

Office of the Independent Monitor

Modified Consent Decree

333 So. Beaudry Avenue, 16th Floor
Los Angeles, California 90017
Tel: (213) 241-6036
Fax: (213) 241-7551

DR. CARL A. COHN
Independent Monitor

JAY R. ALLEMAN
Administrative Coordinator

ARUN K. RAMANATHAN
Research Director

JAIME E. HERNANDEZ
Outreach Coordinator

Memo

To: Robert Myers, Catherine Blakemore, Maria Ott

From: Carl A. Cohn

Re: Outcome No. 2: Performance in the Statewide Assessment Program

Date: May 3, 2004

CC: Roy Romer, Donnalyn Jaque-Anton, Allyn Kreps, Brigitte Ammons,
Thomas Hehir, Jay Alleman

Section 6, Subsection A (37) of the Modified Consent Decree states that “The Independent Monitor shall develop an outcome measure to require that the performance of students with disabilities increases on the statewide assessment program and that the disparity with general education performance on the statewide assessment program is reduced. The outcome measure shall be based on key performance indices.”

This section requires the Independent Monitor to identify the disparity or “gap” between the performance of students with disabilities and general education students on the key performance indices of the statewide assessment program. Once the performance gap has been established, the Independent Monitor must develop an outcome to ensure that this gap decreases and that the performance of students with disabilities increases.

Background

The state-wide assessment system is referred to as the California Standardized Testing and Reporting (STAR) program. During the 2002-03 school year, most students in Los Angeles Unified School District (LAUSD) took the California Standards Tests (CST) and the California Achievement Test, Sixth Edition Survey (CAT6). Students with the most significant cognitive disabilities, who were unable to take the CST or the CAT6, took the California Alternative Performance Assessment (CAPA). While the CST and CAT6 have a number of different components, nearly all students take the English Language Arts and Mathematics sections. Because of their high participation rates and broad grade distribution, these sections would constitute the key performance indices that could be used to develop an outcome measure. These are also the sections that form the basis of the No Child Left Behind (NCLB) standards.

Over the past two years, the results of the CST have been given a greater weight in the determination of the Academic Performance Index (API) than the results of the CAT6. For the 2003-04 school year, the CST comprised 80% of the weight for the grade 2-8 API calculations and 73% of the weight for the grade 9-11 API calculations. As the CST has assumed greater weight, it should be considered the key performance index. Since it is possible that the CAT6 may be phased out before the termination of the Modified Consent Decree, the results of the CST should be used to develop an outcome measure. However, since the CAT6 is a nationally norm referenced test, its results can provide a valuable reference for LAUSD performance vs. the performance of students nationally. As such, we requested an analysis of both the CST and the CAT6 from Dr. Peter Goldschmidt of the National Center for Research on Evaluation, Standards and Student Testing (See Appendix A). The CST results are presented in the body of his report and the CAT6 in the appendices. In this memo, I will primarily refer to the CST results. These results will be reported as scaled scores rather than using the proficiency categories (% advanced, proficient, basic, below basic, far below basic) derived from the requirements of No Child Left Behind. This is because proficiency categories cannot be used to establish an overall performance gap. However, we have attached the LAUSD STAR summary reports containing the % proficiency breakdown for your information (See Appendix B).

Performance Data

Dr. Goldschmidt found that the distribution of the students with disabilities who took the mathematics and English language arts (ELA) sections of the CST and CAT6 was similar to the distribution of all students with disabilities who took each test. He found that the two groups had similar ethnic, grade, language proficiency and gender distributions. He also found that the disability distribution of special education students who took the CAPA was similar to that of special education students who took the CST and the CAT6, with the exception of the MR subgroup that “was overwhelmingly assigned to the CAPA”. Based on these findings, Goldschmidt states that “overall, the students with valid CST data are representative of the District as a whole.”

Goldschmidt’s methodology permitted him to divide the variation in achievement between student and school components. This allowed him to determine how much of the variation or difference in reading scores is associated with a student level effect such as English Language Learner (ELL) status and how much is associated with placement in a particular school – an elementary school in Local District A vs. one in Local District G. Goldschmidt found that roughly 73% of the variation in math scores and 79% in English language arts are attributable to innate student effects within schools. The remainder can be attributed to differences between schools. This means that a comparably small percentage of the variation in student scores is attributable to differences between school level factors such as enrollment. In a practical sense, this means that addressing the overall achievement gap for students with disabilities by addressing differences between schools would have less effect than addressing differences between classrooms within a school. It does not, however, mean that the achievement gap is fixed in stone because of a student’s demographic statistics. Nor does it diminish the likelihood that any such gap could be addressed through instructional, curricular or other classroom or student-level improvements.

While Goldschmidt found little overall variation in the achievement gap for students with disabilities as a whole. for three subgroups, specific learning disability (SLD), speech and

language impaired (SLI), and Deaf-Blind/Multiple Disabilities, the achievement gap did vary to a small extent between schools. As Goldschmidt notes, for these categories, “some schools are doing better (smaller gap) and some schools are doing worse (larger gap). In fact, the results indicate the better schools would have an SLD/Non-SLD achievement gap that is 23 [for math] and 10 points [for ELA] smaller than a school doing poorly with SLD students.” Goldschmidt explores this issue in greater detail on page 16.

Goldschmidt creates three models to analyze the achievement gap. The parameter estimates for these models are presented in Table 9 on page 11. Model 1, identified on the chart as estimate 1, presents the achievement gap without considering the effect of various demographic variables such as ethnicity. For this model, the achievement gap in English Language Arts is -42.69. This means that students with disabilities as a group performed 42.69 points below non-disabled students. That is the “effect of being designated special education” for performance in English language arts. Similarly, in math, the effect of being designated special education is -55.38. However, when one factors in the effects of the demographic variables presented in Model 2 (Estimate 2), the achievement gap “is reduced by approximately 20%” to -32.41 in English language arts and -45.54 in math. This is the baseline achievement gap.¹

Model 3 (estimate 3) “breaks out” the performance gap by disability category and demographic variables. For example, in English language arts, a student with mental retardation performed on average 56.13 points lower than a non-mentally retarded student, controlling for demographic characteristics related to achievement. Similarly, a student with a specific learning disability (SLD) performed 36.96 points lower than a non-SLD student. Interestingly, SLD exhibits the largest gap of any category in English language arts with the exception of MR, an unexpected result since SLD is a high incidence disability with typically less impact on performance than moderate to severe disabilities. As would be expected, speech and language impairment (SLI) has the least effect on performance at -8.36 for English language arts and -8.53 for math.

A second interesting feature of Model 3 is that the effect of each characteristic is additive. For example, the effect of being African American is -23.16 points. I have already noted that the effect of being SLD is -36.96 points. Thus, the effect of being African American and labeled SLD on performance on the English language arts section of the CST would be the sum of the two effects or -60.22. This means that an African American identified as SLD performed 60.22 points lower than a non-African-American, non-SLD student. If one adds the -21.67 point effect of Title I status to the -36.96 effect of being SLD, the performance gap is 58.63 points.

Table 11 on page 16 considers the effect of school context. As noted earlier there is variation between schools in the achievement gap for students in the categories: SLD, ED, MD_DB, and SLI. This table examines the effects of several demographic variables on the *school mean* as opposed to the *student mean*, after accounting for the effects of individual student characteristics such as disability and ethnicity. Interestingly, all of the demographic characteristics with the exception of school size have a positive effect on school mean. Based on the results of earlier analyses, it would be expected that a larger % of GATE students and % of White (non-minority) students would increase a school’s score and the size of a school have a negative relationship with the overall score (larger schools produce lower scores). In addition, this table reveals that

¹ Table 10 on page 13 presents these results in terms of effect sizes. Thus, in English language arts, students with disabilities performed .49 standard deviation below non-disabled student. For math, the gap was .91 standard deviation. The standard deviation for the sample is 65.9 for English language arts and 49.9 for math.

both % special education and % ELL students have a positive effect on a school's overall score, once we control for the effects of disability and ELL status. When the effects are examined by specific disability category, some of these results are reversed. For example, student with SLD, increasing the % of ELL students in a school increases the math performance gap by 3.49 points.

Implications

Dr. Goldschmidt's states that "the existing achievement gaps between special education and non-special education students are substantively large and careful analysis of ensuing years' data needs to be undertaken in order to properly monitor progress toward reducing those gaps." This finding has considerable implications for both general and special education.

Goldschmidt also makes a series of important points regarding the development of any performance outcome for the District. He states that "it would be inappropriate to monitor the special education achievement gap" using the overall gap between special and general education or to monitor categories such as low and high incidence disabilities. This is because of the "substantial difference in the performance gaps among specific disability categories." For example, the gap between in English language arts for the two high incidence categories, SLD and SLI are 36.96 and 8.36 respectively. These are considerably different gaps and combining them would only result in masking those differences. Further, a high incidence gap would make little practical sense, as there are likely to be considerable differences in the manner in which the District would attempt to improve the performance of these different groups of students. The same holds for the moderate to low incidence categories. This does not, however, preclude the possibility of other groupings for the purpose of establishing a performance outcome/s.

In terms of monitoring the gap over time, Goldschmidt states that a growth curve model would be preferable. He also notes that "there exist external estimates of acceptable differences in performance due to actual cognitive differences and not due to school related factors." By this, he is referring the fact that disabilities such as mental retardation and traumatic brain injury can impact students' cognitive ability. Even though many students with cognitive disabilities take the CAPA, for some disability categories, it is likely that a percentage of the performance gap for those that take the CST is associated with the impact of the disability and a percentage with instructional and curricular issues. The District has little control over the former but considerable control over the latter. While the provision of accommodations and modifications can enhance the performance of students with disabilities, it is unlikely that they can entirely overcome the impact of a cognitive disability. The parties should carefully consider the impact of disability as they consider the establishment of a performance outcome.

The development of an outcome to address the achievement gap between general and special education is an important step. It reflects the current focus in education policy and practice on improving the academic performance of all students, especially those students who have traditionally been "left behind". Special education has often been overly focused on the process of education rather than its results. As these statistics reveal, the results for students with disabilities in LAUSD both overall and in comparison to their non-disabled students are poor. This is especially the case when one factors in other demographic characteristics such as ethnicity and Title I status. It is my hope that the parties will closely review and consider these results, and if necessary access the expertise to understand and analyze them, prior to engaging in the negotiations over an outcome.

Analysis of Differences in Student Performance 2002-2003.

By
Pete Goldschmidt, Ph.D.

Two issues need to be considered when addressing the performance gaps among students: one, eliminating as many confounding characteristics associated with achievement (Campbell and Stanley, 1963; Pedhazur, 1982); and two, accounting for the natural data structure of students attending schools (Burstein, 1980; Raudenbush and Bryk, 2002). The former issue can be partially addressed by including student characteristics as covariates (Pedhazur, 1982); while latter issue is addressed by using a hierarchical modeling framework (Raudenbush and Bryk, 2002). Within the underlying question of the existing achievement gaps between special education students and no-special education students, this analysis focuses on three related issues:

- one, is the sample of special education students who have available CST mathematics and English Language Arts (ELA) information representative of special education students in the district in general;
- two, are special education students who take the CAPA similar to special education students who take the CST; and
- three what are the achievement gaps for each special education category.

In order to examine these questions the remainder of this analysis presents the data, methods and results.

Data

The data for this analysis were supplied by LAUSD and consist of the standard data elements produced by the California Department of Education for each school district as part of the Standardized testing and Reporting (STAR) system. This dataset includes student background as well as student outcome measures. The outcome measures of interest are the California Standards Test and California Achievement Test v.6 (CAT6) reading and Mathematics results. We focus on the CST because it is more closely aligned with specific California standards. We focus on Mathematics and English Language Arts (ELA) because these

subject areas are tested across all grades¹. To the extent that progress in closing achievement gaps will be monitored over time, the same subject matter needs to be available to examine achievement growth. Other subjects apply to either primary grades or secondary grades, but not both. Also, Mathematics and ELA are the current basis for No Child Left Behind (NCLB) standards². A supplemental dataset, the California Alternative Performance Assessment (CAPA) was also provided by LAUSD. These data include student background and outcome measures for special education students who did not participate in the STAR tests. Table 1 displays the distribution of special education students by category for three samples. The first dataset, STAR data all, represents the entire STAR dataset as received from LAUSD for students who had a specific disability code. The second dataset represented in Table 1 is for students who had a specific disability code and had both Mathematics and ELA CST scores. The third dataset summarized in

Table 1: Distribution of disability categories within each sample

Disability Code	STAR Data, all		STAR Data - w/ CST		CAPA Data	
	Frequency - N	Percent	Frequency - N	Percent	Frequency - N	Percent
None	33	0.05	17	0.03	3	0.06
MR	2,998	4.53	492	1.01	2,233	44.76
hh-deaf	1,317	1.99	961	1.97	178	3.57
sli	4,381	6.61	4,165	8.55	8	0.16
vi	263	0.40	178	0.37	52	1.04
ed	3,468	5.23	2,171	4.46	120	2.41
oi-tbi	822	1.24	544	1.12	153	3.07
ohi	3,077	4.64	2,471	5.07	88	1.76
sld	46,676	70.45	36,635	75.20	277	5.55
db-md	1,264	1.91	139	0.29	990	19.84
aut	1,952	2.95	946	1.94	813	16.30
U					74	1.48
Total (1)	66,251	100	48,719	100	4,989	100

Notes:

(1) The percentage of Special Education students in the STAR all sample, and the STAR with CST sample is 11.5 and 9.7 respectively.

Table 1 is for students in the CAPA file provided by LAUSD. The proportion of special education students with valid CST test scores is slightly less than the proportion of special

¹ Results using CAT6 Mathematics and Reading are reported in Annex 2.

² Science will be included

education students in the district, in general. However, the majority of the difference is accounted for by the fact that about 10% of special education student took the CAPA instead of the CST. Comparing the distribution of students by specific disability code reveals that students with valid CST scores are distributed in approximately the same proportions as in the district in general. One exception to this is that the MR subgroup is significantly under-represented. This is consistent with the CAPA data, as almost 45% of the CAPA examinees are from the MR subgroup. Overall, the sample used to examine performance gaps is sufficiently representative of the LAUSD special edition disability category distribution to utilize it for the analysis.

Of course, we would consider any parameter estimates pertaining to the performance gaps as lower bounds, as it is likely that the most severely affected students took the CAPA instead of the CST. This issue will be addressed further below. Table 2 continues the comparison among the three data samples.

Table 2: Distribution of Ethnicity categories within each sample of Special Education students

Ethnicity	STAR Data, all		STAR Data - w/ CST		CAPA Data	
	Frequency - N	Percent	Frequency - N	Percent	Frequency - N	Percent
White	7,313	11.04	5,345	10.97	590	11.83
aa	12,427	18.76	8,275	16.99	876	17.56
ai	268	0.40	174	0.36	12	0.24
asian	1,049	1.58	803	1.65	125	2.51
pacisl	754	1.14	518	1.06	150	3.01
hisp	43,970	66.37	33,306	68.36	3,180	63.74
unknown	470	0.71	298	0.61	56	1.12
Total	66,251	100	48,719	100	4,989	100

Table 2 displays the distribution of ethnic/race categories among special education students. The focus here is not whether any particular ethnic/race group is over represented in special education, rather whether special education students with valid CST scores are representative of the ethnic/race distribution of special education students in the district as a whole. The STAR CST distribution is representative, while the CAPA examines tend to slightly over-represent Asians and Pacific Islanders.

Table 3 displays the distribution of gender for Special Education students. The

Table 3: Distribution of gender within each sample among Special Education students

Gender	STAR Data, all		STAR Data - w/ CST		CAPA Data	
	Frequency - N	Percent	Frequency - N	Percent	Frequency - N	Percent
Female	22,877	34.54	16,808	34.50	1,876	37.60
Male	43,352	65.46	31,911	65.50	3,108	62.30
missing	22					
Total	66,229	100	48,719	100	4,989	100

results in Table 3 indicate that gender appears consistent among the three samples.

Table 4 presents the distribution of language status among special education students. As with gender, the results appear to be consistent across the three samples.

Table 4: Distribution of language status categories within each sample Among Special Education students

Language Status	STAR Data, all		STAR Data - w/ CST		CAPA Data	
	Frequency - N	Percent	Frequency - N	Percent	Frequency - N	Percent
EO/IFEP*	31,173	47.05	21,759	44.66	2,211	44.32
ELL*	32,623	49.24	24,996	51.31	2,736	54.84
Redesignated*	2,455	3.71	1,964	4.03	42	0.84
Total	66,251	100	48,719	100	4,989	100

* Students who are designated EO/IFEP = English Only/Initially English Proficient; ELL = English Language Learner; Redesignated = students who were initially ELL but have been redesignated as fully English Proficient.

Finally, Table 5 presents the distribution of grade among special education students.

Table 5: Distribution of grade within each sample among Special Education students

Grade	STAR Data, all		STAR Data - w/ CST		CAPA Data	
	Frequency - N	Percent	Frequency - N	Percent	Frequency - N	Percent
2	5,161	7.79	4,310	8.85	508	10.18
3	6,602	9.97	5,657	11.61	491	9.84
4	7,212	10.89	6,094	12.51	541	10.84
5	6,829	10.31	5,814	11.93	528	10.58
6	8,073	12.19	6,832	14.02	602	12.07
7	8,364	12.62	6,328	12.99	573	11.49
8	6,771	10.22	5,281	10.84	478	9.58
9	7,929	11.97	4,747	9.74	459	9.20
10	5,484	8.28	2,234	4.59	419	8.40
11	3,826	5.78	1,422	2.92	390	7.82
Total	66,251	100	48,719	100	4,989	100

It is not readily apparent from examining the numbers in Table 5 that students with CST scores are representative of LAUSD special education students, in general. This result is more apparent

when examining Figure 1, which displays the proportion representation in each grade. The percentages vary slightly among the samples, but the overall patterns are similar.

Figure 1: Distrubtion of Grade by Sample

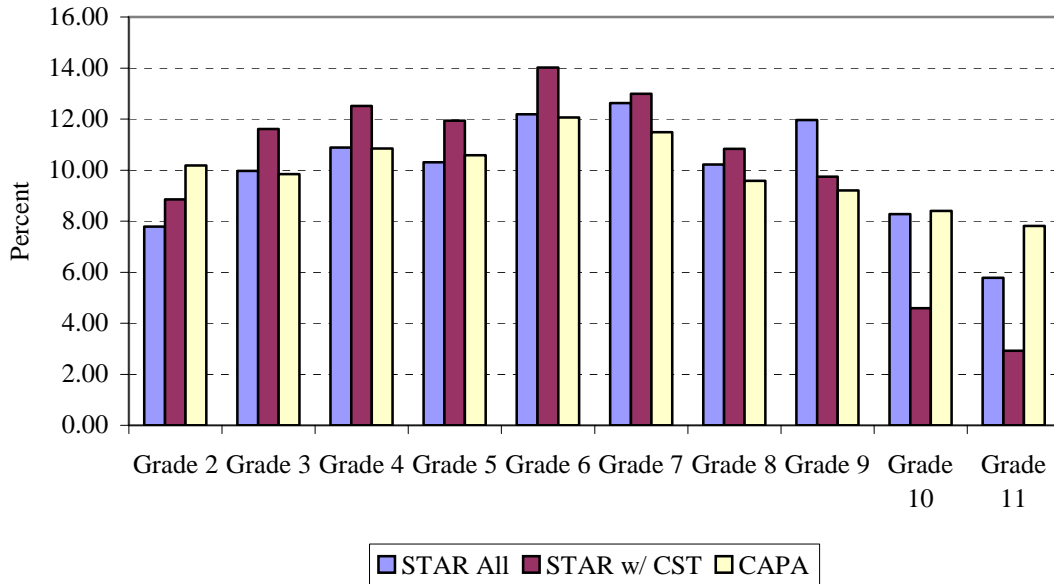


Table 6 examines the bivariate distribution of ethnicity/race and specific disability code for both the STAR and the CAPA dataset. This is primarily presented to examine who is taking the CAPA, as compared to who exists in the district, overall. Again, while there appears to be disproportionate proportions of various ethnic/race categories, the purpose of Table 6 is to examine whether these same proportions exist in the CAPA data. Given the small sample sizes within each cell in the CAPA file, the figures are consistent with the district STAR data source.

Table 6: Disability Category by Ethnicity/Race status (in percent)

STAR							
Disability Category	African American		American Indian		Pacific Islander		Unknown
	White	American	Indian	Asian	Islander	Hispanic	
None	9.41	11.71	0.28	4.20	2.65	71.58	0.18
mr	7.71	16.88	0.33	2.17	2.27	70.51	0.13
hh-deaf	7.97	9.72	0.23	3.42	3.34	75.17	0.15
sli	12.83	11.32	0.32	4.15	1.16	70.01	0.21
vi	13.69	15.21		5.32	3.42	61.98	0.38
ed	18.83	36.45	1.04	0.98	0.63	37.89	4.18
oi-tbi	13.87	10.34	0.36	3.04	2.07	70.07	0.24
ohi	28.14	18.62	0.78	2.14	1.30	48.33	0.68
sld	8.58	18.85	0.35	1.00	0.84	69.88	0.51
db-md	14.56	15.19	0.55	2.61	2.37	64.00	0.71
aut	28.38	17.01	0.46	6.05	4.10	42.01	2.00
Total	9.60	12.52	0.29	3.90	2.47	70.99	0.24

CAPA							
Disability Category	African American		American Indian		Pacific Islander		Unknown
	White	American	Indian	Asian	Islander	Hispanic	
None	33.33					66.67	
mr	7.75	16.03	0.13	2.10	2.55	70.98	0.45
hh-deaf	5.62	17.98		0.56	4.49	71.35	
sli	25.00	25.00			12.50	37.50	
vi	17.31	11.54		3.85	1.92	65.38	
ed	16.67	31.67	1.67	2.50	0.83	45.83	0.83
oi-tbi	11.76	16.34		1.96	4.58	61.44	3.92
ohi	20.45	17.05		3.41	4.55	51.14	3.41
sld	10.11	18.41	1.44	1.81	1.44	62.82	3.97
db-md	12.73	12.93		2.02	2.12	68.99	1.21
aut	21.65	25.46	0.37	4.55	5.04	42.44	0.49
U	12.16	18.92		5.41	6.76	44.59	12.16
Total	11.83	17.56	0.24	2.51	3.01	63.74	1.12

Table 7 presents the distribution of test modifications by specific disability category. These test modifications represent a set of standard modifications as coded by the district. As expected, the vast majority of students in LAUSD do not receive any modifications. What is interesting to note, however, is that between 25% and 93% of special education students taking the CST receive no valid test modification.

The descriptive results indicate that, overall the students with valid CST data are representative of the district as a whole. The caveat to consider are that the MR subgroup is overwhelming assigned to the CAPA. This may be due to the specific nature of this category, but also highlights the fact that there is insufficient information available

Table 7: Distribution of Test Modifications by Disability Category

Disability Category	N	None	If received any test modification, type of Modification received, in Percent					
			IEP	Time	Reading/ELA	Calculator	Dictionary	Other
None	452,578	99.65	0.00	44.96	19.02	3.49	26.76	5.77
Not indicated	716	0.00	100.00	0.00	0.00	0.00	0.00	0.00
mr	492	27.85	30.70	16.90	29.58	8.73	2.82	11.27
hh-deaf	961	54.32	51.25	13.67	23.92	2.51	0.23	8.43
sli	4,165	93.01	40.55	29.55	21.65	2.41	1.72	4.12
vi	178	32.58	35.00	50.00	6.67	2.50	0.00	5.83
ed	2,171	49.42	49.73	26.87	12.02	5.65	2.73	3.01
oi-tbi	544	55.70	40.66	27.80	16.60	5.39	1.24	8.30
ohi	2,471	57.99	43.55	29.67	13.29	7.61	2.02	3.85
slid	36,635	46.07	41.32	28.78	16.13	8.53	2.24	3.00
db-md	139	39.57	40.48	26.19	11.90	10.71	1.19	9.52
aut	946	41.33	32.43	28.29	21.08	7.39	2.70	8.11
Total	501,996	94.77	39.91	28.78	15.77	7.18	4.81	3.56

Note: Data on Braille was missing, hence, some of the students categorized as VI and DB_MD listed as receiving no modification, may have received Braille modification.

with which to discern whether the most severely effected students in any special education category are the ones taking the CAPA (although this may be exactly the role of the CAPA). It is beyond the scope of this study to examine relative performance between the CAPA and the CST. However, the focus of this analysis is to identify baseline performance gaps among specific disability categories and between those and non-special education students. Hence, the next section briefly describes the methodology used to estimate the performance gaps.

Analysis Methodology

As noted above, the natural data structure of students requires that the intra-cluster correlations be taken into account (Raudenbush and Bryk, 2002). Ignoring the data structure leads to under estimation of the standard errors associated with the parameter estimates and inflates type I error rates³. Hence, this analysis will use a two level hierarchical linear model (HLM), with students at level 1 and schools at level 2. By using a two level model, we are able to divide the variation in achievement into between student and between school components. It is particularly important to do so because data containing multiple levels of aggregation can lead to errors in interpretation of results, when these multiple levels are ignored (Aitkin and Longford,

³ The type I error rate is the rate that the null hypothesis is falsely rejected. This is set a-priori, and is referred to as alpha (α). Testing hypotheses entails comparing whether test a statistic (for example, the parameter estimate for the achievement gap between girls and boys) is far enough from 0, to say that it does not have that value due to chance variation. The null hypothesis is rejected when the probability of the test statistic occurring is less than the α value. Dividing the test statistic by the standard error of the test statistic and looking up that value on a normal curve table obtains the probability. If data are clustered (e.g. students within a classroom) and their outcomes are correlated, then this causes the standard error to be too small because the N used to calculate the standard is assumed to represent N independent observations, but in reality represents fewer observations if the observations are correlated. An intra-cluster correlation of (ICC) .05 can double the actual α value – that is increase the likelihood of making a type I error.

1986; Burstein, 1980). For example, ELL (English Language Learner) status, measured at the student level, represents a measure of a student's English language ability, while the aggregated mean student ELL at a school, represents the school or community language environment and the potential effect of encountering, practicing, or reinforcing English or English conventions (Burstein, 1980). These are related, yet, distinct issues. The former variable can be used to examine whether an individual student is affected by not speaking English fluently. The latter addresses whether an ELL student will benefit from peers who speak predominantly English, or predominantly their native language. Ignoring the nested nature of the data and simply analyzing outcomes aggregated to the school level upwardly biases results of student level predictors, because within school student level variation is lost upon aggregation (Freedman, Pisani, and Purves, 1978). Not only are student effect biased, but it becomes unclear whether the estimated effect is due to a group effect or whether the aggregated variable proxies for an un-represented student effect (Burstein, 1980). That is, if the effect of ELL on ELA scores were only measured at the school level, then this would technically address the question of whether peers speaking a native language (more vs. less) effects mean student outcomes. Since the individual effect of ELL status is not measured, it is confounded with the group effect.

For this analysis, the focus is on what the performance gap is for special education students and whether the specific disability categories' effects vary among schools. The HLM methodology presented in Annex 1 was used to generate the results presented in Tables 8 through 10.

Results

Table 8 presents the results of the various models used to identify performance gaps among students. The results displayed under model 0 are the baseline or random

Table 8: Random effects and variance components

Variance Components	Mathematics			ELA		
	0	2	3	0	2	3
Student	3400.383	2589.278	2577.000	2064.099	1409.914	1402.000
School	1288.221	398.056	399.000	536.151	174.655	171.000
SLI						50.8
ED			452.6			180.7
SLD			544.7			103.8
DB_MD			617.5			
	Percent			Percent		
	Attributable	Percent accounted for		Attributable	Percent accounted for	
Student	72.5	23.9	24.2	79.4	31.7	32.1
School	27.5	69.1	69.0	20.6	67.4	68.1

coefficient results. These results indicate that the majority of the variation in CST scores is attributable to students, within schools. In fact 72.5% and 79.4% of the variation in Mathematics and ELA scores are attributable to students, respectively. The variance estimates for students can be used to compare student performance within a school. For example, in mathematics, a student one standard deviation above average would be expected to score about 116 points⁴ higher than a student one standard deviation below average, in the same school. Conversely, about 20.6% to 27.5% of the variation in outcomes is attributable to schools. This means that differences among schools (e.g. school size) can only account for 20%-28% of the variation in achievement scores. Roughly, this means that a hypothetical 10 point achievement gap between girls and boys, in Mathematics, could at best be reduced by 2.75 points by addressing differences *between* schools.

Each of the specific disability categories was tested for randomly varying slopes, but only the four indicated in Table 8 varied randomly. In other words, except for SLI, ED, SLD, and DB_MD⁵, the achievement gaps estimated for each specific disability category does not vary among schools. That is, the average district-wide gaps between special education students (except the specific groups noted above) and non-special education students does not vary among schools – differences between schools (e.g. enrolment) will not account for any of the gap sizes. On the other hand randomly varying slopes (or gaps) for SLD, for example, indicate that some schools are doing better (i.e. smaller gap) and some are doing worse (i.e. larger gap). In fact, these results indicate that better schools would have an SLD-Non-SLD achievement gap

⁴ The student level variance is 3400. The standard deviation = the square root of the variance = 58. 58 X 2= 116.

⁵ Not all disability categories varied significantly in both Mathematics and ELA.

that is 23 and 10 points smaller than a school doing poorly with SLD students, in Mathematics and ELA respectively. This between-school variation is discussed further below.

First student level effects are examined. The results displayed in Table 9 present the parameter estimates derived from three models, each of which build upon the previous model. Separate models are tested for Mathematics and ELA. Model 1 simply categorizes students as being designated as special education or not, but does not account for other student characteristics and potentially confounds the effects of special education status with other student background factors that might affect achievement. Model 2 adds student concomitant variables. Model 3 identifies each specific category separately and includes student concomitant variables as well.

The result from Model 1 indicate that for both Mathematics and ELA, special education students are expected to score approximately 55.4 and 42.7 points lower than non-special education students. This result is not unexpected. Model 2 accounting for differences among students in other concomitant characteristics demonstrates the effect of confounding. The special education parameter estimate (achievement gap) is reduced by approximately 20% once student background characteristics are included. This indicates that comparing unadjusted means over-estimates the gap in special education student achievement by about 20%.

The results of model 3 indicate that there are both substantively large and statistically significant performance gaps between special education and non-special education students. Model 3 further indicates that these performance gaps vary substantially by specific disability category. This implies that grouping special education categories into high and low incidence is likely not appropriate and would incorrectly imply that students within the high or low incidence categories demonstrate similar achievement patterns. Again, the student covariates conform to expectations. Girls perform less well than boys in Mathematics, but better in ELA.

Table 9: HLM parameter estimates for student characteristics

CST Model	Mathematics			ELA		
	1 Estimate	2 Estimate	3 Estimate	1 Estimate	2 Estimate	3 Estimate
Mean Scale Score	329.63 *	322.39 *	322.40 *	323.38 *	320.11 *	320.13 *
Grade		-7.45 *	-7.27 *		-2.40 *	-2.32 *
Effect of Being designated						
Special Education (Any)	-55.38 *	-45.54 *		-42.69 *	-32.41 *	
GATE		55.42 *	55.17 *		47.63 *	47.48 *
MR			-80.28 *			-56.13 *
HH_DEAF			-20.56 *			-21.60 *
SLI			-8.53 *			-8.36 *
VI			-24.87 *			-20.01 *
ED			-46.49 *			-26.88 *
OI_TBI			-37.53 *			-20.74 *
OHI			-38.22 *			-26.02 *
SLD			-52.41 *			-36.96 *
DB_MD			-58.89 *			-31.85 *
AUT			-45.79 *			-27.25 *
Demographic Effects (compared to Whites)						
Girls (vs. Boys)		-3.27 *	-3.19 *		7.23 *	7.30 *
African American		-30.84 *	-30.55 *		-23.51 *	-23.16 *
American Indian		-15.51 *	-15.36 *		-12.47 *	-12.35 *
Asian		21.25 *	21.10 *		4.03 *	4.05 *
Pacific Islander		2.21 *	2.25 *		-0.17	-0.04
Hispanic		-15.65 *	-15.54 *		-12.92 *	-12.71 *
Unknown		-23.63 *	-23.11 *		-22.91 *	-22.56 *
ELL		-15.15 *	-15.04 *		-24.80 *	-24.68 *
Redesignated EL		7.33 *	7.08 *		4.75 *	4.61 *
Title 1 (vs non-Title 1)		-23.79 *	-23.58 *		-21.83 *	-21.67 *

* p < .05

The effect of being an English Language Learner (ELL) is negative, and consistent with previous research is the fact that performance gap between ELL and non-ELL students is smaller

in mathematics because it is less language dependent. There are differences by race/ethnicity categories as well (the reference group is White students). The results also account for which grade a student is in. Generally, students' performance decreases with grade. While results pertaining to student background are not the focus of this analysis, checking these results is important in that it lends some credibility to findings that are of interest.

The importance of including these student covariates, as noted, is to be able to examine the marginal effects of the specific disability categories, net of other characteristics related to achievement. Hence, the estimates presented in Table 9, Model 3 represent the marginal effect, or performance gap for each specific category *ceteris paribus*. Also, these effects are additive. For example, a girl is estimated to score approximately 3.2 points less than a boy in Mathematics, but a girl who is classified with SLI would be expected to score about 11.7 (3.2+8.5) points lower than a boy, not classified with SLI.

This simple additivity depends on whether any interaction effects exist among the variables. Preliminary analyses examined whether there existed any interaction, or joint, effects among some of the pertinent student characteristics and specific disability categories. For example the results in Model 3 suggest that students classified with SLD score approximately 52.4 points lower in Mathematics than students without the special education designation. The results from Model 3 also indicate that girls score approximately 3.2 points lower than boys in Mathematics. Examining the joint effect accounts for potential effects of different distributions of girls and boys in the SLD category and examines whether among SLD students specifically, girls score differently than boys. Hence, preliminary analyses examined whether there were interaction effects between specific disability categories and gender, grade, and race/ethnicity. In general the results indicated that that there were not statistically significant interaction effects, nor were these effects substantively significant. Further, creating interaction effects between specific disability categories and race/ethnicity uses too many degrees of freedom and creates a model that even this large dataset cannot support. These results are likely not robust to variations in specifications. The results, therefore, imply that only main effects need to be considered.

Table 10: Effect size estimates for student characteristics

CST Model	Mathematics			ELA		
	1 <u>Estimate</u>	2 <u>Estimate</u>	3 <u>Estimate</u>	1 <u>Estimate</u>	2 <u>Estimate</u>	3 <u>Estimate</u>
Mean Scale Score						
Grade		-0.15 *	-0.15 *		-0.04 *	-0.04 *
Effect of Being designated Special Education (Any)	-1.11 *	-0.91 *		-0.65 *	-0.49 *	
GATE		1.11 *	1.11 *		0.72 *	0.72 *
MR			-1.61 *			-0.85 *
HH_DEAF			-0.41 *			-0.33 *
SLI			-0.17 *			-0.13 *
VI			-0.50 *			-0.30 *
ED			-0.93 *			-0.41 *
OI_TBI			-0.75 *			-0.31 *
OHI			-0.77 *			-0.39 *
SLD			-1.05 *			-0.56 *
DB_MD			-1.18 *			-0.48 *
AUT			-0.92 *			-0.41 *
Demographic Effects (compared to Whites)						
Girls (vs. Boys)		-0.07 *	-0.06 *		0.11 *	0.11 *
African American		-0.62 *	-0.61 *		-0.36 *	-0.35 *
American Indian		-0.31 *	-0.31 *		-0.19 *	-0.19 *
Asian		0.43 *	0.42 *		0.06 *	0.06 *
Pacific Islander		0.04 *	0.05 *		0.00	0.00
Hispanic		-0.31 *	-0.31 *		-0.20 *	-0.19 *
Unknown		-0.47 *	-0.46 *		-0.35 *	-0.34 *
ELL		-0.30 *	-0.30 *		-0.38 *	-0.37 *
Redesignated EL		0.15 *	0.14 *		0.07 *	0.07 *
Title 1 (vs non-Title 1)		-0.48 *	-0.47 *		-0.33 *	-0.33 *

* p < .05

In order to get a more intuitive estimate, Table 10 presents the results of the three models in terms of effect sizes⁶. The effect sizes are substantively large for each of the specific disability categories. Students designated as MR, for example, are expected to score 1.61 standard deviations below non-MR students in Mathematics, *ceteris paribus*. Students designated with SLD, the largest specific disability category, are expected to score about 1.05 and 0.56 standard deviations below students without a special education designation, in Mathematics and ELA respectively.

The results in Tables 9 and 10 represent the baseline (as of 2002-2003) performance gaps by specific disability categories, net of student background characteristics and also demonstrated that for SLI, ED, SLD and DB_MD students, their performance gaps varied significantly among schools. Table 11 presents results of reduced models for Mathematics and ELA that attempt to account for this variability with a limited number of school context variables. This should be considered a preliminary investigation of factors that might account for the variation in these effects. As is discussed below, many school context variables will not be appropriate for modeling cross-sectional data.

The school context variables considered here include only aggregate proportions of student background characteristics. That is, for each school, the percentage of GATE, Special Education, non-Minority⁷, ELL, and Title 1 are calculated. A proxy for school enrolment is also used. Table 11 presents only those school factors that were statistically significantly related to mean school achievement, and to the three specific disability categories demonstrated, above, to vary significantly between schools (that is the effects of being designated into one of the three categories.). They are interpreted as follows.

For example, as Tables 9 and 11 demonstrate, students designated as SLD demonstrate lower performance by about 54 and 38 points compared to non-SLD students, in Mathematics and ELA respectively. However, as noted, this effect varies among schools. For example in Mathematics, an average student classified as SLD, in a school where SLD students are performing one standard deviation above average, as compared to a similar student attending a school where SLD students are performing one standard deviation below average, would score

⁶ The effect size is estimated as $b1j/(s.d. \text{ of the outcome})$. The standard deviation of this sample is 49.9 for Mathematics and 65.9 for ELA.

⁷ Non-Minority = White.

about 46.6 points higher⁸. Model 4 attempts to use a limited number of school context variables to account for the variation in the SLD achievement gap as well as SLI, ED, and MD_DB.

School Mean achievement is positively related to all of the aggregate student characteristics, after accounting for the effects of individual student characteristics on achievement. For both Mathematics and ELA, school size is inversely associated with mean school performance – although substantively this effect is small. SLI varied only for ELA, and the percentage of Title 1 students and school size are both related to this variation. ED varies for both Mathematics and ELA. Again, school size was positively related to performance. That is the gap between ED students and non-ED students, was smaller in larger schools. The SLD achievement gap is associated with the number percentage of students with several aggregate school factors. Consistent with the specific disability categories, larger schools are estimated to have smaller SLD- Non-SLD achievement gaps. The MD_DB effect was also related to school size in a manner consistent with the other special education categories. It is interesting to note that for specific disability categories, the percentage of ELL students at a school exacerbates the negative consequences of that category. Another consistent result is that the percentage of Students with Disabilities at a school generally does not affect the achievement gap between students with specific categories and students without. The results in Table 11 indicate that magnitude of these limited school context variables is relatively small, compared to the overall gaps. This clearly warrants further research; likely by combining school data from other sources (e.g. the CBEDS data) that could examine programmatic and teacher qualification effects.

⁸ The between school variance in the SLD effect in Mathematics is 544.7 (Table 9). Hence, the s.d. = $\sqrt{544.7} = 23.3$. $23.3 \times 2 = 46.6$.

Table 11: HLM parameter estimates for the effect of school context (1)

CST	Mathematics	Reading
	4	4
	<u>Estimate</u>	<u>Estimate</u>
Mean Scale Score	322.34 *	320.11 *
Effect of School Context (2)		
% GATE students	10.73 *	8.14 *
% Special Education students	2.40 *	1.64 *
% Non-Minority (3)	4.17 *	2.11 *
% ELL Students	1.53 *	0.63 *
% Title 1 Students	3.20 *	2.43 *
School Size (4)	-0.96 *	-0.55 *
SLI	-8.55 *	-8.64 *
Effect of School Context (2)		
% Title 1 Students		-0.44 *
School Size (4)		0.19 *
ED	-48.19 *	-29.04 *
Effect of School Context (2)		
% GATE students	-4.41 *	
% ELL Students	-3.04 *	
School Size (4)	1.40 *	0.55 *
SLD	-54.08 *	-37.89 *
Effect of School Context (2)		
% of GATE students	-3.77 *	-3.32 *
% Non-Minority (3)	-3.38 *	
% ELL Students	-3.49 *	
% Title 1 Students	-1.09 *	
School Size (4)	1.62 *	0.41 *
DB_MD	-75.50 *	-31.75 *
Effect of School Context (2)		
% ELL Students	-3.21 *	
School Size (4)	3.35 *	
Variance Components		
SLI		45.5 *
ED	383.0	151.5 *
SLD	258.1	77.0 *
DB_MD	132.8	
Variance reduction from unconditional Estimates (Table 9)		
SLI		10.5%
ED	15.4%	16.1%
SLD	52.6%	25.8%
DB_MD	78.5%	

Notes:

* p < .05

- (1) Model 4, at the student level, contains all of the variables presented in Table 10. Only the results of the mean and the three specific disability categories that varied among schools are
- (2) The effect of school context presents change in achievement associated with a 10% change in the school context variable.
- (3) Non-Minority = White.
- (4) School Size is a proxy for student enrolment and estimated by the number of valid test scores in each school. Figures displayed represent the change in achievement per 100 students.

Conclusions

It is important to examine performance gaps net of student background characteristics as this reduces the effects of potential confounding factors. However, cross-sectional analyses such as the one conducted to establish the baseline estimates are inadequate to monitor progress in reducing these gaps.

The existing achievement gaps between special education students and non-special education students are substantively large and careful analysis of ensuing years' data needs to be under-taken in order to properly monitor progress towards reducing these gaps. The results suggest that it would be inappropriate to monitor the special education achievement gap using a dichotomous special education or not special education designation. Similarly, the results imply that classifying students into high or low incidence special education categories would be inappropriate. Given that the results indicate substantial differences in performance gaps among specific disability categories. Overall, the results indicate that these performance gaps are relatively uniform across schools (and by extension across local districts). This is supported by the fact that approximately 75% of the variation in student achievement scores is attributable to students within schools. Further, only three specific disability categories demonstrated any statistically significant variability among schools. Of course, one of these is SLD, the largest category, and this warrants further investigation. The simple cross-level models attempting to account for the between-school variation in the achievement gaps of specific disability categories do not present results that allow for conclusive policy and practice recommendations. However, the results are suggestive as to a couple of points pertaining to specific achievement gaps. One, the percentage of students designated as special education at a school does not affect how large the gap is – implying there are no “economies of scale”; two, as the percentage of ELL students at a school increase, so do the performance gaps of specific disability categories – implying that for students in some disability categories schools are not able to successfully handle several competing sources of achievement gaps; and three, while larger schools tend to have lower overall achievement, larger schools, have slightly smaller special education – non-special education achievement gaps -- for some specific disability categories.. Clearly, all of these areas warrant additional research.

While the HLM methodology addresses some issues, the simple two level model cannot address internal validity issues that will arise when attempting to monitor performance gaps over time. Ideally, a hierarchical growth model to assess change (Bryk and Raudenbush, 1987), within a growth curve approach (Rogosa, Brandt, and Zimowski, 1982), should be utilized. Longitudinal modeling has several advantages that can address issues of internal validity (Campbell and Stanley, 1963), as well as specific issues associated with the relationship between baseline and growth. In fact longitudinal hierarchical linear growth models can address spurious negative correlations that standard pre-post designs produce (Raudenbush and Bryk, 2002). Another benefit of longitudinal modeling is that the effect of test modifications or school program variables can be examined as well⁹.

A growth curve approach estimates growth trajectories for each student, which then allows for comparisons among specific disability categories' growth, relative to students without such designations – i.e. how much that gap is closing. Given that there exist external estimates of acceptable differences in performance due to actual cognitive differences, and not due to school related factors, relative growth estimates should be based on the difference between the proposed gap and the current gap divided by the time to meet the objectives. Statistically, a longitudinal growth model can explicitly test whether any specific group is “catching up,” but the model cannot address to what extent this catching up represents acceptable progress towards the objective. A lower bound, then, should be statistically significant growth parameter estimates for each of the specific disability categories.

⁹ This is not possible using a cross-sectional dataset, as students most in need of modifications receive them – generating negative estimates for the effects of the modifications. But in longitudinal studies, students become their own controls and growth rates can be estimated and compared.

References

- Aitkin, M. and Longford, N. (1986) "Statistical Modeling Issues in School Effectiveness Studies," *Journal of the Royal Statistical Society*, v. 149, pt. 1, pp.1-43.
- Bryk, A. S. and S. W. Raudenbush. (1987). Application of Hierarchical Linear Models to Assessing Change, *Psychological Bulletin*, v. 104(3), pp. 396-404.
- Burstein, L. (1980) The analysis of multi-level data in educational research and evaluation, *Review of Research in Education*, 4, pp. 158-233.
- Campbell, D.T. and J. Stanley (1963) *Experimental and Quasi Experimental Designs for Research*, Chicago, Rand McNally.
- Freedman, D., Pisani, R., and Purves, R. (1978) *Statistics*, WW Norton and Company, New York.
- Hardy, Melissa A. (1993) "Regression with Dummy Variables," SAGE University Paper Series on Quantitative Applications in the Social Sciences, 07-093, Newbury Park, CA.
- Pedhazur, E. (1982). *Multiple regression in behavioral research*. Fort Worth: Harcourt Brace Jovanovich.
- Raudenbush, W.S. and A.S. Bryk (2002). Hierarchical Linear Models: Applications and Data Analysis Methods, Sage Publications, Newbury Park.
- Rogosa, D.R., Brandt, D., and Zimowski, M. (1982) "A Growth Curve Approach to the Measurement of Change," *Psychological Bulletin*, v. 92, pp. 726-748.

Annex 1: Statistical Methodology employed in the analysis

While the focus of this analysis is not on school effects, in it is, none-the-less, informative to examine the extent to which specific disability categories' performance gaps vary among schools – as it is not unlikely that some schools may do a better job with some students than others.

Briefly, the HLM at level 1 is

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{1ij} + \dots + \beta_{pj}X_{pij} + r_{ij}, \text{ where } r_{ij} \sim N(0, \sigma^2) \quad (1)$$

Where Y_{ij} is the outcome for student i in school j , and X_{1ij} to X_{pij} are student level covariates. These covariates include student characteristics such as gender, race/ethnicity, or language status. At level 2

$$\beta_{0j} = g_{00} + u_{0j} \text{ where } u_{0j} \sim N(0, \tau_{00}) \quad (2)$$

$$\beta_{1j} = g_{10} + u_{1j} \text{ where } u_{1j} \sim N(0, \tau_{11}) \quad (3)$$

.

.

$$\beta_{pj} = g_{p0} + u_{pj} \text{ where } u_{pj} \sim N(0, \tau_{pp}) \quad (4)$$

where g_{00} is the overall mean of the sample, and u_{0j} is unique increment associated with school j (Raudenbush and Bryk, 2002). g_{10} to g_{p0} are the average slopes (or in the case of categorical variables, the average effect of being assigned that particular category as compared to a reference group (Hardy, 1993). The random effects u_{0j} to u_{pj} are the unique effects associated to school j (Raudenbush and Bryk, 2002). Equation 2 through 4 can be extended to include school level context variables that would be used to account for the significant variation in the unique effects of schools. However, that is not the focus of this analysis. Equations 2-4 are tested to using a χ^2 ((Raudenbush and Bryk, 2002) test to determine whether the variation in the level 1 parameter estimates are significant. That is. X_1 might be gender¹⁰, β_{1j} then represents the CST achievement gap between girl and boys. This gap is simultaneously modeled as a function of an

¹⁰ Gender is coded as 1=girls and 0=boys.

overall grand mean achievement gap (g_{10}) and a unique school effect u_{1j} . If the associated significance test for u_{1j} reveals that this value differs significantly from 0, then this indicates that the size achievement gap between girls and boys depends on which specific school in which the students are enrolled. If the estimate for u_{1j} is not statistically different from 0, then a single mean achievement gap estimate adequately summarizes the relationship between gender and achievement. If this is the case then (3) would be reduced to:

$$\beta_{1j} = g_{10} \tag{3b}$$

Annex 2: Results using CAT6 Mathematics and Reading as the outcome measures.

These results are generally similar to the ones reported above, using the CST. The main differences are that the CST results tend to be less homogeneous within schools.

Table 8 presents the results of the various models used to identify performance gaps among students. The results displayed under model 0 are the baseline or random

Table 8: Random effects and variance components

Variance Components	Mathematics Models			Reading Models		
	<u>0</u>	<u>2</u>	<u>3</u>	<u>0</u>	<u>2</u>	<u>3</u>
Student	420.7	330.0	306.0	341.6	262.6	247.6
School	88.4	25.4	27.8	62.2	16.1	17.0
SLI						3.6
SLD			17.0			11.9
DB_MD			69.3			65.6
	Percent	Percent accounted for		Percent	Percent accounted for	
	Attributable			Attributable		
Student	82.6	21.6	27.3	84.6	23.1	27.5
School	17.4	71.3	68.6	15.4	74.1	72.7

coefficient results. These results indicate that the majority of the variation in CAT6 scores is attributable to students, within schools. In fact 82.6% and 84.6% of the variation in Mathematics and Reading scores are attributable to students, respectively. The variance estimates for students can be used to compare student performance within a school. For example, in mathematics, a student one standard deviation above average would be expected to score about 41 points¹¹ higher than a student one standard deviation below average, in the same school. Conversely, about 15% to 17% of the variation in outcomes is attributable to schools. This means that differences among schools (e.g. school size) can only account for 15% of the variation in achievement scores. Roughly, this means that a hypothetical 10 point achievement gap between girls and boys, could at best be reduced by 1.5 points by addressing differences *between* schools.

¹¹ The student level variance is 420. The standard deviation = the square root of the variance = 20.5. $20.5 \times 2 = 41$.

Each of the specific disability categories was tested for randomly varying slopes, but only the three indicated in Table 8 varied randomly. In other words, except for SLD, DB_MD and (SLI in Reading), the achievement gaps estimated for each specific disability category does not vary among schools. That is, the average district-wide gaps between special education students (except the specific groups noted above) and non-special education students does not vary among schools – differences between schools (e.g. enrolment) will not account for any of the gap sizes. On the other hand randomly varying slopes (or gaps) for SLD, for example, indicate that some schools are doing better (i.e. smaller gap) and some are doing worse (i.e. larger gap). In fact, these results indicate that better schools would have a SLD-Non-SLD achievement gap that is 16.4 and 13.8 points smaller than a school doing poorly with SLD students, in Mathematics and Reading respectively. This between-school variation is discussed further below.

First student level effects are examined. The results displayed in Table 9 present the parameter estimates derived from three models, each of which build upon the previous model. Separate models are tested for Mathematics and Reading. Model 1 simply categorizes students as being designated as special education or not, but does not account for other student characteristics and potentially confounds the effects of special education status with other student background factors that might affect achievement. Model 2 adds student concomitant variables. Model 3 identifies each specific category separately and includes student concomitant variables as well.

The result from Model 1 indicate that for both Mathematics and Reading, special education students are expected to score approximately 20 and 16.4 NCEs lower than non-special education students. This result is not unexpected. Model 2 accounting for differences among students in other concomitant characteristics demonstrates the effect of confounding. The special education parameter estimate (achievement gap) is reduced by approximately 20% once student background characteristics are included. This indicates that comparing unadjusted means over-estimates the gap in special education student achievement by about 20%.

The results of model 3 indicate that there are both substantively large and statistically significant performance gaps between special education and non-special education students. Model 3 further indicates that these performance gaps vary substantially by specific disability category. This implies that grouping special education categories into high and low incidence is likely not appropriate and would incorrectly imply that students within the high or low incidence

categories demonstrate similar achievement patterns. Again, the student covariates conform to expectations. Girls perform less well than boys in Mathematics, but better in Reading.

Table 9: HLM parameter estimates for student characteristics

CAT6 Model	Mathematics			Reading		
	1	2	3	1	2	3
	<u>Estimate</u>	<u>Estimate</u>	<u>Estimate</u>	<u>Estimate</u>	<u>Estimate</u>	<u>Estimate</u>
Mean NCE	46.86 *	43.61 *	45.26 *	41.72 *	39.37 *	40.64 *
Grade		-1.70 *	-1.46 *		-0.93 *	-0.77 *
Effect of Being designated Special Education (Any)	-20.01 *	-16.10 *		-16.37 *	-13.09 *	
GATE		21.08 *	19.93 *		17.89 *	16.95 *
MR			-30.38 *			-22.59 *
HH_DEAF			-9.80 *			-9.54 *
SLI			-2.83 *			-2.86 *
VI			-11.89 *			-7.68 *
ED			-15.91 *			-11.92 *
OI_TBI			-14.08 *			-7.26 *
OHI			-13.50 *			-9.56 *
SLD			-18.49 *			-14.37 *
DB_MD			-20.46 *			-14.23 *
AUT			-15.66 *			-11.56 *
Demographic Effects (compared to Whites)						
Girls (vs. Boys)		-0.20 *	-1.00 *		3.63 *	2.99 *
African American		-12.20 *	-11.46 *		-9.44 *	-8.80 *
American Indian		-5.68 *	-5.67 *		-4.05 *	-4.02 *
Asian		6.77 *	5.95 *		1.07 *	0.42
Pacific Islander		1.54 *	0.44		0.15	-0.71 *
Hispanic		-5.87 *	-6.03 *		-4.37 *	-4.48 *
Unknown		-11.01 *	-9.35 *		-10.65 *	-9.37 *
ELL		-6.89 *	-6.59 *		-10.17 *	-9.91 *
Redesignated EL		4.84 *	3.47 *		2.89 *	1.78 *
Title 1 (vs non-Title 1)		-9.34 *	-8.73 *		-8.07 *	-7.57 *

* p < .05

The effect of being an English Language Learner (ELL) is negative, and consistent with previous research is the fact that performance gap between ELL and non-ELL students is smaller in mathematics because it is less language dependent. There are differences by race/ethnicity categories as well (the reference group is White students). The results also account for which grade a student is in. Generally, students' performance decreases with grade. While results

pertaining to student background are not the focus of this analysis, checking these results is important in that it lends some credibility to findings that are of interest.

The importance of including these student covariates, as noted, is to be able to examine the marginal effects of the specific disability categories, net of other characteristics related to achievement. Hence, the estimates presented in Table 9, Model 3 represent the marginal effect, or performance gap for each specific category *ceteris paribus*. Also, these effects are additive. For example, a girl is estimated to score approximately 1.0 NCE less than a boy, but a girl who is classified with SLI would be expected to score about 3.8 NCEs lower than a boy, not classified with SLI.

This simple additivity depends on whether any interaction effects exist among the variables. Preliminary analyses examined whether there existed any interaction, or joint, effects among some of the pertinent student characteristics and specific disability categories. For example the results in Model 3 suggest that students classified with SLD score approximately 18.5 NCEs lower in Mathematics than students without the special education designation. The results from Model 3 also indicate that girls score approximately 1.0 NCE lower than boys in Mathematics. Examining the joint effect accounts for potential effects of different distributions of girls and boys in the SLD category and examines whether among SLD students specifically, girls score differently than boys. Hence, preliminary analyses examined whether there were interaction effects between specific disability categories and gender, grade, and race/ethnicity. In general the results indicated that that there were not statistically significant interaction effects, nor were these effects substantively significant. Further, creating interaction effects between specific disability categories and race/ethnicity uses too many degrees of freedom and creates a model that even this large dataset cannot support. These results are likely not robust to variations in specifications. The results, therefore, imply that only main effects need to be considered.

Table 10: Effect size estimates for student characteristics

CAT6 Model	Mathematics			Reading		
	1 <u>Estimate</u>	2 <u>Estimate</u>	3 <u>Estimate</u>	1 <u>Estimate</u>	2 <u>Estimate</u>	3 <u>Estimate</u>
Mean NCE						
Grade		-0.09 *	-0.08 *		-0.05 *	-0.05 *
Effect of Being designated Special Education (Any)	-1.05 *	-0.85 *		-0.96 *	-0.77 *	
GATE		1.11 *	1.05 *		1.05 *	1.00 *
MR			-1.60 *			-1.33 *
HH_DEAF			-0.52 *			-0.56 *
SLI			-0.15 *			-0.17
VI			-0.63 *			-0.45 *
ED			-0.84 *			-0.70 *
OI_TBI			-0.74 *			-0.43 *
OHI			-0.71 *			-0.56 *
SLD			-0.97 *			-0.85 *
DB_MD			-1.08 *			-0.84 *
AUT			-0.82 *			-0.68 *
Demographic Effects (compared to Whites)						
Girls (vs. Boys)		-0.01 *	-0.05 *		0.21 *	0.18 *
African American		-0.64 *	-0.60 *		-0.56 *	-0.52 *
American Indian		-0.30 *	-0.30 *		-0.24 *	-0.24 *
Asian		0.36 *	0.31 *		0.06 *	0.02 *
Pacific Islander		0.08 *	0.02 *		0.01 *	-0.04 *
Hispanic		-0.31 *	-0.32 *		-0.26 *	-0.26 *
Unknown		-0.58 *	-0.49 *		-0.63 *	-0.55 *
ELL		-0.36 *	-0.35 *		-0.60 *	-0.58 *
Redesignated EL		0.25 *	0.18 *		0.17 *	0.10 *
Title 1 (vs non-Title 1)		-0.49 *	-0.46 *		-0.47 *	-0.45 *

* p < .05

In order to get a more intuitive estimate, Table 10 presents the results of the three models in terms of effect sizes¹². The effect sizes are substantively large for each of the specific disability categories. Students designated as MR, for example, are expected to score 1.66 standard deviations below non-MR students in Mathematics, ceteris paribus. Students

¹² The effect size is estimated as $b1j/(s.d. \text{ of the outcome})$. The standard deviation of this sample is 19 for Mathematics and 17 for Reading, which are somewhat below the metric calibrated s.d. of 21. That is, Normal Curve Equivalent are scaled so that the norming population mean is 50 and the s.d. is 21.

designated with SLD, the largest specific disability category, are expected to score about 0.97 and 0.85 standard deviations below students without a special education designation, in Mathematics and Reading respectively.

The results in Tables 9 and 10 represent the baseline (as of 2002-2003) performance gaps by specific disability categories, net of student background characteristics and also demonstrated that for SLI, SLD and DB_MD students, their performance gaps varied significantly among schools. Table 11 presents results of reduced models for Mathematics and Reading that attempt to account for this variability with a limited number of school context variables. This should be considered a preliminary investigation of factors that might account for the variation in these effects. As is discussed below, many school context variables will not be appropriate for modeling cross-sectional data.

The school context variables considered here include only aggregate proportions of student background characteristics. That is, for each school, the percentage of GATE, Special Education, non-Minority¹³, ELL, and Title 1 are calculated. A proxy for school enrolment is also used. Table 11 presents only those school factors that were statistically significantly related to mean school achievement, and to the three specific disability categories demonstrated, above, to vary significantly between schools (that is the effects of being designated into one of the three categories.). They are interpreted as follows.

For example, as Tables 9 and 11 demonstrate, students designated as SLD demonstrate lower performance by about 18.6 and 14.4 NCEs compared to non-SLD students. However, as noted, this effect varies among schools. Model 4 attempts to use a limited number of school context variables to account for the variation in the SLD achievement gap. The SLD achievement gap is associated with the number percentage of students who are designated as ELL in a school. It is important to reiterate that this a separate effect from the individual student being designated as ELL. As the percentage of ELL students increases, the SLD achievement gap increases (irrespective of whether or not the SLD student is designated ELL). In fact, comparing schools at either end of the possible ELL percent range (95% C.I) a school the upper end (+ two standard deviations), as compared to a school at the lower end (- two standard deviations) would

¹³ Non-Minority = White.

Table 11: HLM parameter estimates for the effect of school context (1)

CAT6	Mathematics 4 Estimate	Reading 4 Estimate
Mean NCE	45.21	40.56
Effect of School Context (2)		
% GATE students	3.03 *	2.24 *
% Special Education students	0.52 *	
% Non-Minority (3)	0.68 *	0.77 *
% ELL Students	0.42 *	
% Title 1 Students	1.03 *	0.83 *
School Size (4)	-0.21 *	-0.08 *
SLI	-2.83 *	-2.92 *
Effect of School Context (2)		
% Title 1 Students		-0.18 *
School Size (4)		0.06 *
SLD	-18.62 *	-14.38 *
Effect of School Context (2)		
% of GATE students		-0.45 *
% ELL Students	-0.46 *	
School Size (4)	0.10 *	
DB_MD	-22.86 *	
Effect of School Context (2)		
% Title 1 Students	-0.89 *	
School Size (4)	0.72 *	
Variance Components	Variance	Variance
SLI		3.2
SLD	14.2	11.7
DB_MD	15.1	70.4
Variance reduction from unconditional Estimates (Table 9)		
SLI		12.1%
SLD	16.6%	1.9%
DB_MD	78.2%	0.0%

Notes:

* $p < .05$

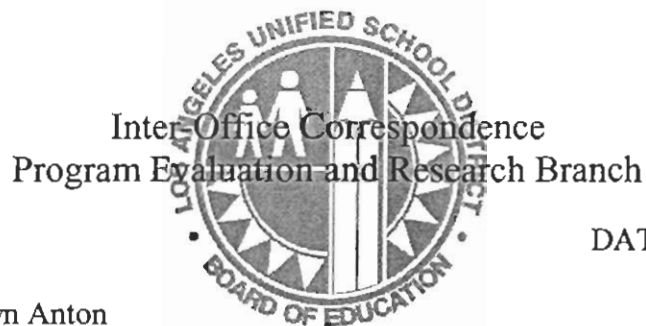
(1) Model 4, at the student level, contains all of the variables presented in Table 10. Only the results of the mean and the three specific disability categories that varied among schools are displayed.

(2) The effect of school context presents change in achievement associated with a 10% change in the school context variable.

(3) Non-Minority = White.

(4) School Size is a proxy for student enrolment and estimated by the number of valid test scores in each school. Figures displayed represent the change in achievement per 100 students.

have a SLD achievement gap that is about 4 NCEs larger (effect size = .21). The results in Table 11 indicate that magnitude of these limited school context variables is relatively small, compared to the overall gaps. This clearly warrants further research; likely by combining school data from other sources (e.g. the CBEDS data) that could examine programmatic and teacher qualification effects.



DATE: December 30, 2003

TO: Donnalyn Anton

FROM: Jeffrey A. White

CC: Esther Wong
Ted BartellSUBJECT: MATCHED CST SCORES FOR SPECIAL EDUCATION STUDENTS

Attached are matched individual student California Standards Test (CST) scores for Special Education Students. The Special Education identification comes from the primary disability codes on the 2003 STAR data layout and includes:

- Mental Retardation
- Hard of Hearing
- Deaf
- Speech/Language Impairment
- Visual Impairment
- Emotional Disturbance
- Orthopedic Impairment
- Other Health Impairment
- Established Medical Disability
- Deaf-Blindness
- Multiple Disabilities
- Autism
- Traumatic Brain Injury

Students who took the CSTs with exclusionary accommodations and/or modifications are not included in this report. Please contact me if you have questions or would like me to walk you through this report (213-241-8266).

**Appendix A:
2002 to 2003 Matched Individual Student California Standards Test Scores by District and Special Education Identification**

District B	ENGLISH LANGUAGE ARTS					MATHEMATICS					SCIENCE				HISTORY						
	Grade/ Subject	% Proficient/Above	2002	2003	Change	% Basic/Above	% Turnover In Rank	Loss	Gain	Grade/ Subject	% Proficient/Above	2002	2003	Change	Grade/ Subject	% Proficient/Above	2002	2003	Change		
Special Education	3	391	9.7	8.7	-1.0	-0.5	27.4	21.0		3	457	20.4	16.4	-3.9	-3.7	26.5	27.4				
	4	300	7.0	10.3	3.3	15.0	8.7	44.7		4	406	11.8	14.5	2.7	7.1	15.3	36.9				
	5	237	8.0	9.3	1.3	3.8	18.6	25.7		5	275	8.7	12.0	3.3	-2.5	27.6	21.8				
	6	306	3.9	4.6	0.7	-1.0	26.1	18.6		6	324	2.5	4.8	2.2	-1.2	16.0	30.2				
	7	317	2.2	3.8	1.6	6.0	11.0	32.2		7	411	2.7	2.4	-0.2	6.8	20.4	28.0				
	8	288	2.1	1.4	-0.7	4.2	9.0	32.3			8	611		1.3							
	9	418	2.6	4.5	1.9	9.8	6.9	41.4				446		0.9							
	10	292	3.1	3.8	0.7	3.1	8.9	38.0				118		2.5							
	11	154	1.3	1.9	0.6	0.6	14.3	14.9				19		31.6							
	Not Identified for Special Education																				
	District C																				
Special Education																					
3	403	14.4	18.6	4.2	3.0	23.1	26.3		3	456	25.7	30.3	4.6	2.6	21.7	34.2					
4	389	11.8	16.7	4.9	23.4	8.5	40.3		4	462	15.8	26.8	11.0	13.6	12.3	43.5					
5	292	20.5	19.2	-1.3	4.8	22.6	26.7		5	322	21.1	23.6	2.5	-0.9	31.4	26.1					
6	430	6.7	8.1	1.4	4.2	24.2	19.8		6	449	7.6	8.7	1.1	-1.8	19.2	23.2					
7	410	2.2	6.1	3.9	6.8	10.0	36.3		7	475	3.4	3.4	0.0	3.6	22.9	28.4					
8	394	4.7	5.7	1.0	3.9	13.5	28.1			732		4.2									
9	396	4.0	7.8	3.8	12.6	11.1	45.5			280		3.9									
10	267	7.1	5.6	-1.5	5.2	11.6	33.7			62		11.3									
11	173	6.9	4.0	-2.9	-1.7	24.9	11.6			25		16.0									
Not Identified for Special Education																					
3	4,217	33.7	36.8	3.1	1.4	21.0	26.1		3	4,328	47.0	55.0	8.1	5.6	17.4	36.7					
4	4,192	34.1	43.4	9.3	14.9	10.0	39.9		4	4,220	40.1	58.6	18.5	14.1	10.6	49.4					
5	3,933	36.4	40.4	4.0	3.9	17.3	23.8		5	4,038	42.1	47.8	5.7	1.4	22.3	27.7					
6	4,868	25.6	29.0	3.5	0.8	17.9	22.7		6	4,902	27.4	27.7	0.3	2.2	20.4	25.7					
7	4,794	22.7	32.3	9.6	10.6	9.5	35.2		7	4,890	24.2	25.5	1.3	3.7	20.9	23.1					
8	4,407	26.2	26.7	0.5	6.2	14.9	26.8			5,930		16.1									
9	5,109	24.6	35.4	10.8	10.3	9.4	41.6			3,895		20.9									
10	3,944	35.9	37.4	1.5	4.3	16.3	27.8			2,440		24.2									
11	2,857	39.6	41.3	1.7	0.7	20.2	23.2			1,677		32.4									
HS Summative																					
General Math																					
Algebra I																					
Geometry																					
Algebra II																					
HS Summative																					
General Math																					
Algebra I																					
Geometry																					
Algebra II																					
HS Summative																					
Earth Sci.																					
Biology																					
Chemistry																					
Physics																					
Integrated 1																					
Integrated 2																					
Integrated 3																					
Integrated 4																					
Earth Sci.																					
Biology																					
Chemistry																					
Physics																					
Integrated 1																					
Integrated 2																					
Integrated 3																					
Integrated 4																					
Earth Sci.																					
Biology																					
Chemistry																					
Physics																					
Integrated 1																					
Integrated 2																					
Integrated 3																					
Integrated 4																					
Earth Sci.																					
Biology																					
Chemistry																					
Physics																					
Integrated 1																					
Integrated 2																					
Integrated 3																					
Integrated 4																					
Earth Sci.																					
Biology																					
Chemistry																					
Physics																					
Integrated 1																					
Integrated 2																					
Integrated 3																					
Integrated 4																					

**Appendix A:
2002 to 2003 Matched Individual Student Standards Test Scores by District and Special Education Identification**

District	ENGLISH LANGUAGE ARTS					MATHEMATICS					SCIENCE					HISTORY					
	Grade/ Subject	% Proficient/Above N 2002	2003	Change	% Basic/ Above Change	% Turnover In Rank Less Gain	Grade/ Subject	% Proficient/Above N 2002	2003	Change	% Basic/ Above Change	% Turnover In Rank Less Gain	Grade/ Subject	% Proficient/Above N 2002	2003	Change	Grade/ Subject	% Proficient/Above N 2002	2003	Change	
Special Education	3	333	17.7	18.6	0.9	-5.1	26.4	21.9													
	4	336	12.5	18.2	5.7	19.3	9.5	46.7													
	5	324	24.5	20.6	-0.9	5.6	19.3	25.9													
	6	331	8.5	12.1	3.6	4.5	24.2	23.9													
	7	370	6.2	8.1	1.9	5.4	13.8	32.4													
	8	263	6.1	6.1	0.0	2.3	16.3	25.5													
	9	350	5.7	9.1	3.4	11.7	30.3	41.1													
	10	238	5.9	5.5	-0.4	4.6	11.8	29.4													
	11	165	7.3	9.7	2.4	1.2	12.1	22.4													
	Not Identified for Special Education																				
	District E	3	3,748	45.3	46.1	0.8	-0.3	22.0	23.5												
4		3,586	46.6	53.2	6.6	11.8	9.6	36.2													
5		3,524	47.0	50.1	3.1	4.8	17.1	22.4													
6		3,493	35.5	38.9	4.4	0.0	16.0	23.5													
7		3,388	33.9	40.6	6.7	4.5	13.9	29.1													
8		2,744	36.1	34.8	-1.3	1.4	17.7	22.8													
9		3,523	34.6	42.3	10.7	9.3	9.7	39.3													
10		2,863	37.9	39.7	1.8	4.4	17.1	26.9													
11		2,210	40.7	41.9	1.2	-2.1	23.7	23.9													
Not Identified for Special Education																					
Special Education		3	279	7.2	8.6	1.4	-3.9	26.9	20.8												
	4	299	6.7	6.7	2.0	20.4	9.4	49.5													
	5	276	8.0	11.2	3.2	8.0	17.0	33.0													
	6	242	3.7	7.0	3.3	-1.7	24.8	20.7													
	7	301	2.7	3.3	0.6	2.3	13.3	23.3													
	8	293	1.7	1.4	-0.3	3.1	10.2	22.5													
	9	288	3.8	4.9	1.1	5.9	9.4	36.5													
	10	200	2.5	3.5	1.0	5.5	11.0	36.5													
	11	114	6.1	1.8	-4.3	4.4	21.9	14.0													
	Not Identified for Special Education																				
	Not Identified for Special Education	3	4,929	27.6	29.3	1.7	1.2	21.8	26.2												
4		4,733	26.9	33.8	6.9	17.3	9.1	41.3													
5		4,238	28.2	30.8	2.6	5.6	17.3	25.5													
6		4,512	21.9	26.9	5.0	-0.1	18.0	23.1													
7		4,182	20.5	26.6	6.1	6.6	13.4	30.3													
8		3,554	19.8	20.1	0.3	3.0	18.1	27.3													
9		4,017	15.1	22.8	7.7	10.1	11.8	38.4													
10		3,124	22.6	23.8	1.2	4.4	16.5	27.3													
11		2,293	25.8	24.7	-1.1	-1.0	26.1	21.1													
Not Identified for Special Education																					
HS Summative		3	5,017	37.5	48.5	11.0	7.1	16.4	38.8												
	4	4,817	38.7	53.7	17.1	14.6	10.9	48.8													
	5	4,398	36.2	39.2	3.0	-1.5	26.9	24.2													
	6	4,531	23.9	27.0	3.1	4.6	16.8	31.6													
	7	4,251	24.5	22.6	-2.0	3.8	24.8	21.7													
	8	4,399			6.9																
	9	3,951			7.6																
	10	2,566			8.5																
	11	1,236			8.6																
	HS Summative																				
	HS Summative	3	377	11.9	17.8	5.8	8.5	18.0	36.9												
4		414	9.2	20.5	11.4	13.8	13.0	43.5													
5		332	13.9	15.1	1.2	-2.4	25.0	20.2													
6		273	7.3	7.3	0.0	-5.5	25.6	23.4													
7		395	2.8	3.8	1.0	0.0	27.1	24.1													
8		598			0.8																
9		178			3.4																
10		70			2.9																
11		13			15.4																
HS Summative																					
Integrated 4		3	377	11.9	17.8	5.8	8.5	18.0	36.9												
	4	414	9.2	20.5	11.4	13.8	13.0	43.5													
	5	332	13.9	15.1	1.2	-2.4	25.0	20.2													
	6	273	7.3	7.3	0.0	-5.5	25.6	23.4													
	7	395	2.8	3.8	1.0	0.0	27.1	24.1													
	8	598			0.8																
	9	178			3.4																
	10	70			2.9																
	11	13			15.4																
	Integrated 4																				
	Integrated 3	3	377	11.9	17.8	5.8	8.5	18.0	36.9												
4		414	9.2	20.5	11.4	13.8	13.0	43.5													
5		332	13.9	15.1	1.2	-2.4	25.0	20.2													
6		273	7.3	7.3	0.0	-5.5	25.6	23.4													
7		395	2.8	3.8	1.0	0.0	27.1	24.1													
8		598			0.8																
9		178			3.4																
10		70			2.9																
11		13			15.4																
Integrated 3																					
Integrated 2		3	377	11.9	17.8	5.8	8.5	18.0	36.9												
	4	414	9.2	20.5	11.4	13.8	13.0	43.5													
	5	332	13.9	15.1	1.2	-2.4	25.0	20.2													
	6	273	7.3	7.3	0.0	-5.5	25.6	23.4													
	7	395	2.8	3.8	1.0	0.0	27.1	24.1													
	8	598			0.8																
	9	178			3.4																
	10	70			2.9																
	11	13			15.4																
	Integrated 2																				
	Integrated 1	3	377	11.9	17.8	5.8	8.5	18.0	36.9												
4		414	9.2	20.5	11.4	13.8	13.0	43.5													
5		332	13.9	15.1	1.2	-2.4	25.0	20.2													
6		273	7.3	7.3	0.0	-5.5	25.6	23.4													
7		395	2.8	3.8	1.0	0.0	27.1	24.1													

**Appendix A:
2002 to 2003 Matched Individual Student California Standards Test Scores by District and Special Education Identification**

ENGLISH LANGUAGE ARTS					MATHEMATICS					SCIENCE					HISTORY		
Grade/ Subtest	% Proficient/Above			% Basic/ Above	% Turnover in Rank	Grade/ Subtest	% Proficient/Above			% Basic/ Above	% Turnover in Rank	Grade/ Subtest	% Proficient/Above			Grade/ Subtest	% Proficient/Above
	N	2002	2003				Change	N	2002				2003	Change	N		
District F																	
Special Education																	
3	238	4.6	6.7	2.1	2.1	23.5	19.3										
4	209	3.8	7.7	3.9	19.6	7.7	54.1										
5	160	5.0	4.4	-0.6	6.9	18.1	32.5										
6	193	5.2	4.1	-1.1	1.0	29.5	15.5										
7	145	2.1	2.8	0.7	4.8	9.7	31.0										
8	157	0.0	1.9	1.9	1.9	7.6	23.6										
9	317	1.3	1.3	0.0	4.1	6.9	38.8										
10	190	0.5	1.6	1.1	3.2	10.0	35.3										
11	120	4.2	2.5	-1.7	3.3	15.8	25.0										
Not Identified for Special Education																	
3	4,388	19.8	20.2	0.4	-2.1	24.9	23.2										
4	4,214	18.4	25.0	6.6	19.4	9.4	44.6										
5	3,635	18.9	23.1	4.2	6.6	15.7	27.9										
6	3,407	17.4	12.6	-4.8	-0.9	22.2	19.9										
7	2,961	9.5	14.6	5.1	9.2	13.4	33.4										
8	2,849	10.2	10.7	0.5	6.6	14.8	29.2										
9	3,627	12.9	21.8	8.9	12.7	9.3	41.9										
10	2,580	22.5	23.6	1.1	6.4	16.0	28.8										
11	2,078	23.7	25.9	2.2	5.0	19.3	26.1										
District G																	
Special Education																	
3	194	2.6	4.6	2.0	-2.6	25.3	17.0										
4	203	3.9	4.9	1.0	20.2	8.9	47.8										
5	149	4.0	8.1	4.1	1.3	23.5	36.6										
6	239	2.5	0.8	-1.7	-5.0	36.1	16.7										
7	280	1.1	1.1	0.0	2.9	12.1	25.0										
8	248	1.2	2.0	0.8	3.6	6.9	27.0										
9	209	0.5	1.4	0.9	8.6	7.7	36.4										
10	174	3.4	3.4	0.0	1.1	10.3	25.3										
11	133	0.8	0.8	0.0	-1.5	12.8	18.0										
Not Identified for Special Education																	
3	4,515	19.1	18.3	-0.8	-3.0	27.0	23.3										
4	4,464	17.9	19.9	2.0	16.0	12.0	39.8										
5	3,915	16.4	19.8	3.4	7.7	16.9	29.6										
6	3,595	10.1	11.4	1.3	-1.6	23.5	20.0										
7	3,161	8.6	14.4	5.8	5.8	13.9	31.1										
8	2,804	11.4	12.0	0.6	4.5	16.4	27.9										
9	2,295	10.2	16.6	6.4	8.0	13.2	35.6										
10	1,927	14.8	16.6	1.8	3.3	18.7	28.5										
11	1,464	17.9	18.1	0.2	-0.8	24.0	21.6										
District H																	
Special Education																	
3	255	4.7	5.9	1.2	0.8	20.0	27.5										
4	279	3.6	6.5	2.9	9.0	13.3	36.9										
5	190	4.7	3.2	-1.6	2.1	23.2	23.7										
6	299	1.2	2.7	1.5	-0.4	20.5	30.5										
7	338	0.6	0.3	-0.3	2.1	24.9	24.0										
8	372		0.3														
9	290		0.0														
10	163		0.0														
11	37		0.0														
Not Identified for Special Education																	
3	4,720	26.3	30.0	3.7	2.0	23.9	33.2										
4	4,570	21.3	28.7	7.5	10.2	16.0	40.0										
5	4,217	20.0	21.0	1.1	-2.2	29.8	23.6										
6	3,645	9.9	7.8	-2.1	-1.4	22.9	25.6										
7	3,218	8.0	6.3	-1.7	-1.8	28.1	19.9										
8	2,239		4.2														
9	2,827		2.4														
10	1,909		1.5														
11	1,255		2.2														
District I																	
Special Education																	
3	255	4.7	5.9	1.2	0.8	20.0	27.5										
4	279	3.6	6.5	2.9	9.0	13.3	36.9										
5	190	4.7	3.2	-1.6	2.1	23.2	23.7										
6	299	1.2	2.7	1.5	-0.4	20.5	30.5										
7	338	0.6	0.3	-0.3	2.1	24.9	24.0										
8	372		0.3														
9	290		0.0														
10	163		0.0														
11	37		0.0														
Not Identified for Special Education																	
3	4,720	26.3	30.0	3.7	2.0	23.9	33.2										
4	4,570	21.3	28.7	7.5	10.2	16.0	40.0										
5	4,217	20.0	21.0	1.1	-2.2	29.8	23.6										
6	3,645	9.9	7.8	-2.1	-1.4	22.9	25.6										
7	3,218	8.0	6.3	-1.7	-1.8	28.1	19.9										
8	2,239		4.2														
9	2,827		2.4														
10	1,909		1.5														
11	1,255		2.2														
District J																	
Special Education																	
3	255	4.7	5.9	1.2	0.8	20.0	27.5										
4	279	3.6	6.5	2.9	9.0	13.3	36.9										
5	190	4.7	3.2	-1.6	2.1	23.2	23.7										
6	299	1.2	2.7	1.5	-0.4	20.5	30.5										
7	338	0.6	0.3	-0.3	2.1	24.9	24.0										
8	372		0.3														
9	290		0.0														
10	163		0.0														
11	37		0.0														
Not Identified for Special Education																	
3	4,720	26.3	30.0	3.7	2.0	23.9	33.2										
4	4,570	21.3	28.7	7.5	10.2	16.0	40.0										
5	4,217	20.0	21.0	1.1	-2.2	29.8	23.6										
6	3,645	9.9	7.8	-2.1	-1.4	22.9	25.6										
7	3,218	8.0	6.3	-1.7	-1.8	28.1	19.9										
8	2,239		4.2														
9	2,827		2.4														
10	1,909		1.5														
11	1,255		2.2														

**Appendix A:
2002 to 2003 Matched Individual Student California Standards Test Scores by District and Special Education Identification**

ENGLISH LANGUAGE ARTS										MATHEMATICS										SCIENCE										HISTORY	
Grade/ Subtest	% Proficient/Above			% Basic/ Above	% Turnover In Rank	Grade/ Subtest	% Proficient/Above			% Basic/ Above	% Turnover In Rank	Grade/ Subtest	% Proficient/Above			Grade/ Subtest	% Proficient/Above														
	N	2002	2003				Change	N	2002				2003	Change	N		2002	2003	Change	N	2002	2003	Change								
District J																															
Special Education																															
3	365	3.0	3.8	0.8	-3.6	20.5	17.0																								
4	340	2.1	3.8	1.7	12.4	8.2	47.9																								
5	280	2.5	4.3	1.8	6.1	15.4	35.0																								
6	324	1.5	1.9	0.4	-2.5	29.0	14.5																								
7	203	2.0	2.0	0.0	2.0	14.3	29.6																								
8	172	0.6	0.6	0.0	4.7	7.0	31.4																								
9	380	0.8	1.1	0.3	6.8	7.4	42.1																								
10	251	0.4	0.8	0.4	3.6	6.4	33.9																								
11	161	0.0	1.9	1.9	3.1	19.3	19.3																								
Not Identified for Special Education																															
District K																															
Special Education																															
3	4,774	18.0	19.5	1.5	0.1	23.4	24.4																								
4	4,673	19.3	24.1	4.8	19.0	10.3	43.0																								
5	4,098	18.0	20.9	2.9	6.8	16.6	28.0																								
6	3,696	11.2	12.8	1.6	-1.3	23.2	19.7																								
7	3,548	10.4	14.8	4.4	4.8	14.4	28.6																								
8	2,473	10.8	10.1	-0.7	2.9	17.8	25.1																								
9	3,797	9.2	17.1	7.9	14.7	9.4	44.6																								
10	2,581	15.6	18.0	2.4	7.9	15.0	31.5																								
11	1,952	18.4	18.6	0.2	2.5	21.5	22.4																								
Not Identified for Special Education																															
Special Education																															
3	322	9.3	6.2	-3.1	0.3	26.4	20.8																								
4	231	11.3	14.3	0.0	16.9	16.5	41.6																								
5	266	4.5	6.4	1.9	1.5	19.6	26.4																								
6	295	2.7	3.4	0.7	-3.1	32.2	18.3																								
7	344	2.6	3.2	0.6	5.5	12.2	27.3																								
8	256	1.6	2.3	0.7	2.0	13.7	25.0																								
9	216	1.9	2.8	0.9	10.6	7.9	39.8																								
10	151	2.0	3.3	1.3	6.0	6.6	37.7																								
11	84	3.6	2.4	-1.2	2.4	11.9	21.4																								
Not Identified for Special Education																															
Special Education																															
3	4,856	33.1	32.4	-0.7	-1.6	24.6	22.2																								
4	4,636	31.9	37.2	5.3	14.5	11.5	37.9																								
5	4,156	29.8	32.3	2.5	5.6	17.9	24.4																								
6	4,462	23.0	26.4	3.4	-1.3	20.3	21.4																								
7	4,256	19.3	25.4	6.1	5.2	13.6	27.9																								
8	3,809	23.6	22.2	-1.4	1.8	19.3	23.3																								
9	3,597	18.4	27.4	9.0	10.2	11.5	38.8																								
10	2,888	25.8	26.5	0.7	5.5	16.7	27.4																								
11	1,826	31.0	29.9	-1.1	0.8	22.6	20.8																								
Not Identified for Special Education																															
Special Education																															
3	4,956	43.1	49.7	6.6	4.0	20.0	34.1																								
4	4,682	37.7	51.2	13.5	11.6	13.0	43.2																								
5	4,353	36.8	35.9	-0.9	-3.8	29.6	21.0																								
6	4,476	23.2	22.7	-0.5	-0.8	23.4	23.5																								
7	4,296	18.9	19.1	0.2	1.7	22.6	20.6																								
8	4,663		12.0																												
9	3,006		11.6																												
10	1,940		13.7																												
11	1,195		14.1																												
Not Identified for Special Education																															
Special Education																															
3	357	15.7	15.4	-0.3	0.0	22.7	27.2																								
4	341	10.3	14.7	4.4	7.0	11.4	33.1																								
5	307	9.4	8.8	-0.7	-3.6	31.8	18.9																								
6	310	3.5	3.2	-0.3	-0.6	20.3	31.3																								
7	404	2.2	1.7	-0.5	1.2	26.5	23.5																								
8	469		1.7																												
9	194		1.5																												
10	45		8.9																												
11	15		6.7																												
Not Identified for Special Education																															
Special Education																															
3	357	15.7	15.4	-0.3	0.0	22.7	27.2																								
4	341	10.3	14.7	4.4	7.0	11.4	33.1																								
5	307	9.4	8.8	-0.7	-3.6	31.8	18.9																								
6	310	3.5	3.2	-0.3	-0.6	20.3	31.3																								
7	404	2.2	1.7	-0.5	1.2	26.5	23.5																								
8	469		1.7																												
9	194		1.5																												
10	45		8.9																												
11	15		6.7																												
Not Identified for Special Education																															
Special Education																															
3	357	15.7	15.4	-0.3	0.0	22.7	27.2																								
4	341	10.3	14.7	4.4	7.0	11.4	33.1																								
5	307	9.4	8.8	-0.7	-3.6	31.8	18.9																								
6	310	3.5	3.2	-0.3	-0.6	20.3	31.3																								
7	404	2.2	1.7	-0.5	1.2	26.5	23.5																								
8	469		1.7																												
9	194		1.5																												
10	45		8.9																												
11	15		6.7																												
Not Identified for Special Education																															
Special Education																															
3	357	15.7	15.4	-0.3	0.0	22.7	27.2																								
4	341	10.3	14.7	4.4	7.0	11.4	33.1																								
5	307	9.4	8.8	-0.7	-3.6	31.8	18.9																								
6	310	3.5	3.2	-0.3	-0.6	20.3	31.3																								
7	404	2.2	1.7	-0.5	1.2	26.5	23.5																								
8	469		1.7																												
9	194		1.5																												
10	45		8.9																												
11	15		6.7																												
Not Identified for Special Education																															
Special Education																															
3	357	15.7	15.4	-0.3	0.0	22.7	27.2																								
4	341	10.3	14.7	4.4	7.0	11.4	33.1																								
5	307	9.4	8.8	-0.7	-3.6	31.8	18.9																								
6	310	3.5	3.2	-0.3	-0.6	20.3	31.3																								