DOCUMENT RESUME

ED 477 931 TM 035 027

AUTHOR	DeMauro, Gerald E.
TITLE	Differential Consequential Validity and the Stability of Inferences across Ethnicity and Community on New York State Large Scale Tests.
PUB DATE	2001-06-00
NOTE	57p.; Version of a paper presented at the Annual Conference of the Council of Chief State School Officers (Houston, TX, June 2001).
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01/PC03 Plus Postage.
DESCRIPTORS	Academic Achievement; *Achievement Tests; Disadvantaged Youth; Elementary Secondary Education; English; Expectation; *Inferences; Language Arts; *Large Scale Assessment; *Reliability; State Programs; Testing Programs; *Validity
IDENTIFIERS	*New York

ABSTRACT

The consequences of large state testing are often uniformity of expectations for achievement. The largest impact of higher standards, then, are realized by traditionally disenfranchised student populations, particularly the least affluent who are most likely to bear the yoke of low expectation. This paper advances S. Messick's (1981) fundamental validity concerns in a way that precludes a sharp distinction between evidentiary and consequential validity, and evaluates validity information from 3 years of fourth grade English Language Arts assessment in New York State. The analyses considered more than 200,000 students in each of the 3 years. The results attest to the positive impact of higher expectations for students in groups that have been disenfranchised historically. An appendix shows the general linear model regression by year, ethnicity, and needs resource category. (Author/SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 477 931

Differential Consequential Validity and the Stability of Inferences across Ethnicity and Community On New York State Large Scale Tests

Gerald E. DeMauro
Coordinator of State Assessment
New York State Education Dept.
March, 2002

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

G. E. Demauro

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

TM035027

Revision of a Paper presented as part of a symposium at the Annual Conference of the Council of Chief State School Officers, Houston, Texas, June 2001

BEST COPY AVAILABLE

Table of Contents

Table of Tables and Figures

Abstract

The consequences of large state testing are often uniformity of expectations for achievement. The largest impact of higher standards, then, are realized by traditionally disenfranchised student populations, particularly the least affluent who are most likely to bear the yoke of low expectation. This paper advances Messick's (1981) fundamental validity concerns in a way that precludes a sharp distinction between evidentiary and consequential validity, and evaluates validity information from three years of fourth grade English Language Arts assessment in New York State. The results attest to the positive impact of higher expectations for students in groups that have been disenfranchised historically.

Differential Construct Validity and the Stability of Inferences Made on Large Scale State Tests

Gerald E. DeMauro

Coordinator of State Assessment

In 1993, New Jersey instituted a more difficult series of graduation examinations to replace the ninth grade instrument they had been using. Many in the state's educational community were concerned that this development would disenfranchise the students in the least affluent districts. The developmental research (c.f. Good and Brophy, 1990) however, suggests that the uniform high standards and consequences for all students would remove the insidious burden of low expectations from these children.

Examinations in reading, mathematics, and writing were offered four times, beginning in the fall of the junior year, and thereafter in the spring of the junior year, the fall of the senior year, and the spring of the senior year for all students not passing earlier. The students from the 30 poorest school districts in the state initially had a passing rate of about half that of all other school districts in the state. The cumulative passing rate, using the initial cohort size as the denominator, for students from the 30 neediest school districts after the senior spring administration was about the same as the rate of all other districts a year and a half earlier. Clearly, the test was having positive impact on student's in the state's least affluent school districts.

Many states have increased the demands on children for high school graduation, each concerned that there is an institutional expectation of lower achievement for the students from the poorest school districts. New York State, which has a testing system that dates back to 1861, has also adopted higher learning standards and has revamped the testing system to be less course-bound and more standards-bound. Every child will be

expected to pass five examinations: Comprehensive English; Mathematics; Global History and Geography; United States History and Government; and one of four science examinations from among Physical Setting/Physics, Physical Setting/Chemistry, Physical Setting/Earth Science, or Living Environment.

Associated with that commencement-level program, New York requires testing in fourth and eighth grade and a large number of additional high school examinations related to course achievement and language proficiency. In fourth grade, students take examinations in English Language Arts (ELA-4), Mathematics (Math-4), Science, and Social Studies (actually administered at the start of fifth grade). In eighth grade, students take examinations in English Language Arts (ELA-8), Mathematics (Math-8), Science, Social Studies, Technology, and Language Proficiency.

Purpose of the Assessment System

The State assessment system is responsible for certifying that the students have achieved the New York State Learning Standards at the elementary, intermediate, and commencement levels. These Learning Standards are designed to delineate the knowledge and skills needed for achievement in either the next set of grades (for the elementary and intermediate standards) or postsecondary work and academic environments. Therefore, the primary validity obligation is to support inferences based on test scores that students have achieved the required skill and knowledge.

In that context, predictive evidence of success subsequent to the tests adds to the documentation that the tests are sensitive to acquisition of the Learning Standards but is neither necessary nor sufficient to sustain validity. Clearly, some tests will measure some areas of achievement and others will measure other areas. Validity documentation using predictive sources of evidence, then, becomes a reasoned judgment about the information provided by the test scores and the appropriate application of that information.

Conceptualizing Construct Validity for State Testing

Overview: two reconstructions. Validity documentation for state purposes, must be mindful of the levels of interpretation: child, school, program, district, and state. The State's responsibility extends to each level to insure access to the Learning Standards for every student through the instructional program. This responsibility has profound implications for validity.

For example, Messick (1981) proposed that two issues must be addressed to evaluate the use of a test for a specific purpose: the quality of the test for the interpretation made of its use and the social consequences of the test use. Messick saw the first issue as evidentiary and the second as ethical.

The first reconstruction of a validity paradigm for state testing depends on recognizing that both issues are ultimately ethical and both must be evaluated by the weight of evidence. Both evidence and consequences are based on student performance on test questions, and that same performance is evidentiary that is, provides the evidence from which ethical judgments must derive. Different performance implies different consequences; performance is never free of consequence. Consequences, as well, generalize from the performance of individuals to the performance of institutions. Traditionally, some evidence of test and item quality has been viewed as support for testing. It is clear, however, that all evidence on test and item quality, both before and after the administration of an instrument, must be evaluated for its support of the usefulness of this test at this time, for this purpose, and for these examinees. Ethical judgment is obliged first to the truths revealed by this evaluation.

The second validity reconstruction involves the false separation of differential validity as an independent evaluation of test utility. In fact, validity must be concerned with the distribution of error. That distribution, when systematically related to group

membership, threatens validity of the instrument at that time, for that use, and for all children because it limits the capacity for generalization and contaminates the evidentiary basis on which judgments are made. Although this sounds theoretical, it is, in fact, observable. Every child in the state is a member of several demographic groups. When the interpretation of the test score is differential with regard to group membership, the interpretation of every child's score is less certain and the consequences of testing are inequitable.

Purpose of this Study

The state's role in giving universal access and in implementing the promise of higher learning standards to reduce differential expectations based on socioeconomic conditions requires first an address of the test scores. The current analysis focuses on the results of three years of New York State testing in English Language Arts and in Mathematics for grade four and eight, the end grades for the elementary and intermediate Learning Standards, respectively. In particular, the performance of populations identified by ethnicity and by socioeconomic status of the school district is considered to determine if the consequence of universal access reduces disparities among populations. Simply put, if the test is expected to be valid for all groups, equity of expectations for performance on the test should stimulate equity of instructional intervention and ultimately of achievement. Therefore, the consistency of implementation of State Learning Standards and the consequential validity of the associated testing should have this most profound effect on the populations that have been traditionally disenfranchised, particularly in the urban centers. It is these populations that will make the most progress when equity of achievement, that is, equity of consequence, is realized.

Methods

The Examination

The examination (ELA-4) consists of 28 multiple-choice questions, three scores given analytically, and four scores of 4, 3, 3, and 4 points, respectively, awarded holistically. The examination is administered in late January to the State's fourth grade population. For the purposes of this study, three measures of achievement were first evaluated as possible dependent variables: the scale score means, the percentage of children achieving level 3 or higher, and the school performance index.

The scale scores range from 455 to 800. The examination is pattern scored using partial credit models specified by three parameters for multiple-choice questions (no partial credit is awarded for these) and of two parameters for the holistically scored questions.

The standard-setting studies for state examinations divided the score range into four levels, of which level 3 was considered adequate achievement of the standards. The second possible dependent variable, then, was achievement of level 3 or higher on the examination.

Finally, for school accountability purposes, the state employs an index which awards 100 points for a score of level 2 (602-644) and 200 points for a score of level 3 (645-691) or level 4 (692-800). On this index, schools must achieve a minimum average, which was 140 in 2001.

Table 1 shows that the correlations among these three variables were very high. The additional information gained by using the whole scale score, then, recommended that it serve as the study's dependent variable.

Table 1a

Intercorrelations among Scale Scores, Probability of Reaching Level 3, and School Accountability Index on the Grade 4 English Language Arts Test (ELA-4), 1999 (n=210,563), 2000 (n=216,467), and 2001 (n=215,091)

				Correlations		
Year	Variable	Mean	Standard Deviation	Scale Score	Prob. 3+	Account. Index
1999	Scale Score	640.87	35.07	-----	.865*	.853
	P(3+)	0.48	0.50	.865*	-----	.955*
	Account. Ind.	136.65	67.85	.853	.955*	-----
2000	Scale Score	652.48	42.21	----	.853*	.821
	P(3+)	0.59	0.49	.853*	-----	.975*
	Account. Ind.	148.94	66.82	.821	.975*	-----
2001	Scale Score	653.47	43.60	-----	.848*	.826
	P(3+)	0.60	0.49	.848*	-----	.975*
	Account. Ind.	149.71	67.65	.826	.975*	-----

*Spearman correlation. All other correlation coefficients are Pearson product moment.

Table 1b

Intercorrelations among Scale Scores, Probability of Reaching Level 3, and School Accountability Index on the Grade 8 English Language Arts Test (ELA-8), 1999 (n=187,089), 2000 (n=195,503), and 2001 (n=195,225)

				Correlations		
Year	Variable	Mean	Standard Deviation	Scale Score	Prob. 3+	Account. Index
1999	Scale Score	699.31	30.31	-----	.775*	.873
	P(3+)	0.28	0.45	.775*	-----	.835*
	Account. Ind.	102.02	72.97	.873	.835*	-----
2000	Scale Score	696.64	33.88	----	.861*	.857
	P(3+)	0.59	0.49	.861*	-----	.943*
	Account. Ind.	131.41	69.52	.857	.943*	-----
2001	Scale Score	697.76	34.95	-----	.862*	.848
	P(3+)	0.45	0.50	.862*	-----	.943*
	Account. Ind.	131.43	69.75	.848	.943*	-----

*Spearman correlation. All other correlation coefficients are Pearson product moment.

Table 1c

Intercorrelations among Scale Scores, Probability of Reaching Level 3, and School Accountability Index on the Grade 4 Mathematics Test (Math-4), 1999 (n=215,574), 2000 (n=219,845), and 2001 (n=213,879)

				Correlations		
Year	Variable	Mean	Standard Deviation	Scale Score	Prob. 3+	Account. Index
1999	Scale Score	651.53	40.57	-----	.816*	.804
	P(3+)	0.67	0.47	.816*	-----	.983*
	Account. Ind.	156.73	66.73	.804	.983*	-----
2000	Scale Score	648.11	37.44	----	.827*	.804
	P(3+)	0.65	0.48	.827*	-----	.982*
	Account. Ind.	155.68	65.68	.804	.982*	-----
2001	Scale Score	654.29	40.83	-----	.803*	.782
	P(3+)	0.69	0.46	.803*	-----	.986*
	Account. Ind.	160.01	64.30	.782	.986*	-----

*Spearman correlation. All other correlation coefficients are Pearson product moment.

Table 1d

Intercorrelations among Scale Scores, Probability of Reaching Level 3, and School Accountability Index on the Grade 8 Mathematics (Math-8), 1999 (n=193,411), 2000 (n=216,467), and 2001 (n=215,091)

Year	Variable	Mean	Standard Deviation	Correlations: Scale Score	Correlations: Prob. 3+	Correlations: Account. Index
1999	Scale Score	701.39	42.82	-----	.840*	.871
	P(3+)	0.38	0.49	.840*	-----	.893*
	Account. Ind.	108.60	81.48	.871	.893*	-----
2000	Scale Score	704.24	41.44	----	.849*	.858
	P(3+)	0.40	0.49	.849*	-----	.906*
	Account. Ind.	115.02	79.42	.858	.906*	-----
2001	Scale Score	703.61	44.04	-----	.848*	.852
	P(3+)	0.40	0.49	.848*	-----	.904*
	Account. Ind.	113.60	80.03	.852	.904*	-----

*Spearman correlation. All other correlation coefficients are Pearson product moment.

Samples

The samples were divided into six socioeconomic categories called "needs/ resource" categories. These designations are based on community type and district wealth factors: New York City, Big Four (Buffalo, Rochester, Syracuse, and Yonkers), high-need urban/suburban, high-need rural, average, and low need. Ethnicity was grouped as: African American, European American, Hispanic American, and Others (Asian American, American Indian/Native American, Pacific Islander American, and others). The representation by ethnicity and needs/resource category is given in Tables 2a (grade 4) and 2b (grade 8).

Table 2a

Representation in the 1999, 2000, and 2001 Examinee Samples by Ethnicity and School District Needs Resource (Socioeconomic) Category (Grade 4)

Year=1999

Needs/ Resource		Ethnicity: African American	European American	Hispanic American	Other	Total
New York City	N	27,957	12,185	27,150	8,060	75,352
	Prop.	.371	.162	.360	.107	.358
Big Four	N	3,339	1,600	1,052	3,571	9,652
	Prop.	.349	.167	.110	.374	.045
High-Need Urb./Sub.	N	4,749	8,114	1,978	1,533	16,374
	Prop.	.290	.496	.121	.094	078
High-Need Rural	N	384	12,589	269	639	13,881
	Prop.	.028	.907	.019	.046	.066
Average	N	3,717	55,558	2,378	4,232	65,885
	Prop.	.056	.843	.036	.064	.313
Low Need	N	510	25,168	729	3,102	29,509
	Prop.	.017	.853	.025	.105	.140
Total 1999	N	40,656	115,214	33,556	21,137	210,563
	Prop.	.193	.547	.159	.100	1.000

Table 2a

Representation in the 1999, 2000, and 2001 Examinee Samples by Ethnicity and School District Needs Resource (Socioeconomic) Category (Grade 4)

Year=2000

Needs/ Resource		Ethnicity: African American	European American	Hispanic American	Other	Total
New York City	N	28,082	11,916	28,693	8,583	77,274
	Prop.	.363	.154	.371	.111	.357
Big Four	N	5,701	2,727	1,499	260	10,187
	Prop.	.560	.268	.147	.026	.047
High-Need Urb./Sub.	N	4,946	8,488	2,039	294	15,767
	Prop.	.314	.538	.129	.019	.073
High-Need Rural	N	377	13,066	313	235	13,991
	Prop.	.027	.933	.022	.017	.065
Average	N	3,884	60,410	2,723	1,274	68,291
	Prop.	.057	.885	.040	.019	.316
Low Need	N	760	27,831	936	1,430	30,957
	Prop.	.025	.899	.030	.046	.143
Total 2000	N	43,750	124,438	36,203	12,076	216,467
	Prop.	.202	.575	.167	.056	1.000

Table 2a

Representation in the 1999, 2000, and 2001 Examinee Samples by Ethnicity and School District Needs Resource (Socioeconomic) Category (Grade 4)

Year=2001

Needs/ Resource		Ethnicity: African American	European American	Hispanic American	Other	Total
New York City	N	29,054	11,672	29,719	8,491	78,936
	Prop.	.368	.148	.377	.108	.367
Big Four	N	5,598	2,464	1,496	282	9,840
	Prop.	.569	.250	.152	.029	.046
High-Need Urb./Sub.	N	5,372	8,449	2,322	352	16,495
	Prop.	.326	.512	.141	.021	.077
High-Need Rural	N	424	12,489	284	215	13,412
	Prop.	.032	.931	.021	.016	.062
Average	N	3,950	575,848	2,846	1,386	66,030
	Prop.	.060	.876	.043	.021	.307
Low Need	N	755	27,184	1,021	1,418	30,378
	Prop.	.025	.895	.034	.047	.141
Total 2001	N	45,153	120,106	37,688	12,144	215,091
	Prop.	.210	.558	.175	.057	1.000

Table 2b

Representation in the 1999, 2000, and 2001 Examinee Samples by Ethnicity and School District Needs Resource (Socioeconomic) Category (Grade 8)

Year=1999

Needs/ Resource		Ethnicity: African American	European American	Hispanic American	Other	Total
New York City	N	22,029	10,585	20,750	6,396	59,760
	Prop.	.369	.177	.347	.107	.319
Big Four	N	3,708	2,313	1,122	223	7,366
	Prop.	.503	.314	.152	.030	.039
High-Need Urb./Sub.	N	3,935	7,688	1,566	1,018	14,207
	Prop.	.277	.542	.110	.072	.076
High-Need Rural	N	298	13,717	220	364	14,599
	Prop.	.020	.940	.015	.025	.078
Average	N	2,882	58,302	1,998	2,958	66,140
	Prop.	.044	.882	.030	.045	.354
Low Need	N	450	20,629	596	3,342	25,017
	Prop.	.018	.825	.024	.134	.134
Total 1999	N	33,302	113,324	26,252	14,301	187,089
	Prop.	.178	.605	.140	.076	1.000

Table 2b

Representation in the 1999, 2000, and 2001 Examinee Samples by Ethnicity and School District Needs Resource (Socioeconomic) Category (Grade 8)

Year=2000

Needs/ Resource		Ethnicity: African American	European American	Hispanic American	Other	Total
New York City	N	23,786	10,724	22,049	6,705	63,264
	Prop.	.376	.170	.349	.106	.324
Big Four	N	4,169	2,356	1,249	238	8,012
	Prop.	.520	.294	.156	.030	.041
High-Need Urb./Sub.	N	4,206	8,522	1,550	309	14,587
	Prop.	.288	.584	.106	.021	.075
High-Need Rural	N	321	13,745	248	244	14,558
	Prop.	.022	.944	.017	.017	.075
Average	N	3,657	59,820	2,463	1,537	67,477
	Prop.	.054	.887	.037	.023	.345
Low Need	N	761	24,687	912	1,245	27,605
	Prop.	.028	.894	.033	.045	.141
Total 2000	N	36,900	119,854	28,471	10,278	195,503
	Prop.	.189	.613	.146	.053	1.000

Table 2b

Representation in the 1999, 2000, and 2001 Examinee Samples by Ethnicity and School District Needs Resource (Socioeconomic) Category (Grade 8)

Year=2001

Needs/ Resource		Ethnicity: African American	European American	Hispanic American	Other	Total
New York City	N	23,256	10,713	21,955	6,961	62,885
	Prop.	.370	.170	.349	.111	.322
Big Four	N	4,111	2,158	1,267	397	7,933
	Prop.	.518	.272	.160	.050	.046
High-Need Urb./Sub.	N	4,385	7,959	1,867	431	14,642
	Prop.	.300	.544	.128	.029	.075
High-Need Rural	N	313	12,841	239	317	13,710
	Prop.	.023	.937	.017	.023	.070
Average	N	3,778	59,226	2,710	1,999	67,713
	Prop.	.056	.875	.040	.030	.347
Low Need	N	751	25,344	962	1,285	28,342
	Prop.	.027	.894	.034	.045	.145
Total 2001	N	36,594	118,241	29,000	11,390	195,225
	Prop.	.187	.606	.149	.058	1.000

Analyses

Two major analyses were undertaken of the results. The first was a general linear model in which year of administration, ethnicity, and needs/resource category were treated as independent variables and the scale score was the dependent variable. Post hoc analyses included quantitative contrasts to determine functions of growth over the three-year span. The errors of prediction, based on projections from linear and quadratic coefficients from the post hoc quantitative contrasts, were computed by ethnicity and needs/resource categories.

The second analysis employed analyses of standardized mean differences for ethnicity and needs/resource category. These were computed by:

1. identifying the highest scoring of the 24 ethnic groups within each needs/resource category in 1999;

2. finding the standardized difference with each of the other 23 groups based on the highest 199 group's mean and standard deviation each year.

These measures provide an estimate of the gain of each group over time with respect to the performance of the group that was highest initially, and therefore directly address the issue of closing the gap in achievement.

Group Differences, 1999-2001

General findings. Appendix A provides the general linear model results for the analysis of scale score by ethnicity, needs resource category of school district, year of administration, and the first- and second-order interactions of these three variables. All the main effects and the interactions proved significant for each of the four tests. Individual test results are summarized below.

ELA-4. Scheffe post hoc contrasts revealed that each year's scores were significantly higher than the previous year's (means of 640.87, 652.48, and 653.47, respectively), and that European American students scored highest in each year, followed by other students, and then by African American and Hispanic American students. Each year, the Hispanic American and African American populations scored virtually the same, each significantly lower than each of the other two ethnic categories (see Table 3). Scheffe post hoc analyses also revealed that in each year, each ethnic group scored significantly higher than it had in the previous year. These analyses also showed that students from the low need school districts scored significantly higher each year than those from average, high need rural, high need urban/suburban, New York City, and the Big Four school districts.

ELA-8. Each main and interactive effect was statistically significant. The mean scores dropped from 699.31 in 1999 to 696.64 in 2000. The 2001 mean (697.76) represented a significant recovery compared to 2000, but it was still significantly below the 1999 mean. The patterns of scores are given in Table 3.

The community types scored in the same order each year: Low need, average, high need rural, high need urban/suburban, New York City, and Big Four. For the three years combined, needs resource category was significantly higher than the next lowest.

Means for each of the four ethnic designations were in the same order each year, as well (see Table 3): Others, European Americans, Hispanic Americans, and African Americans. Again, each group scored significantly higher than the next over the three year period. By group, African Americans actually dropped from 1999 (683.77) to 2000 (678.58) and regained some of the loss in 2001 (679.91). The same pattern held for Hispanic Americans (684.29, 679.55, and 681.26, respectively) and others (707.77, 706.67, and 707.10). For European Americans, the first year drop (706.29 to 705.40) was regained in 2001 (706.42).

By needs/resource category, there were three patterns of scoring means. The first was to drop from 1999 to 2000 and then to recover in 2001 but not enough to regain the first year loss. This pattern characterized New York City, high need urban/suburban districts, high need rural districts, and average districts. The second pattern was to drop in 2000 and then drop again in 2001. This pattern characterized the Big Four districts. Finally, the third pattern is to increase each year, and this pattern characterizes the low need school districts (see Table 3 for the means).

Math 4. Again, all of the main effects and interactions were statistically significant. The highest mean was achieved in 2001 (654.29). From 1999 to 2000 the mean score fell from 651.53 to 648.11. That pattern held for each of the six needs resource district types and for each of the four ethnic groups. Statewide, the differences among the three means are all significant. Low need districts (677.43) scored highest, followed by average (662.16), high need rural (653.34), high need urban/suburban (647.20), New York City (635.13), and the Big Four districts (634.79). Only New York City and Big Four districts were not significantly different. The first four types of districts had means in the same order over all three years, while New York City had the lowest means in 1999, but the Big Four had the lowest 2000 and 2001 means.

Among the ethnic categories, European Americans (664.80) scored highest in the three years, followed by other ethnic groups (664.42), Hispanic Americans (630.99), and African Americans (628.94), each group being significantly lower than the next higher group. By year, European American students scored highest in 1999 (665.27) and in 2001 (668.37), while the other ethnic groups scored higher in 2000 (660.94). The means are shown in Table 3.

Math 8. As with the other examinations, all main effects and interactions proved to be statistically significant. Students made large mean gains between 1999 and 2000 (701.39 and 704.24), and then dropped somewhat in 2001 (703.61). That pattern characterized African American students and European American students, while Hispanic American students and those of other ethnic groups both gained steadily over the three test administrations.

The pattern of large gains in 2000 and then a smaller regression in 2001 also characterized students in the average, high need rural, high need urban/suburban, and New York City districts. In the Big Four districts, after the initial gain in 2000, the 2001 drop actually brought the mean below the 1999 mean. Among the low need districts, after an initial rise in 2000, there was also a nominal gain in 2001 (see Table 3).

Overall, the other ethnic groups achieved the highest means (718.81 over three years), followed by European Americans (716.60), Hispanic Americans (678.27), and African Americans (675.02). By needs resource categories, the overall mean scores were: Low need (730.70), average (713.87), high need rural (705.36), high need urban/suburban (694.56), New York City (684.43, and Big Four (680.79) districts.

Each year, the overall score means remained in the same order for ethnicity as described by the overall pattern above. While Hispanic Americans and other ethnic groups rose each year in mean scores, European Americans and African Americans first

rose from 1999 to 2000 and then regressed in 2001, although for European Americans the drop was nominal.

The order of scoring was also the same in each year by needs resource category as the pattern described above. For all five of these categories, the scoring pattern was to rise from 1999 to 2000 and then to regress in 2001. The exception was for the low need districts that made nominal gains between 2000 and 2001 (see Table 3).

Table 3

Observed Means by Needs Resource Category and Ethnicity on the Grade Four English Language Arts (ELA-4), Grade Four Mathematics (Math-4), Grade Eight English Language Arts (ELA-8), and Grade Eight Mathematics (Math-8) Examinations 1999-2001

			Means			
Needs/Resource	Ethnicity	Year	ELA-4	Math-4	ELA-8	Math-8
New York City	African Am.	1999	621.49	625.31	683.30	671.52
		2000	628.69	622.44	677.74	674.08
		2001	629.58	627.96	679.20	671.30
	Eur. Am.	1999	648.60	661.03	708.25	708.40
		2000	662.28	655.50	706.70	709.97
		2001	664.43	663.13	707.26	710.07
	Hispanic Am.	1999	619.01	627.03	683.14	674.65
		2000	626.58	624.65	677.49	676.08
		2001	628.68	631.13	679.29	676.20
	Others	1999	643.92	659.76	706.41	714.48
		2000	656.29	658.59	704.26	715.86
		2001	657.71	664.51	705.29	717.64
	All	1999	627.38	635.13	690.14	683.45
		2000	636.15	632.07	685.37	685.19
		2001	637.42	638.12	686.90	684.63
Big Four	African Am.	1999	621.20	627.82	682.11	674.05
		2000	627.56	625.41	675.43	674.28
		2001	628.43	630.03	676.43	670.43
	Eur. Am.	1999	640.81	652.02	697.37	697.16
		2000	648.83	645.01	693.53	700.14
		2001	650.53	654.03	692.67	695.97
	Hispanic Am.	1999	627.02	629.96	681.93	671.45
		2000	637.60	628.90	678.82	673.93
		2001	640.72	635.79	678.47	671.79
	Others	1999	628.70	653.61	696.45	698.22
		2000	648.89	649.28	698.03	696.90
		2001	653.95	659.20	687.93	700.87
	All	1999	627.92	615.26	687.31	681.48
		2000	635.28	636.70	681.95	682.38
		2001	636.52	637.96	681.75	678.56
Urban/sub.	African Am.	1999	625.97	633.68	682.30	673.89
		2000	637.31	634.55	677.51	675.98
		2001	642.00	642.18	677.70	682.77
	Eur. Am.	1999	644.18	654.91	699.85	704.26
		2000	656.31	652.63	697.04	707.45
		2001	657.55	659.86	693.38	707.64
	Hispanic Am.	1999	627.26	635.98	684.45	675.98
		2000	640.52	636.98	682.11	682.77
		2001	643.53	642.32	682.81	682.62
	Others	1999	636.98	650.45	695.49	698.63
		2000	661.47	658.57	698.55	707.78
		2001	658.61	664.33	696.22	708.22
	All	1999	636.18	636.19	692.98	691.75
		2000	648.41	647.61	689.85	696.81
		2001	650.53	649.65	690.14	695.01

Rural	African Am.	1999	628.90	631.87	682.85	673.22
		2000	636.45	631.45	680.30	683.53
		2001	634.53	636.20	683.17	679.61
	Eur. Am.	1999	644.15	654.31	699.12	704.30
		2000	654.39	651.74	695.95	707.96
		2001	653.81	657.18	696.66	707.14
	Hispanic Am.	1999	632.52	640.92	686.70	682.97
		2000	643.58	639.54	685.53	688.57
		2001	646.49	643.64	682.75	683.31
	Others	1999	637.91	647.77	694.39	694.93
		2000	649.88	647.62	692.40	702.69
		2001	644.84	653.27	689.75	703.55
	All	1999	643.21	642.37	698.48	703.15
		2000	653.59	653.45	695.37	706.93
		2001	652.90	647.10	695.95	706.01
Average	African Am.	1999	634.45	642.29	689.21	682.73
		2000	645.19	640.34	685.03	690.43
		2001	645.12	645.93	686.92	688.39
	Eur. Am.	1999	651.86	664.38	705.73	713.80
		2000	664.83	660.10	704.51	717.85
		2001	665.40	667.70	705.17	716.96
	Hispanic Am.	1999	638.35	647.59	692.72	688.12
		2000	650.09	645.03	689.15	694.18
		2001	652.11	650.88	690.86	694.95
	Others	1999	654.58	667.87	706.75	712.96
		2000	668.72	666.60	709.54	721.91
		2001	664.32	671.36	712.00	727.85
	All	1999	650.57	648.86	704.66	711.50
		2000	663.20	661.36	703.00	715.52
		2001	663.69	660.65	703.78	714.56
Low Need	African Am.	1999	644.02	654.76	699.06	699.44
		2000	659.17	654.12	696.14	702.61
		2001	660.14	660.91	696.87	702.94
	Eur. Am.	1999	660.19	678.94	715.05	729.54
		2000	677.63	673.67	716.28	732.74
		2001	681.14	682.63	717.66	732.75
	Hispanic Am.	1999	645.31	657.79	699.28	701.43
		2000	659.59	654.43	698.52	706.88
		2001	664.32	661.26	699.52	706.18
	Others	1999	658.66	684.13	717.24	734.63
		2000	681.51	678.45	722.60	746.03
		2001	686.23	695.83	723.16	746.11
	All	1999	659.38	678.16	714.68	728.77
		2000	676.81	672.89	715.42	731.61
		2001	680.29	681.87	716.74	731.65
All	African Am.	1999	623.53	623.55	683.77	673.47
		2000	631.58	632.08	678.58	677.14
		2001	632.83	633.38	679.91	674.36
	Eur. Am.	1999	651.80	649.55	706.29	714.25
		2000	665.42	663.06	705.40	717.92
		2001	666.80	664.47	706.42	717.55
	Hispanic Am.	1999	621.80	628.62	684.29	676.30
		2000	630.59	638.24	679.55	679.03
		2001	632.94	640.17	681.26	679.36
	Others	1999	644.96	641.66	707.77	716.45
		2000	660.42	659.73	706.67	719.16
		2001	662.08	661.11	707.10	721.60
	All[1]	1999	640.87	651.53	699.31	701.39
		2000	652.48	648.11	696.64	704.24
		2001	653.47	654.29	697.76	703.61

Post hoc quantitative contrasts. Post hoc quantitative contrasts (Myers, 1973) were used to examine the nature of the growth functions on the four examinations. For each of the four tests, these revealed that the growth function over the three years had significant linear and quadratic components (see Appendices).

For all four examinations, the linear and quadratic functions were evident in each community type and ethnic group, as the initial first year growth leveled off considerably in the second year for ELA-4 and for Math 8. For ELA-8 and for Math 4, the initial drop in means between 1999 and 2000 was mitigated by a rise in mean scores in 2001.

The linear and quadratic coefficients, respectively, were : 6.25 and –1.75 for ELA-4, -0.78 and 0.63 for ELA-8, 1.38 and 1.60 for Math 4, and 1.11 and –0.58 for Math 8. The overall means for the three years were: 649.00 (ELA-4), 697.88 for ELA-8, 651.28 for Math 4, and 703.09 for Math 8. Figures 1-4 plot the coefficients.

Because the quantitative post hoc contrasts are orthogonal (additive), they can be separated as linear, quadratic, or both. Based on these separated coefficients, standard errors of prediction were computed by ethnicity and needs/resource grouping. This was accomplished by:

1. computing the projected linear, quadratic, and both functions growth based the respective coefficients and the added (both) linear and quadratic coefficients;
2. subtracting the observed scores for each year and squaring the differences;
3. taking the mean of the squared differences; and
4. taking the square root of that mean.

Tables 4a-4d present the results of these analyses. These errors can be interpreted as the degree of concurrence for each group between the actual, observed performance of that group, and the performance projected based on the linear, quadratic, and summed coefficients.

The standard errors of prediction are about the same for linear, quadratic, and summed components for all groups, and for the three functions, with some tendency to be lower for students from the high need urban/suburban, high need rural, and average districts. In each case, the ethnic and community type that is best characterized by linear or quadratic, or by the sum of both types of patterns are the same group. That is, the smallest errors are found for the same group for all three patterns. For ELA-4 this group is for students in the other ethnic group category in high need rural districts. For ELA-8, this pattern holds for Hispanic American students in low need districts. For Math 4 the smallest errors were for students in the other ethnic group category in low need districts, and for Math-8, the smallest errors were for European American students in the rural school districts.

Table 4a

Overall, Linear, and Quadratic Standard Errors of Prediction on the ELA-4, by Needs/Resource and Ethnicity

ETHNIC	RESOURCE	BOTH	CURVE	LINEAR
African American	New York City	48.24	46.13	49.20
African American	Big Four	43.04	40.64	44.12
African American	High Need US	38.94	37.34	39.73
African American	High Need Rural	37.75	35.84	38.66
African American	Average	35.90	34.96	36.44
African American	Low Need	33.76	34.57	33.58
African American	All	45.38	43.43	46.29
European American	New York City	43.41	44.30	43.17
European American	Big Four	39.84	39.64	40.06
European American	High Need US	35.21	35.64	35.20
European American	High Need Rural	33.31	33.50	33.40
European American	Average	35.29	36.69	34.86
European American	Low Need	40.54	43.04	39.59
European American	All	37.27	38.67	36.83
Hispanic American	New York City	49.64	47.43	50.63
Hispanic American	Big Four	39.21	37.69	39.97
Hispanic American	High Need US	37.57	36.16	38.28
Hispanic American	High Need Rural	32.84	31.71	33.46
Hispanic American	Average	33.36	33.09	33.64
Hispanic American	Low Need	33.35	34.40	33.07
Hispanic American	All	47.14	45.18	48.04
Other	New York City	42.31	42.66	42.29
Other	Big Four	38.87	36.83	39.83
Other	High Need US	35.98	35.29	36.42
Other	High Need Rural	32.80	31.78	33.38
Other	Average	35.24	36.55	34.84
Other	Low Need	39.65	41.94	38.81
Other	All	40.15	40.61	40.09
All	New York City	47.44	45.98	48.15
All	Big Four	41.24	39.50	42.07
All	High Need US	36.75	36.23	37.11
All	High Need Rural	33.42	33.49	33.56
All	Average	35.25	36.45	34.90
All	Low Need	40.15	42.57	39.24
All	All	40.98	40.94	41.14

Table 4b

Overall, Linear, and Quadratic Standard Errors of Prediction on the ELA-8, by Needs/Resource and Ethnicity

ETHNIC	RESOURCE	BOTH	CURVE	LINEAR
African American	New York City	38.17	35.82	39.04
African American	Big Four	35.29	32.42	36.32
African American	High Need US	35.31	32.62	36.29
African American	High Need Rural	34.49	32.17	35.36
African American	Average	31.75	29.97	32.46
African American	Low Need	29.19	28.93	29.43
African American	All	36.76	34.40	37.64
European American	New York City	35.14	36.28	34.89
European American	Big Four	32.63	32.04	32.97
European American	High Need US	30.41	30.36	30.58
European American	High Need Rural	28.99	28.76	29.23
European American	Average	29.54	30.54	29.36
European American	Low Need	33.73	36.10	33.05
European American	All	31.03	32.11	30.81
Hispanic American	New York City	37.97	35.59	38.84
Hispanic American	Big Four	36.03	33.48	36.97
Hispanic American	High Need US	32.41	30.09	33.29
Hispanic American	High Need Rural	31.17	29.04	31.99
Hispanic American	Average	30.66	29.44	31.21
Hispanic American	Low Need	28.60	28.67	28.75
Hispanic American	All	36.70	34.46	37.53
Other	New York City	36.83	37.66	36.69
Other	Big Four	31.47	30.61	31.90
Other	High Need US	31.25	30.89	31.53
Other	High Need Rural	30.77	29.78	31.24
Other	Average	33.48	34.89	33.15
Other	Low Need	36.37	38.97	35.60
Other	All	35.65	36.75	35.41
All	New York City	37.45	36.02	38.03
All	Big Four	34.52	32.42	35.32
All	High Need US	32.16	31.02	32.67
All	High Need Rural	29.20	28.87	29.47
All	Average	29.83	30.62	29.73
All	Low Need	33.68	35.96	33.04
All	All	33.31	33.20	33.49

Table 4c

Overall, Linear, and Quadratic Standard Errors of Prediction on the Math-4, by Needs/Resource and Ethnicity

ETHNIC	RESOURCE	BOTH	CURVE	LINEAR
African American	New York City	44.64	44.86	45.29
African American	Big Four	38.52	38.75	39.20
African American	High Need US	36.49	36.63	36.93
African American	High Need Rural	33.98	34.18	34.57
African American	Average	32.89	32.98	33.16
African American	Low Need	33.01	32.92	32.78
African American	All	41.85	42.05	42.45
European American	New York City	41.06	40.96	40.78
European American	Big Four	36.95	36.94	36.95
European American	High Need US	34.51	34.44	34.33
European American	High Need Rural	31.86	31.80	31.71
European American	Average	35.95	35.79	35.49
European American	Low Need	44.44	44.19	43.70
European American	All	38.00	37.85	37.56
Hispanic American	New York City	43.01	43.22	43.63
Hispanic American	Big Four	38.89	39.08	39.45
Hispanic American	High Need US	34.88	35.01	35.28
Hispanic American	High Need Rural	32.28	32.39	32.60
Hispanic American	Average	32.63	32.66	32.72
Hispanic American	Low Need	33.08	32.99	32.82
Hispanic American	All	41.30	41.48	41.85
Other	New York City	42.58	42.48	42.29
Other	Big Four	36.42	36.38	36.30
Other	High Need US	38.99	38.91	38.78
Other	High Need Rural	29.21	29.22	29.26
Other	Average	42.13	41.95	41.61
Other	Low Need	51.09	50.82	50.30
Other	All	43.25	43.12	42.86
All	New York City	43.27	43.41	43.68
All	Big Four	38.15	38.30	38.62
All	High Need US	35.30	35.33	35.40
All	High Need Rural	31.88	31.84	31.77
All	Average	35.77	35.64	35.38
All	Low Need	44.29	44.04	43.56
All	All	39.74	39.37	39.71

Table 4d

Overall, Linear, and Quadratic Standard Errors of Prediction on the Math-8, by Needs/Resource and Ethnicity

ETHNIC	RESOURCE	BOTH	CURVE	LINEAR
African American	New York City	54.01	51.54	54.26
African American	Big Four	48.46	45.74	48.74
African American	High Need US	47.68	45.29	47.93
African American	High Need Rural	45.73	43.43	45.97
African American	Average	41.41	39.71	41.60
African American	Low Need	35.05	34.72	35.10
African American	All	51.21	48.83	51.46
European American	New York City	42.03	42.48	42.00
European American	Big Four	38.97	38.25	39.07
European American	High Need US	36.92	37.12	36.93
European American	High Need Rural	34.03	34.24	34.04
European American	Average	35.81	37.08	35.71
European American	Low Need	42.75	45.22	42.52
European American	All	37.85	39.09	37.75
Hispanic American	New York City	50.37	48.00	50.62
Hispanic American	Big Four	50.58	47.94	50.85
Hispanic American	High Need US	46.27	44.13	46.50
Hispanic American	High Need Rural	42.06	40.15	42.26
Hispanic American	Average	39.30	38.07	39.44
Hispanic American	Low Need	35.73	35.78	35.75
Hispanic American	All	48.82	46.59	49.05
Other	New York City	47.12	48.08	47.04
Other	Big Four	39.74	39.12	39.82
Other	High Need US	40.70	40.47	40.74
Other	High Need Rural	39.47	38.93	39.54
Other	Average	42.12	43.47	42.01
Other	Low Need	51.09	53.78	50.83
Other	All	46.30	47.51	46.19
All	New York City	50.12	48.47	50.30
All	Big Four	46.06	43.92	46.21
All	High Need US	41.68	40.69	41.79
All	High Need Rural	34.62	34.70	34.64
All	Average	36.49	37.48	36.41
All	Low Need	43.03	45.39	42.81
All	All	42.96	42.81	42.99

Closing the Gap

Note from Table 3 that, with little exception, within each group the 1999 means are ordered with students from New York City and the Big Four scoring the lowest, followed by high need urban/suburban, high-need rural, average, and low-need students. The two exceptions to this are: (1) other students, among whom New York City students outscored Big Four, high-need rural and high-need urban/suburban students; and (2) European American students, among whom New York City students also outscored these other three groups. These discrepancies may well be related to the variation of socioeconomic status among groups from community type to community type.

The findings indicate that there has been substantial progress in closing the scoring disparities over these community types on the ELA-4 test, and somewhat mixed results on the other examinations.

ELA-4. The following findings with regard to ELA-4 illustrate the extent to which the gaps among population groups are closing:

1. Within each ethnic group, the mean for New York City students in 2001 was greater than the mean for students in the Big Four or in high-need urban/suburban districts in 1999;

2. Within each ethnic group, the 2001 mean for Big Four students was greater than the mean for students in high-need urban/suburban districts in 1999;

3. Within each ethnic group, the 2001 mean for students from the high-need urban/suburban districts was greater than the 1999 mean for either students from the high-need rural or average districts in 1999;

4. Within each ethnic group, the 2001 mean for students from the average districts was greater than the 1999 mean for students from the low-need (affluent) districts;

5. Within each ethnic group, the 2001 mean for students from the low-need districts was greater than the means for students from any of the other districts.

For all students except those in the other classification, the 2001 means for the high-need rural school districts surpassed the 1999 means for students from the average school districts. Thus, in general, the results show that within a two-year period, the socioeconomic disparities evident after four and a half years of school are closing.

Perhaps most pertinent to the issue of closing the performance gap are the changes in the group means from year to year. Tables 5a through 5d present these changes for each year's 24 scale score means among the four ethnic groups within the six needs resource categories. These mean differences means are given in three ways: changes in the 2000 mean from the 1999 mean, changes in the 2001 mean from the 1999 mean, and changes in the 2001 mean from the 1999 mean.

It is interesting that the students in the other classification from the Big Four and high-need urban/suburban districts grew most. These students were followed by African American students from the low-need districts and African American students from the high-need urban/suburban districts. These two African American groups made the largest gains from 1999 to 2000 and from 2000 to 2001. These gains for groups that scored so low in 1999 constitute the very definition of closing the gap between performance and higher expectation.

Table 6a shows the mean standardized differences by ethnicity and needs resource category for each ethnic group, matched with the European American group within needs resource category and year of testing. The mean differences are standardized with reference to the standard deviation for the appropriate European American group. In other words, these differences can be viewed as progress toward the mean for matched European American children. In particular, positive changes in the standardized mean diffrences over the two-year period (last column) are an indication of closing the gap. As shown, the gains are positive for New York City students of the other ethnicity group, for New York City African American students, and for the rural students of the other

ethnicity group classification. Overall, students in the other ethnic categories (-0.10) came closest to the mean scores for matched European American students, followed by African American students (-0.68) and Hispanic American students (-0.72). The lowest standardized mean differences between ethnic groups was for students in the low need districts (-0.22), followed by students in average districts (-0.35), in rural districts (-0.36), the Big Four districts (-0.45), high need urban/suburban districts (-0.46), and New York City (-0.70).

ELA-8. The pattern of students in poorer districts surpassing the initial performance of students in more affluent districts over time was not evident on this examination. The pattern of initial large loss in 2000 followed by partial recovery in 2001 held for students within each ethnic classification, and overall for students in New York City, the urban/suburban districts, the rural districts, and the average districts. Students in the Big Four districts dropped each consecutive year, while those in the low need districts rose each year.

Standardized mean differences were computed by comparing students within each ethnic group to European American students in the same needs resource categories and years (averaged differences between individual student scores and means for European American students divided by the standard deviation for the European American students). These means are given in Table 6b, and reveal overall that students in low need districts had the lowest overall discrepancies after three years (-0.24), followed by students is: rural (-0.36), average (-0.39), Big Four (-0.45), high need urban/suburban (-0.54), and New York City (-0.71) districts. The smallest differences were for students in the other ethnic categories (-0.04), followed by African American students (-0.72) and Hispanic American students (-0.73).

Math 4. On the Math-4 examination, the initial standardized differences between European American and other students decreased each year, from –0.68 in 1999 to –0.65

in 2000 to –0.64 in 2001. Students of the other ethnic category scored higher than European American students (0.06), while Hispanic American (-0.75) and African American (-0.78) scored lower. The students in the other ethnicity group scored higher than European American students in 2000 and in 2001 in all but the rural schools, and higher in all three years in the average, low need, and Big Four districts.

In order, the smallest mean scoring differences were for students in: low need (-0.21), rural (-.045), average (-0.46), high need urban/suburban (-0.52), Big Four (-0.56), and New York City (-0.74) school districts. Table 6c contains the standardized mean differences.

Math-8. The smallest standardized mean difference with matched groups of European American students was on the 1999 administration (-0.64). The largest standardized difference was in 2000 (-0.70), with the 2001 results (-0.66) between 1999 and 2000. Again, students of the other ethnic category scored higher than European American students (0.14), while Hispanic American (-0.78) and African American (-0.85) scored lower. The smallest standardized differences were for students in the low need districts (-0.22), followed by students in the rural (-0.51), average (-0.55), Big Four (-0.63), high need urban/suburban (-.070), and New York City (-0.72). Table 6d shows all of the standardized mean differences.

Table 5a

Improvements in Rankings of Means on the ELA-4, 1999-2001
by Needs/Resource Category and Ethnicity

		1999-2001		2000-2001		1999-2000	
RESOURCE	ETHNIC	Mean Change	Rank	Mean Change	Rank	Mean Change	Rank
High Need Rural	African American	5.63	1	-1.92	3	7.54	3
High Need Rural	Other	6.93	2	-5.04	1	11.97	12
Big Four	African American	7.23	3	0.87	8	6.36	1
New York City	African American	8.09	4	0.89	9	7.20	2
High Need Rural	European American	9.66	5	-0.58	4	10.24	6
New York City	Hispanic American	9.67	6	2.11	15	7.56	4
Big Four	European American	9.72	7	1.69	13	8.03	5
Average	African American	10.67	8	-0.07	5	10.74	8
High Need US	European American	13.36	9	1.24	11	12.13	13
Average	European American	13.53	10	0.56	6	12.97	15
Big Four	Hispanic American	13.70	11	3.13	19	10.58	7
Average	Hispanic American	13.76	12	2.02	14	11.74	11
New York City	Other	13.79	13	1.42	12	12.36	14
High Need Rural	Hispanic American	13.98	14	2.91	17	11.06	9
Average	Other	14.75	15	0.61	7	14.14	18
New York City	European American	15.83	16	2.14	16	13.69	17
High Need US	African American	16.03	17	4.69	21	11.34	10
Low Need	African American	16.12	18	0.98	10	15.15	20
High Need US	Hispanic American	16.27	19	3.01	18	13.26	16
Low Need	Hispanic American	19.01	20	4.73	23	14.28	19
Low Need	European American	20.96	21	3.51	20	17.44	21
High Need US	Other	21.63	22	-2.86	2	24.49	24
Big Four	Other	25.25	23	5.06	24	20.19	22
Low Need	Other	27.57	24	4.72	22	22.85	23

Table 5b

Improvements in Rankings of Means on the ELA-8, 1999-2001
by Needs/Resource Category and Ethnicity

		1999-2001		1999-2000		2000-2001	
RESOURCE	ETHNIC	Mean Change	Rank Change	Mean Change	Rank Change	Mean Change	Rank Change
Big Four	Other	-8.52	1	-10.09	1	1.57	21
Big Four	African American	-5.68	2	1.00	14	-6.69	1
Big Four	European American	-4.70	3	-0.87	5	-3.83	6
High Need Rural	Other	-4.65	4	-2.65	3	-2.00	15
High Need US	African American	-4.59	5	0.19	7	-4.78	4
New York City	African American	-4.10	6	1.42	19	-5.56	3
High Need Rural	Hispanic American	-3.95	7	-2.78	2	-1.17	18
New York City	Hispanic American	-3.86	8	1.80	21	-5.66	2
Big Four	Hispanic American	-3.46	9	-0.35	6	-3.10	9
High Need Rural	European American	-2.46	10	0.71	12	-3.16	8
Average	African American	-2.29	11	1.89	22	-4.18	5
Low Need	African American	-2.20	12	0.73	13	-2.93	10
Average	Hispanic American	-1.86	13	1.71	20	-3.57	7
High Need US	Hispanic American	-1.64	14	0.70	11	-2.34	13
High Need US	European American	-1.47	15	1.34	17	-2.81	11
New York City	Other	-1.12	16	1.03	16	-2.15	14
New York City	European American	-0.99	17	0.55	8	-1.54	16
Average	European American	-0.56	18	0.66	10	-1.22	17
Low Need	Hispanic American	0.25	19	1.00	15	-0.76	19
High Need Rural	African American	0.32	20	2.87	24	-2.55	12
High Need US	Other	0.73	21	-234	4	3.06	23
Low Need	European American	2.61	22	1.38	18	1.27	20
Average	Other	5.25	23	2.46	23	2.78	22
Low Need	Other	5.92	24	0.57	9	5.35	24

Table 5c

Improvements in Rankings of Means on the Math-4, 1999-2001
by Needs/Resource Category and Ethnicity

		1999-2001		1999-2000		2000-2001	
RESOURCE	ETHNIC	Mean Change	Rank Change	Mean Change	Rank Change	Mean Change	Rank Change
Big Four	European American	2.01	1	9.02	22	-7.01	1
New York City	European American	2.10	2	7.63	20	-5.53	3
Big Four	African American	2.22	3	4.67	2	-2.41	11
New York City	African American	2.65	4	5.51	7	-2.86	8
High Need Rural	Hispanic American	2.72	5	4.10	1	-1.38	15
High Need Rural	European American	2.86	6	5.44	6	-2.58	9
Average	Hispanic American	3.29	7	5.84	11	-2.55	10
Average	European American	3.32	8	7.60	18	-4.28	6
Low Need	Hispanic American	3.47	9	6.82	15	-3.36	7
Average	Other	3.49	10	4.76	4	-1.27	16
Average	African American	3.64	11	5.59	8	-1.95	14
Low Need	European American	3.69	12	8.95	21	-5.26	4
New York City	Hispanic American	4.10	13	6.48	13	-2.37	12
High Need Rural	African American	4.37	14	4.74	3	-0.41	20
New York City	Other	4.75	15	5.92	12	-1.17	17
High Need US	European American	4.95	16	7.23	17	-2.28	13
High Need Rural	Other	5.50	17	5.65	9	-0.15	21
Big Four	Other	5.59	18	9.92	23	-4.33	5
Big Four	Hispanic American	5.83	19	6.89	16	-1.06	18
Low Need	African American	6.15	20	6.79	14	-0.64	19
High Need US	Hispanic American	6.34	21	5.34	5	1.00	23
High Need US	African American	8.50	22	7.63	19	0.87	22
Low Need	Other	11.70	23	17.38	24	-5.68	2
High Need US	Other	13.89	24	5.76	10	8.12	24

Table 5d

Improvements in Rankings of Means on the Math 8, 1999-2001 by Needs/Resource Category and Ethnicity

		1999-2001		1999-2000		2000-2001	
RESOURCE	ETHNIC	Mean Change	Rank Change	Mean Change	Rank Change	Mean Change	Rank Change
Big Four	African American	-3.62	-4	-3.84	4	0.23	2
Big Four	European American	-1.19	-2	-4.17	2	2.98	8
New York City	African American	-0.22	-2	-2.78	6	2.56	7
High Need Rural	Hispanic American	0.34	-2	-5.26	1	5.60	15
Big Four	Hispanic American	0.35	-1	-2.14	7	2.49	6
New York City	Hispanic American	1.55	-1	0.13	16	1.42	4
New York City	European American	1.67	-1	0.10	15	1.57	5
Big Four	Other	2.65	-1	3.97	23	-1.32	1
High Need Rural	European American	2.84	-1	-0.82	10	3.66	12
New York City	Other	3.16	-1	1.79	22	1.38	3
Average	European American	3.16	0	-0.89	9	4.06	13
Low Need	European American	3.22	0	0.02	13	3.20	11
High Need US	European American	3.38	0	0.19	17	3.19	10
Low Need	African American	3.50	0	0.33	18	3.17	9
High Need US	African American	3.59	0	-3.79	5	7.39	18
Low Need	Hispanic American	4.74	0	-0.70	11	5.44	14
Average	African American	5.67	0	-2.04	8	7.71	19
High Need Rural	African American	6.39	1	-3.92	3	10.31	23
High Need US	Hispanic American	6.64	1	-0.15	12	6.79	17
Average	Hispanic American	6.83	2	0.77	20	6.06	16
High Need Rural	Other	8.62	2	0.86	21	7.76	20
High Need US	Other	9.59	3	0.44	19	9.15	22
Low Need	Other	11.48	3	0.07	14	11.40	24
Average	Other	14.89	4	5.94	24	8.95	21

Table 6a

Standardized Growth by Ethnicity and Community Type
For ELA-4 1999-2001
Criterion Group = Low Need European American

Needs/Resource	Ethnicity	Stand –1999 Mean Diff.	Stand - 2000 Mean Diff.	Stand - 2001 Mean Diff.	Two Year Changes in Diff.
High Need Rural	Other	-0.83	-0.80	-1.00	-0.17
High Need Rural	European American	-0.60	-0.67	-0.75	-0.15
Average	Other	-0.21	-0.26	-0.33	-0.12
Big Four	European American	-0.72	-0.83	-0.84	-0.12
Average	European American	-0.31	-0.37	-0.43	-0.12
High Need Rural	African American	-1.17	-1.19	-1.28	-0.11
High Need US	European American	-0.60	-0.62	-0.65	-0.05
New York City	Other	-0.61	-0.62	-0.65	-0.04
New York City	European American	-0.43	-0.44	-0.46	-0.03
Average	African American	-0.96	-0.94	-0.99	-0.03
Big Four	African American	-1.45	-1.45	-1.45	0.00
Average	Hispanic American	-0.81	-0.80	-0.80	0.01
New York City	African American	-1.44	-1.42	-1.42	0.02
Low Need	African American	-0.60	-0.53	-0.58	0.02
High Need Rural	Hispanic American	-1.03	-0.99	-0.95	0.08
Low Need	Hispanic American	-0.55	-0.52	-0.46	0.09
New York City	Hispanic American	-1.53	-1.48	-1.44	0.09
Big Four	Hispanic American	-1.24	-1.16	-1.11	0.13
High Need US	African American	-1.27	-1.17	-1.08	0.19
High Need US	Hispanic American	-1.23	-1.07	-1.04	0.19
Low Need	Other	-0.06	0.11	0.14	0.20
High Need US	Other	-0.86	-0.47	-0.62	0.24
Big Four	Other	-1.17	-0.83	-0.75	0.42

Table 6b

Standardized Growth by Ethnicity and Community Type
For ELA-8 1999-2001
Criterion Group = Low Need Others

Needs/Resource	Ethnicity	Stand – 1999 Mean Diff.	Stand - 2000 Mean Diff.	Stand - 2001 Mean Diff.	Two Year Changes in Diff.
Big Four	Other	-0.70	-0.75	-1.06	-0.36
Big Four	European American	-0.67	-0.89	-0.92	-0.25
High Need Rural	Other	-0.77	-0.92	-1.01	-0.24
Big Four	African American	-1.19	-1.44	-1.41	-0.22
High Need US	African American	-1.18	-1.37	-1.37	-0.19
High Need Rural	European American	-0.61	-0.81	-0.80	-0.19
High Need Rural	Hispanic American	-1.03	-1.13	-1.22	-0.19
Low Need	African American	-0.61	-0.81	-0.79	-0.18
New York City	European American	-0.30	-0.48	-0.48	-0.18
New York City	African American	-1.15	-1.37	-1.32	-0.17
New York City	Hispanic American	-1.15	-1.38	-1.32	-0.17
New York City	Other	-0.37	-0.56	-0.54	-0.17
Big Four	Hispanic American	-1.19	-1.33	-1.35	-016
High Need US	European American	-0.59	-0.78	-0.75	-0.16
Average	European American	-0.39	-0.55	-0.54	-0.15
Average	African American	-0.95	-1.15	-1.09	-0.14
Average	Hispanic American	-0.83	-1.02	-0.97	-0.14
High Need US	Hispanic American	-1.11	-1.23	-1.22	-0.11
Low Need	European American	-0.07	-0.19	-0.17	-0.10
Low Need	Hispanic American	-0.61	-0.73	-0.71	-0.10
High Need US	Other	-0.74	-0.73	-0.81	-0.07
High Need Rural	African American	-1.16	-1.29	-1.20	-0.04
Average	Other	-0.35	-0.40	-0.34	0.01

Table 6c

Standardized Growth by Ethnicity and Community Type
For Math-4 1999-2001
Criterion Group = Low Need Others

Needs/Resource	Ethnicity	Stand - 1999 Mean Diff.	Stand - 2000 Mean Diff.	Stand - 2001 Mean Diff.	Two Year Changes in Diff.
Low Need	European American	-0.14	-0.14	-0.34	-0.20
New York City	European American	-0.63	-0.65	-0.83	-0.20
Big Four	European American	-0.88	-0.95	-1.06	-0.18
Average	Other	-0.45	-0.34	-0.62	-0.17
Average	European American	-0.54	-0.52	-0.71	-0.17
High Need Rural	European American	-0.82	-0.76	-0.98	-0.16
Low Need	Hispanic American	-0.72	-0.68	-0.88	-0.16
High Need Rural	Hispanic American	-1.18	-1.11	-1.33	-0.15
Average	Hispanic American	-1.00	-0.95	-1.14	-0.14
New York City	Other	-0.67	-0.56	-0.80	-0.13
Big Four	African American	-1.54	-1.51	-1.67	-0.13
Average	African American	-1.15	-1.08	-1.27	-0.12
High Need US	European American	-0.80	-0.73	-0.91	-0.11
New York City	African American	-1.61	-1.59	-1.72	-0.11
Big Four	Other	-0.84	-0.83	-0.93	-0.09
High Need Rural	African American	-1.43	-1.34	-1.52	-0.09
Low Need	African American	-0.80	-0.69	-0.89	-0.09
High Need Rural	Other	-1.00	-0.88	-1.08	-0.08
New York City	Hispanic American	-1.56	-1.53	-1.64	-0.08
Big Four	Hispanic American	-1.48	-1.41	-1.53	-0.05
High Need US	Hispanic American	-1.32	-1.18	-1.36	-0.04
High Need US	African American	-1.38	-1.25	-1.36	0.02
High Need US	Other	-0.92	-0.57	-0.80	0.12

Table 6d

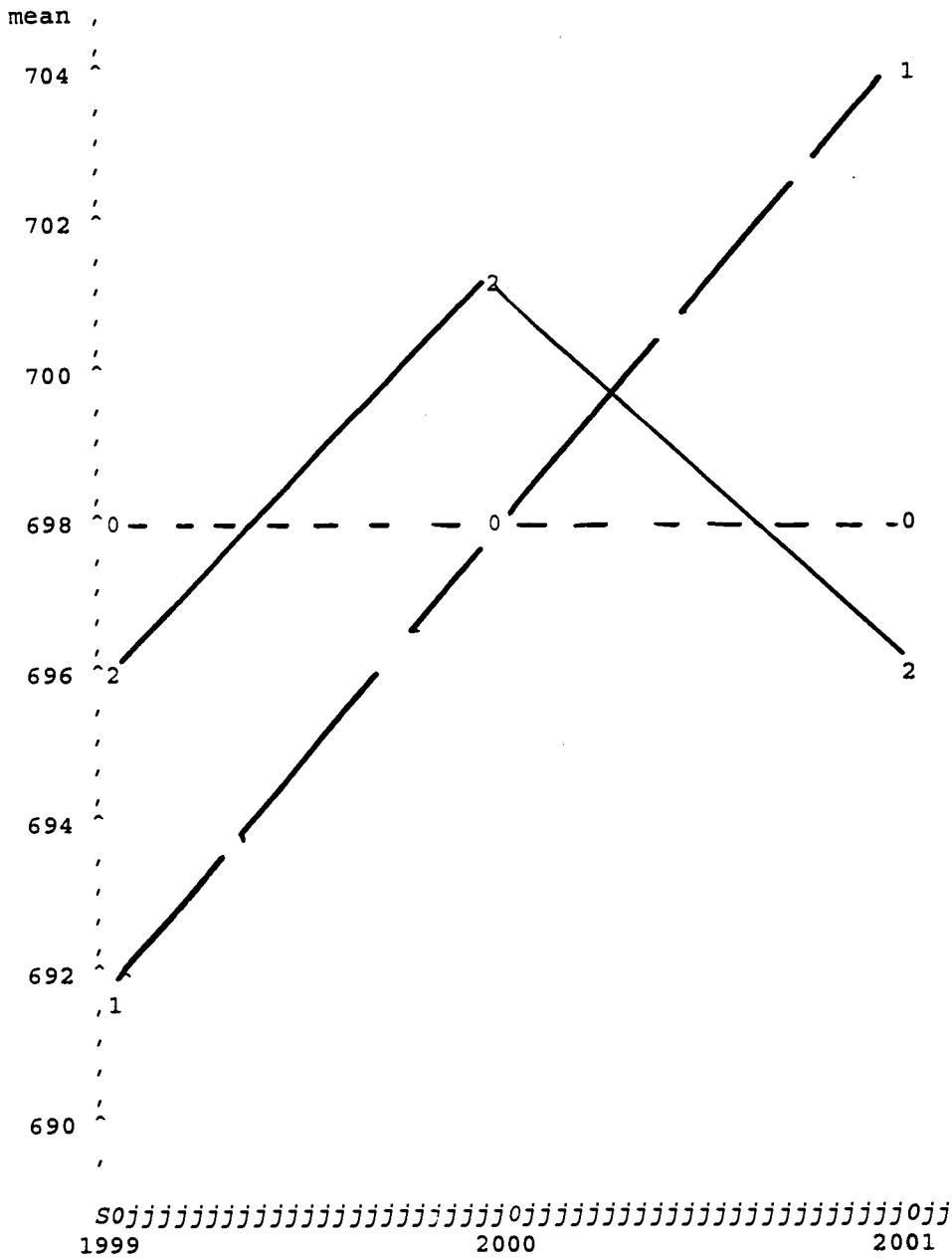
Standardized Growth by Ethnicity and Community Type
For Math - 8 1999-2001
Criterion Group = Low Need Others

Needs/Resource	Ethnicity	Stand - 1999 Mean Diff.	Stand - 2000 Mean Diff.	Stand - 2001 Mean Diff.	Two Year Changes in Diff.
Big Four	African American	-1.53	-1.98	-1.98	-0.45
Big Four	European American	-0.95	-1.27	-1.31	-.036
New York City	African American	-1.60	-1.99	-1.96	-0.36
Big Four	Hispanic American	-1.60	-1.99	-1.95	-0.35
High Need Rural	Hispanic American	-1.31	-1.59	-1.64	-0.33
New York City	Hispanic American	-1.52	-1.93	-1.83	-0.31
New York City	European American	-0.66	-1.00	-0.94	-0.28
High Need US	African American	-1.55	-1.81	-1.81	-0.26
Big Four	Other	-0.92	-1.36	-1.18	-0.26
High Need Rural	European American	-0.77	-1.05	-1.02	-0.25
New York City	Other	-0.51	-0.83	-0.75	-0.24
High Need US	European American	-0.77	-1.07	-1.01	-0.24
Low Need	African American	-0.89	-1.20	-1.13	-0.24
Average	European American	-0.53	-0.78	-0.76	-0.23
Low Need	European American	-0.13	-0.37	-0.35	-0.22
Low Need	Hispanic American	-0.84	-1.08	-1.05	-0.21
Average	African American	-1.31	-1.54	-1.51	-0.20
High Need Rural	African American	-1.55	-1.73	-1.74	-0.19
High Need US	Hispanic American	-1.48	-1.75	-1.66	-0.18
Average	Hispanic American	-1.18	-1.43	-1.34	-0.16
High Need Rural	Other	-1.00	-1.20	-1.11	-0.11
High Need US	Other	-0.91	-1.06	-0.99	-0.08
Average	Other	-0.55	-0.67	-0.48	0.07

Finally, the performance of the various ethnic groups in the low-need (affluent) school districts is interesting. African American students, students in the other classification, and Hispanic American students from these districts placed among the top ten groups in change in mean scores. The 2001 mean of even the highest scoring group, students in the other classification (686.23), from these school districts is not near the test ceiling of 800, and only 88 percent of this group have reached level 3 or higher of the test. The difference in change in relative position, therefore, cannot be easily ascribed to an asymptotic ceiling effect that would restrict the growth for the European American or students in the other classification in these districts or anywhere else in the state. Clearly, this growth indicates that the implementation of the learning standards and the associated assessments have had a profound impact on minority students in these affluent districts.

Figure 1

ELA-4 1999 - 2001
Linear ("1") and Quadratic ("2")
Coefficients



NOTE: 1 obs hidden.

Figure 2

ELA-8 1999 - 2001
Linear ("1") and Quadratic ("2")
Coefficients

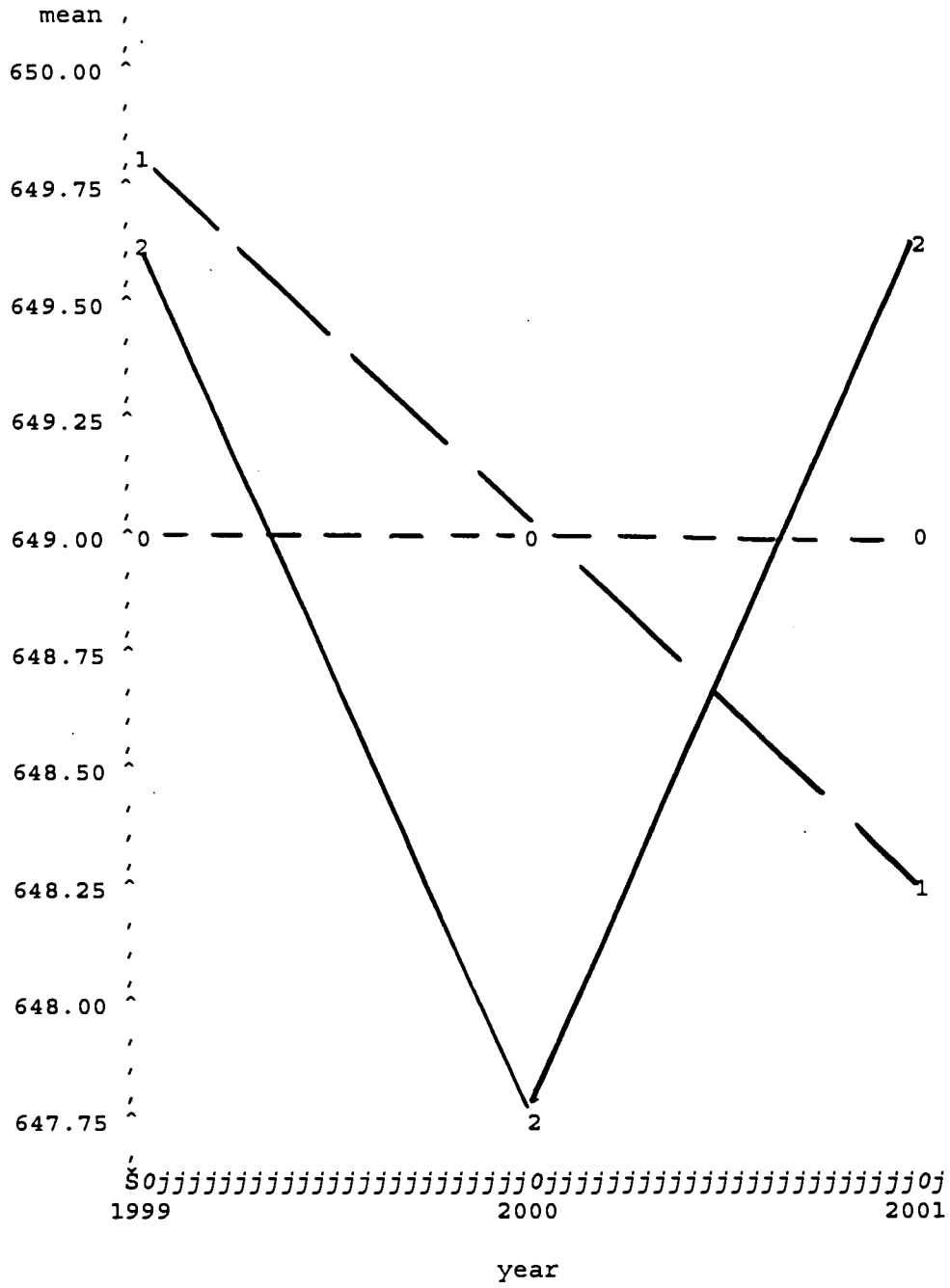


Figure 3

Math-4 1999 - 2001
Linear ("1") and Quadratic ("2")
Coefficients

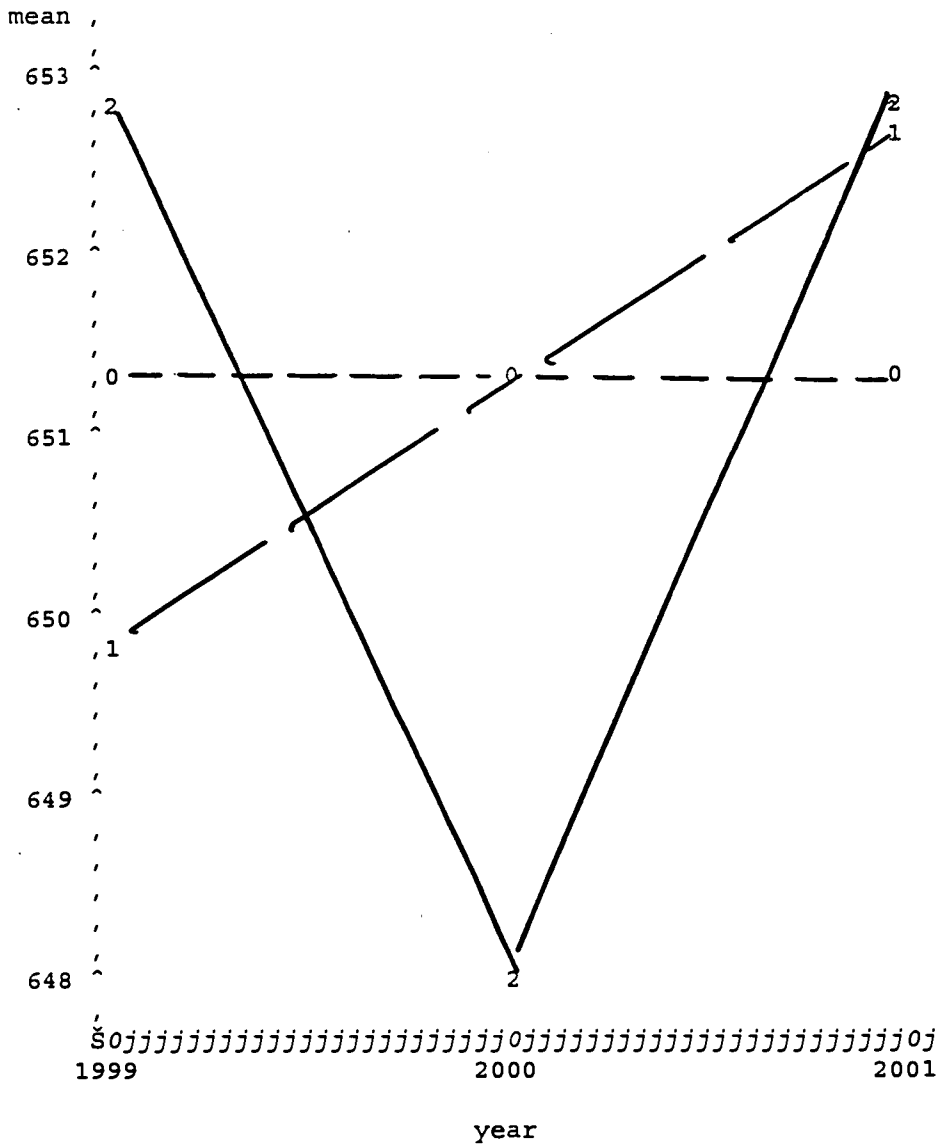
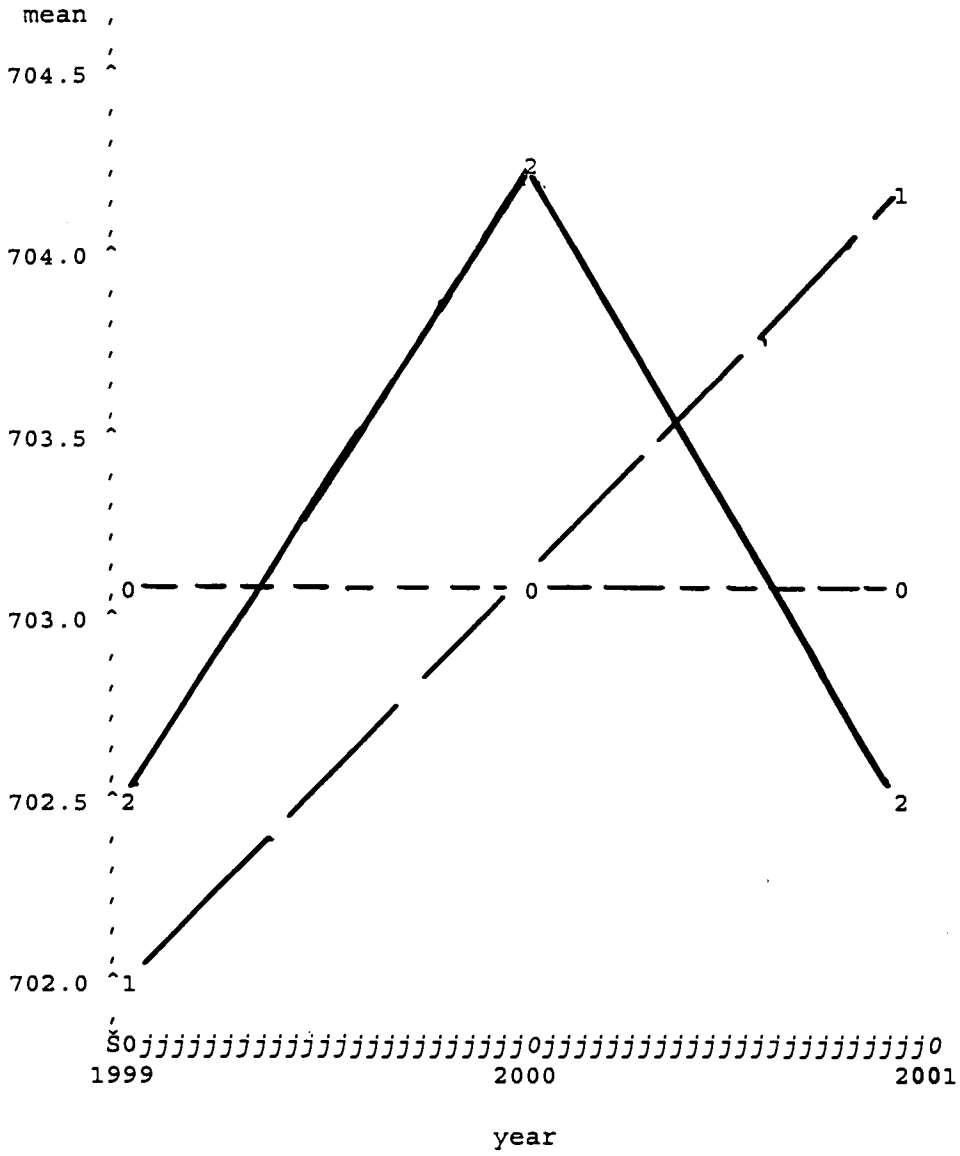


Figure 4

Math-8 1999 - 2001
Linear ("1") and Quadratic ("2")
Coefficients



NOTE: 1 obs hidden.

Discussion

Conclusions

The imperative of equity of expectation and equity of access should have its most profound effects on student populations that have traditionally been disenfranchised. If achievement gaps are to close, there should be large movement in these populations in response to the growing demand for excellence. Three years of data suggest that there is progress in closing these gaps, but that progress is mixed, both by year and by test. It remains for us to make a best judgment about the story these data reveal.

If we examine the rankings of mean gains shown in Tables 6a-6d, a pattern begins to unfold. For example, students in the other category of ethnic groupings made consistently high gains, especially those in the low need, urban/suburban, and average districts. Large gains were also consistently made for African American and for Hispanic American students, across tests, for students in low need and urban/suburban districts. A simple sum of the ranks reveals that these four groups (low need and urban/suburban African American and Hispanic American students) were among the ten groups that made the largest gains in the two year period across all four tests. Clearly, these groups are examples of closing the gap.

References

Good, T. L. & Brophy, J. E. (1990). *Educational Psychology: A Realistic Approach.* White Plains, N. Y.: Longman

Messick, S. (1981). "Evidence and ethic in the evaluation of tests." Invited address to the Division of Measurement and Research Methodology of the American Educational Research Association, Los Angeles, April 1981 (Reprinted at Princeton, New Jersey: Educational Testing Service).

Myers, J. L. (1973). *Fundamentals of Experimental Design* (Second Edition). Boston: Allyn and Bacon, Inc.

Appendix A

General Linear Model Regression:

Year by Ethnicity by Needs resource Category

On the Grade 4 English Language Arts

Assessment, 1999-2001 (R-Square = 0.196)

Source	Sum of Squares	Degrees of Freedom	Mean Square	F-Ratio
Year	3,743,596.77	2	1,871,798.38	1,391.11*
Ethnic	11,528,504.99	3	3,842,835.00	2,855.98*
Resource	12,206,084.36	5	2,441,216.87	1,814.30*
Yr. × Eth.	139,265.97	6	23,211.00	17.25*
Yr. × Res.	304,393.41	10	30,439.41	22.62*
Eth. × Res.	4,921,277.21	15	328,085.15	243.83*
Yr. × Eth. × Res.	289,282.64	30	9,642.75	7.17*
Error	863,922,762.14	642,063	1,345.54	
Total	1,074,390,914.02	642,134		
Year Quantitative Contrasts:				
Linear	2,931,244.13	1	2,931,244.13	2,178.49*
Quadratic	537,358.27	1	537,358.27	399.36*

*Exceeds the p<.001 level of significance.

Appendix A

General Linear Model Regression:

Year by Ethnicity by Resource Needs Category

On the Grade 8 English Language Arts

Assessment, 1999-2001 (R-Square = 0.150)

Source	Sum of Squares	Degrees of Freedom	Mean Square	F-Ratio
Year	53,881.21	2	26,940.60	29.03*
Ethnic	9,089,970.42	3	3,029,990.14*	3,264.89*
Resource	4,470,009.93	5	894,801.99	964.17*
Yr. × Eth.	41,435.59	6	6,905.93	7.44*
Yr. × Res.	117,259.92	10	11,725.99	12.64*
Eth. × Res.	1,957,850.32	15	130,523.62	140.64*
Yr. × Eth. × Res.	86,368.00	30	2,878.93	3.10*
Error	536,177,214.90	577,745	928.10	
Total	635,342,723.50	577,816		
Year Quantitative Contrasts:				
Linear	39,274.15	1	39,274.15	42.32*
Quadratic	15,810.83	1	15,810.83	17.04*

*Exceeds the $p<.001$ level of significance.

Appendix A

General Linear Model Regression:

Year by Ethnicity by Resource Needs Category

On the Grade 4 Mathematics

Assessment, 1999-2001 (R-Square = 0.419)

Source	Sum of Squares	Degrees of Freedom	Mean Square	F-Ratio
Year	32,340,675.93	2	16,170,337.98	12,816.80*
Ethnic	17,293,638.07	3	5,764,546.02	4,569.03*
Resource	11,963,541.97	5	2,392,708.39	1,896.48*
Yr. × Eth.	411,762.82	6	68,627.14	54.39*
Yr. × Res.	557,560.00	10	55,756.00	44.19*
Eth. × Res.	2,893,893.59	15	192,926.24	152.92*
Yr. × Eth. × Res.	435,113.85	30	14,503.79	11.50*
Error	791,137,441.00	627,063	1,262.00	
Total	1,361,276,466.00	627,134		
Year Quantitative Contrasts:				
Linear	20,361,517.34	1	20,361,517.34	16,138*
Quadratic	12,490,219.90	1	12,490,219.90	9,899.87*

*Exceeds the $p<.001$ level of significance.

Appendix A

General Linear Model Regression:

Year by Ethnicity by Resource Needs Category

On the Grade 8 Mathematics

Assessment, 1999-2001 (R-Square = 0.238)

Source	Sum of Squares	Degrees of Freedom	Mean Square	F-Ratio
Year	323,728.08	2	161,864.04	115.87*
Ethnic	22,308,521.89	3	7,436,173.96	5,323.24*
Resource	13,522,844.40	5	2,704,568.88	1,936.09*
Yr. × Eth.	82,691.62	6	13,781.94	9.87*
Yr. × Res.	321,903.69	10	32,190.37	23.04*
Eth. × Res.	2,012,628.75	15	134,175.25	96.05*
Yr. × Eth. × Res.	153,673.78	30	5,122.46	3.67*
Error	824,029,220.00	589,881	1,397.00	
Total	1,080,924,027.00	589,952		
Year Quantitative Contrasts:				
Linear	203,836.79	1	203,836.79	145.92*
Quadratic	120,093.81	1	120,093.81	85.97*

*Exceeds the p<.001 level of significance.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

REPRODUCTION RELEASE

(Specific Document)

TM035027

I. DOCUMENT IDENTIFICATION:

Title: Differential Consequential Validity and the Stability of Inferences across Ethnicity and Community On NYS Large Scale Tests

Author(s): Gerald E. DeMauro, Ph.D.

Corporate Source: NYS Education Dept.

Publication Date: March 2002

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

Level 1

☒

Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.

The sample sticker shown below will be affixed to all Level 2A documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2A

Level 2A

☐

Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only

The sample sticker shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2B

Level 2B

☐

Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, please

Signature: Gerald E. DeMauro

Printed Name/Position/Title: Gerald E. DeMauro

Organization/Address: NYSED
89 Washington Ave., 775 EBA, Albany, NY 12234

Telephone: 518-474-5902

FAX: 518-474-1989

E-Mail Address: gdemauro@mail.nysed.gov

Date: 6/4/03

(Over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

ERIC CLEARINGHOUSE ON ASSESSMENT AND EVALUATION
UNIVERSITY OF MARYLAND
1129 SHRIVER LAB
COLLEGE PARK, MD 20742-5701
ATTN: ACQUISITIONS

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
4483-A Forbes Boulevard
Lanham, Maryland 20706

Telephone: 301-552-4200
Toll Free: 800-799-3742
FAX: 301-552-4700
e-mail: ericfac@inet.ed.gov
WWW: http://ericfacility.org

(Rev. 2/2001)