?
19. Have you noticed any changes in your school as a result of being state-monitored?

District supports

20. We're interested in learning about your district's influence on your improvement efforts. On a scale of 0-10, with 0 meaning not at all important and 10 meaning the most important factor, how important was your district's support and assistance to your school's improvement efforts? Why? How?
21. How did the district affect the most critical factor to your progress, if at all?
22. What other things could the district have done to assist your improvement efforts?

Wrap-up

23. Do you think your school will meet its growth targets? What do you think will happen in the next few years if you don't meet them?
24. Are there any aspects of the II/USP program that you feel should have been different to make it more effective? Are there any lessons learned about the program that you would like to share?

Appendix C-5:
District-Level Interview Protocol

District-Level Interview Protocol

Respondent Background, District Background

1. I'd like to start by talking a bit about your background in this district, especially as it relates to school improvement and accountability. How many years have you been involved with this school district? What was your prior experience? What is your role in this district?

Enablers

2. We have chosen several schools in your district to be in our sample for this study. I'd like to first ask you a bit about these schools and what has helped them make progress in improving student learning.

   - **For growth schools:** We chose [schools] since they have [consistently met growth targets/made high average API growth] since beginning II/USP implementation. Do you agree that these schools have had high growth?

     We are interested in learning what are the crucial elements to these schools' success. What would you say have been the three most important factors that have contributed to these schools' improvement?

   - **For comparison schools or state-monitored schools:** We chose [schools] since they [are state-monitored/have not met their growth targets since beginning II/USP implementation.] Do you think these schools have made some progress in improving student outcomes over the past few years?

3. I would like to focus our discussion on one of the three factors you mentioned. Which of the three factors would you say was the most critical for these schools' improvement? I'm interested in learning more about how this factor has been critical to your II/USP schools' progress. How has this factor led to their success? Has the district contributed in any way to this factor? How do these schools compare to other schools in your district with respect to this factor? Is this factor unique to these schools?

Challenges

4. I am also interested in learning what you feel are the greatest challenges to your efforts to improve II/USP and other similar schools in your district. We recognize there are likely multiple challenges. But if you had to limit it, what would you list as the three greatest challenges to improving II/USP schools and other low performing schools in this district?

5. Which of the three factors you mentioned would you say has been the greatest challenge? I'm interested in learning more about this challenge. How has this been a challenge for your schools? Is the district helping to address this challenge? How?

District Role, Policies, and Support for Struggling Schools

6. Overall, how would you describe the role of the district central office with regard to schools that are trying to undertake improvement efforts? Is there a coherent vision or set of goals in place for how to assist these schools?

7. I'd now like to speak with you about the types of policies in place for II/USP and similar schools in your district. Are there specific policies or programs you've developed in your district to support these schools? Could you briefly describe the policy? Why did your district decide to adopt this policy? Do these priorities address schools beyond those that participated in the II/USP program? Which ones?

8. Are there other key supports that the district provides to schools? Can you talk about them briefly.

Influence of II/USP[7]

9. Do you think participation in II/USP has had an impact on schools' improvement? Why or why not?

10. What aspects of II/USP could have been improved to better support schools' improvement efforts?

11. I am interested in the district's role with II/USP schools. To what extent has the district been involved in the planning and implementation phases of II/USP? With monitoring of II/USP schools?

12. To what extent have schools in your district been able to maintain the strategies that were implemented as a result of the II/USP action plan, after II/USP funding ended? Can you give an example?

State-Monitoring Process[8]

13. How was the SAIT chosen? What qualifications did you feel to be important for a SAIT member?

14. What has been the district's role in working with the SAIT(s) assigned to the school(s)?

15. Has the SAIT been effective in supporting your schools' improvement efforts? In what ways?

16. Will the process help to improve the schools in your district? What kinds of on-going support have they provided to schools in your district? In what ways could the process be more effective?

17. What kinds of support or guidance, if any, do you receive at the district level from the SAIT?

18. What did the SAIT suggest as areas for improvement at the district level?

19. Of the schools that are state-monitored, how many are expected to meet their targets this year? What is likely to happen to those schools if they do not meet their targets?

[7] The questions in this section were only asked of districts with II/USP schools.
[8] The questions in this section were only asked of districts with schools that are state monitored and are working with a SAIT.

Balancing Multiple Accountability Systems

20. Are you feeling conflicting pressures from the federal and state accountability systems?

21. How is the district balancing the pressures of accountability from the state and federal governments? Are you focusing more on one than the other?

22. How is the district helping schools balance these pressures?

Wrap-Up

23. Is there anything else you would like to tell me about your experiences with II/USP? What do you think are the strengths and weaknesses of the policy?

Appendix C-6:
School Assistance Intervention Team Member Interview Protocol

School Assistance Intervention Team Member Interview Protocol

Background & qualifications

1. I'd like to first ask you about your background. Could you tell me briefly about your background in education? What experience do you have in teaching and/or coaching?

2. How many people comprise your SAIT?

3. How did you (or your organization) come to be part of a SAIT? Did your organization serve as an External Evaluator at the beginning of the II/USP process?

4. Describe the training you received from the CDE in becoming a SAIT. How was the training structured? Did you find the training valuable?

5. How many II/USP schools are you working with? How did you come to work as a SAIT for these schools?

Role of SAIT

6. In general terms, how would you describe your role as a SAIT?

Audit process

7. I have some specific questions about the work you have done at [school], which is in our study sample. How were you received by the school community when you began your work in this school?

8. What were the school's expectations of your role? The district's expectations?

9. What did you do when you first started working with the school? Who at the school did you work with most? How did your work change over time?

10. What were the major challenges you faced when conducting your audit?

11. How often did you visit to complete the audit? For how long? What did do you do during the visits?

12. Was this process similar to what you did at other schools?

13. Did you find [school]'s responses to the Academic Program Survey to be accurate? If not, what was different about what you witnessed?

14. How central were the 9 Essential Program Components (EPC) in guiding your work at [school]? Which components were fully or substantially in place? Which components were partially or minimally in place?

15. Do you think the EPCs adequately measure or reflect what needs to be in place at this school for it to improve?

16. To what extent were the EPCs useful in guiding your work? Why? To what extent were they limiting? Why?

17. Was this the same at other schools you worked with as well?

Recommendations

18. What primary strategies did you recommend this school undertake to improve?

19. Who provided input to these recommendations?

20. To what extent were these corrective actions or strategies based on the EPCs?

21. What progress has the school made so far towards implementing the corrective actions and strategies you recommended?

22. Does the school have the resources necessary to implement the reforms outlined in the audit? If not, how will it obtain these resources?

23. Are you familiar with [school]'s Action Plan from the II/USP planning year? How do the corrective actions recommended in your audit differ from the recommendations in the school's Action Plan, developed with the External Evaluator? Is the Action Plan still being used at all at this school?

Follow up

24. Has a clear plan been agreed upon for your role at [school] as they implement the changes you suggested? If yes, what is the plan?

25. Have you been involved in implementing the corrective actions or providing other supports? How?

26. Have you monitored the school's progress? How?

27. How often do you visit the school? How long is each visit? What is the nature of these visits? Is this typical for the other schools you have been working with?

28. Do you feel that all of the stakeholder groups are invested or committed to following the corrective actions outlined in your audit?

29. Based on what you've observed so far, do you think this school has made progress as a result of participation in the state-monitoring process? In what way? How do you know?

30. Do you anticipate that the school will meet its growth targets over the next few years? Why or why not?

31. How does their progress compare to other schools with which you have been working?

32. How concerned do teachers and administrators appear to be about further sanctions if academic performance still does not improve?

Overlap with External Evaluator

33. Have you ever served as both an External Evaluator and SAIT. How has your role as a SAIT differed from your role as an External Evaluator?

District relationship

34. How was the district [of the case study school] involved in the audit and recommendation process? What responsibilities do districts in general have in ensuring the implementation of corrective actions?

35. How often does the School/District Liaison Team meet? What occurs at these meetings?

36. How much did you focus on reform at the district level when conducting your work at [school]?

Wrap-Up

37. How has the federal NCLB program affected, if at all, your work in schools?

38. Do you think the state-monitoring process is targeting an appropriate population of schools?

39. Overall, what do you think are the strengths and weaknesses of the state-monitoring process, and II/USP in general? What suggestions would you make to future policy makers?

Appendix C-7:
State-Level Respondent Interview Protocol

State-Level Respondent Interview Protocol

State-level Accountability Goals

1. We would like to first focus on the state's accountability system, including II/USP and High Priority Schools Grant Program (HPSGP). What do you see as the state's goals for accountability? For the improvement of low-performing schools?

2. How does the current system work to fulfill those goals?

3. Could you tell me about the transition from II/USP to HP? What were the major changes made in this transition? What were the driving factors behind these changes? How do the goals of the two programs differ, if at all?

4. How do II/USP and HP fit in with other programs focused on the improvement of low performing schools (like CSRD, Reading First, etc.)? How does this play out for schools that participate in more than one of these programs?

5. What are the strengths and weaknesses of the state accountability system (including II/USP and HP)? Do you think the system is helping to improve student learning?

6. What challenges has the state faced in implementing this system?

State-Monitoring Process

7. What do you see as the goals for the state-monitoring process?

8. How central are the 9 Essential Components to this process? How were the 9 Essential Program Components identified?

9. What is the process for approving state-monitoring teams (SAITs)?

10. Do you think this process has worked to help these schools improve?

11. We know there has been discussion about which schools should enter this process – what is your opinion of this? Why?

12. The Education Code specifies that for HPSGP schools which do not meet growth targets after three years and failed to make significant growth, the district must enter into a contract with a school assistance and intervention team. However, our understanding is that the plan for HP schools that do not make significant growth is not yet clear. Is that correct? If no, what is the plan? If yes, are there discussions about what will happen to these schools?

State vs. Federal Accountability Goals

13. We'd now like to ask you about the state system, in relation to the federal NCLB legislation. What is your view of the alignment, or lack of alignment between the state accountability system and the federal NCLB system?

14. How does the state perceive tension between the state accountability system and NCLB at the district and school level? To what extent does this tension play out at the state level?

15. How is the state dealing with this situation (district/local and state levels)? Is there a long-term plan in mind for integrating the systems? How do you see this being resolved over time?

16. Do you think districts and schools are focused more on the state targets (API) or the federal targets (AYP)? Why?

17. We understand there will be an increase in AYP targets next year. This will raise the bar significantly for schools. What do you think the implications are for this?

The Future of State Accountability System

18. What do you see as the long-term state role in regards to underperforming schools? Will the state continue to invest resources in the state accountability system, for example in HPSGP? Will resources be targeted differently? Do you see a state intervention or accountability program continuing in the long term?

19. How has the current budget crisis affected plans for accountability and school improvement efforts?

Districts

State Accountability

20. How are the current provisions of the state's accountability program designed to enhance districts' ability to assist schools?

21. Do you think there should be a system of rewards and sanctions at the district level within the state system?

Federal Accountability

22. We understand many districts have been identified for Program Improvement under NCLB. What types of assistance and support is the state providing to these districts under NCLB?