Education Policy Analysis Archives

Exploring the Achievement Gap Between White and Minority Students in Texas: A Comparison of the 1996 and 2000 NAEP and TAAS Eighth Grade Mathematics Test Results

Thomas H. Linton
Texas A&M University-Corpus Christi

Donald Kester
Texas A&M University-Corpus Christi

Abstract

The Texas Assessment of Academic Skills (TAAS) has been used to document and track an achievement gap between white and minority students in Texas. Some educators have credited the TAAS with fueling

a drive to close the achievement gap while others suggest that TAAS scores may be misleading because of factors such as score inflation and a possible ceiling effect. The purpose of this study was to analyze the gap in mathematics achievement for eighth grade students. The study compared TAAS and National Assessment of Educational Progress (NAEP) test results to determine if the achievement gap between white, Hispanic, and African-American Students had narrowed between 1996 and 2000. Results indicate that TAAS mean scores increased significantly for all three ethnic groups between 1996 and 2000. Comparison of the TAAS test score frequency distributions for each ethnic group indicated that white students' scores shifted from the middle to the upper portion of the test score range while minority students' scores shifted from the lower to the middle and higher score range. Both white and minority students' TAAS test score distributions were significantly more negatively skewed in 2000 than in 1996. Comparisons between white and minority students' TAAS scores showed that white students had significantly higher scores than either Hispanic or African-American students in both 1996 and 2000. Comparison of mean score differences in 1996 and 2000 indicated that the achievement gap between white and minority students had narrowed. NAEP scores increased significantly from 1996 to 2000 for Hispanic students, but not for white or African-American students. However, test score distribution patterns showed small positive changes for all three ethnic groups. Comparisons between ethnic groups indicated that there were significant differences between white and minority students' scores in both 1996 and 2000. Comparison of mean score differences in 1996 and 2000 indicated that the achievement gap between Hispanic white students had narrowed slightly but that there was no change in the achievement gap between white and African-American students. Analysis of the TAAS test score distribution patterns indicated the likelihood that a ceiling effect had impacted students' scores. The evidence for a ceiling effect was strongest for white students. In 2000, 60.4% of white students had a TAAS score that fell in the top 10% of the score range. In contrast, there was no evidence of a ceiling effect for the NAEP. Mean score gains on the TAAS are only partially substantiated by the NAEP data. Furthermore, there is a very strong possibility that a ceiling effect artificially restricted the 2000 TAAS scores for white students and created the illusion that the achievement gap between minority and white students had been narrowed.

The Texas Assessment of Academic Skills (TAAS) has been administered as a measure of student achievement in reading, mathematics, and writing in Texas since 1990. The tests have been praised for providing disaggregated data on different ethnic groups and requiring each group to meet the same standard of proficiency. The disaggregated scores have been used to document and track an achievement gap between the scores of white and minority students. Test results have typically been reported as percent of students passing (i.e. meeting minimum expectations) the TAAS (Texas Education Agency, 2000 b). These results show that passing rates for minority students have increased at a faster rate than the passing rates for white students.

Table 1
TAAS Grade 8 Mathematics Test: A Comparison of Percent Passing
for 1996 and 2000 by Ethnic Group

Ethnic Group	Percent Passing		Percent Gain
	1996	2000	
African American	43.8	80.9	37.1
Hispanic	51.4	85.5	34.1
White	77.7	94.8	17.1

Many educators in Texas have credited the tests as fueling the drive to close this achievement gap. Several studies have cited the increased percent of minority students passing the TAAS as evidence that the achievement gap is being closed (Hurley, Chamberlain, Slavin, & Madden, 2001; Jerald, 2000; Jerald, 2001; Texas Education Agency, 2000 a). Other researchers disagree and suggest that increases in passing rates for minority students may be linked to other factors. Haney (2000) argues that increases in the TAAS passing rate for minority high school students is at least partially explained by higher dropout rates among minority students. Toenjes and Dworkin (2002) cite evidence that challenges Haney's assertions and conclude that dropout rates and special education exemptions do not explain the large increases in the TAAS passing rate.

It should be noted that none of these studies directly address the actual achievement level of students. Comparing the passing rate for minority and white students does not provide information about comparative achievement levels. Rather, the passing rate is simply the percentage of students that have attained the minimum achievement level deemed necessary to "pass" the test.

Rather than depending solely on TAAS data, some researchers have used National Assessment of Educational Progress (NAEP) results to analyze the achievement levels of Texas students. These studies have consistently found either no gain or small achievement gains for both minority and white students, but have been unable to substantiate the large gains reported on the TAAS. Amrein and Berliner, after examining test results for 18 states with high-stakes testing programs, reported that student achievement remained at the same level or went down after the high-stakes testing policies were instituted. Camilli, using a cohort analysis of gains on the NAEP math test from grade 4 to grade 8 found that Texas ranked 17th among 35 states. In a study that directly compared TAAS and NAEP test results (Klein, Hamilton, McCaffrey, & Stecher, 2000), TAAS data indicated that the achievement gap between minority and white students was closing while NAEP data did not. Klein, et. al. suggested that the large gains of minority students relative to white students on the TAAS were misleading and could be due, at least in part, to a ceiling effect and teaching to the test. However, Klein's study compared 1992 and 1996 NAEP scores with TAAS data for 1994 and 1998. Because of the disparity in the test administration dates for the NAEP and the TAAS, questions were raised about the conclusions of the study. Furthermore, Klein's study analyzed mean achievement gains for the TAAS and the NAEP but did not examine the actual distributions of test scores for evidence of a ceiling effect.

The National Center for Educational Statistics, in its report, *The Nation's Report Card: Mathematics 2000,* (U. S. Department of Education, 2001a) released national and state reports of NAEP scores for eighth grade mathematics in August 2001. To date, there have been no studies comparing the NAEP 2000 results and the TAAS results for Texas.

Purpose of the Study

The purpose of the present study is to present an analysis of the Texas TAAS and NAEP results for 1996 and 2000 and to explore the mathematics achievement gap between eighth grade minority students and white students in Texas. Specific research questions for the study are:

- Did Texas eighth grade African American, Hispanic, and white student math scores on the TAAS and NAEP increase significantly between 1996 and 2000?
- DO TAAS and NAEP data show that the achievement gap between white and minority students decreased from 1996 to 2000?
- Is there evidence that a ceiling effect artificially restricted the distribution of students' scores on the TAAS or the NAEP?

Methods

A causal-comparative research design was used to analyze the variables in the study. The sample for the TAAS data consisted of all African American, Hispanic, and white students in Texas who were in the Texas Education Agency's accountability subset of TAAS scores in grade 8 in 1996 and 2000 (TEA, 2000 b). (Note 1) The sample for the NAEP-Texas data consisted of a random sample of approximately 2,500 eighth grade students selected according to NAEP specifications in 1996 and 9,600 were selected in 2000. The national results for the NAEP test were included for comparative purposes. (Note 2)

Two sets of analyses were used to compare scores within and between ethnic groups. First, confidence intervals were calculated and used to analyze differences between mean scores. A *d* statistic (Green & Akey, 2000) was calculated to measure effect size for differences between mean scores. (Note 3) Second, a chi-square goodness-of-fit test was used to compare TAAS 1996 and 2000 test score distributions and NAEP 1996 and 2000 test score distributions for white and minority students. The Cramer's V coefficient (Green & Akey, 2000) was calculated to determine effect size for the chi-square analysis. (Note 4)

Research Question #1

Did Texas eighth grade African American, Hispanic, and white student math scores on the TAAS and NAEP increase significantly between 1996 and 2000?

The 1996 and 2000 TAAS scores and NAEP scores for each ethnic group were analyzed in two ways. First, mean scores for each ethnic group were compared to determine if there were significant differences between the 1996 results and the 2000 results. Second, the test score distributions for each ethnic group for 1996 and 2000 were compared to determine changes in the score distribution pattern over time.

TAAS Results

The comparison of 1996 and 2000 mean scores for each ethnic group on the TAAS is presented in Table 2. Results indicate that:

- Mean scores increased significantly for each ethnic group from 1996 to 2000.
- Effect sizes as measured by the d-statistic were moderate, ranging from .472 (white) to .646 (African American).
- Mean score gain for white students less than two-thirds as large as the gain for Hispanic and African American students.

Table 2
Comparison of TAAS TLI Mean Scores for 1996 and 2000 By Ethnic Group

| Group | Ethnicity | Mean TLI Scores | | Gain | Effect Size (d Statistic) |
		1996	2000		
TAAS	African American	65.0	77.2	+12.2[*]	.916
	Hispanic	67.8	79.3	+11.5[*]	.894
	White	77.2	84.2	+7.0[*]	.667

* p < .05

The second set of TAAS analyses compared the 1996 and 2000 score distributions for each ethnic group. A chi-square analysis indicated that the score distributions for each ethnic group changed significantly from 1996 to 2000. Results of the analyses are presented in Tables 3 and 4.

- For each ethnic group, the distribution of scores in 1996 was significantly different from the distribution of scores in 2000.
- Effect size as measured by the phi coefficient indicated that there was a large gain for African American and a moderate gain for Hispanic students than for white students
- Scores for Hispanic and African American students tended to increase from the lowest score range to middle and upper portion of the score range.
- Scores for white students were concentrated in the upper portion of the score range.

Table 3
TAAS TLI Score Distributions in Percent for 1996 and 2000 by Ethnic Group

| Ethnic Group | | TLI Score | | | | | |
		0-69	70-74	75-79	80-84	85-89	90-94
African Am.	1996	56.2	11.8	12.0	10.6	7.4	2.0
	2000	19.1	11.3	18.4	24.4	20.9	6.0
Hispanic	1996	48.6	11.9	13.1	12.9	10.3	3.3

	2000	14.5	9.0	16.3	24.9	26.0	9.4
White	1996	22.3	10.0	13.9	19.6	23.2	11.0
	2000	5.2	3.9	9.1	21.4	37.9	22.5

Table 4
Comparison of TAAS Score Distributions for 1996 and 2000
Within Each Ethnic Group

Test	Group	Significance Level	Effect Size (Cramer's V)
TAAS	White 1996-2000	<.01	.332
	African American 1996-2000	<.01	.405
	Hispanic 1996-2000	<.01	.410

NAEP Results

Two sets of NAEP mean scores were reported: NAEP results for Texas and NAEP results for the nation. The comparison of 1996 and 2000 mean scores within each ethnic group is presented in Table 5.

Results indicated that:

- For the NAEP-Texas, mean scores increased significantly for Hispanic students but not for African American or white students.
- For the National NAEP, mean scores increased significantly for white students but not for Hispanic or African American students.

Table 5
Comparison of NAEP Mean Scores for 1996 and 2000 within each Ethnic Group

Group	Ethnicity	Mean Scale Scores		Gain
		1996	2000	
NAEP-Texas	African American	249	252	+3.0
	Hispanic	256	266	+10.0*
	White	285	288	+3.0
NAEP-Nation	African American	242	246	+4.0
	Hispanic	250	252	+2.0
	White	281	285	+4.0*

* p <.05

The second set of analyses used a chi-square analysis to compare the 1996 and 2000 score

distributions for each ethnic group. Results of the analyses are presented in Tables 6 and 7.

- For the NAEP-Texas, there was a significant change in the score distributions of all three ethnic groups from 1996 to 2000.
- The effect size (phi coefficient) for the NAEP-Texas was very small for white (.067) and African American students (.127) and slightly larger for Hispanic students (.176). This indicates that there was a more of a change in the Hispanic students' score distribution than for white or African American students.
- For the National NAEP, white students' score distribution changed significantly from 1996 to 2000 but the score distributions for Hispanic and African American students did not.
- The effect size for the National NAEP was very small for all three groups (.030 to .053), indicating that the changes from 1996 to 2000 were minimal.

Table 6
Distribution of NAEP Score in Percent for 1996 and 2000 by Ethnic Group

Ethnicity	Year	Below Basic	Basic	Proficient	Advanced
Texas					
African Am.	1996	69	26	4	1
	2000	60	34	6	0
Hispanic	1996	58	34	7	1
	2000	41	45	13	1
White	1996	22	45	29	4
	2000	17	46	33	4
Nation					
African Am.	1996	73	23	4	0
	2000	68	27	5	0
Hispanic	1996	63	29	7	1
	2000	60	31	8	1
White	1996	27	43	25	5
	2000	23	43	28	6

Table 7
Comparison of NAEP Score Distributions for 1996 and 2000 within Ethnic Group

Test	Group	Significance Level	Effect Size (Cramer's V)
NAEP-Texas	White 1996-2000	<.05	.067
	Hispanic 1996-2000	<.01	.176

|			
	African Am. 1996-2000	<.05	.127
NAEP-National	White 1996-2000	<.01	.052
	Hispanic 1996-2000	N.S.	.030
	African Am. 1996-2000	N.S.	.053

Conclusions

Comparison of the TAAS results and NAEP-Texas results show a significant mean score increase with a moderate to large effect size for all three ethnic groups on the TAAS. In contrast, only Hispanic students had a significant mean score increase on the NAEP-Texas, with a small effect size. The Hispanic students' increase on the Texas NAEP was not reflected on the National NAEP, indicating that the NAEP-Texas result was not part of a national trend.

The score distributions for both the TAAS and the NAEP-Texas changed significantly from 1996 to 2000. Effect sizes for changes in the TAAS score distributions were much larger than those found for the NAEP-Texas (Figure 1). In addition, the pattern of change in the distributions was different for the TAAS than for the NAEP-Texas. The TAAS distributions for African American and Hispanic students showed a very large decrease in students who failed the test (i.e. scored below 70 TLI) and an increase in scores in the middle to high range. For white students, the percentage at the lower and middle range decreased and the percentage at the top of the test score range showed a very large increase. In contrast, the NAEP-Texas distributions had changes primarily at the lower end of the score range for all three ethnic groups; that is, from "below basic" to "basic".

Figure 1.
   Effect Size Comparison for 1996 and 2000 Score Distributions
   for TAAS and NAEP
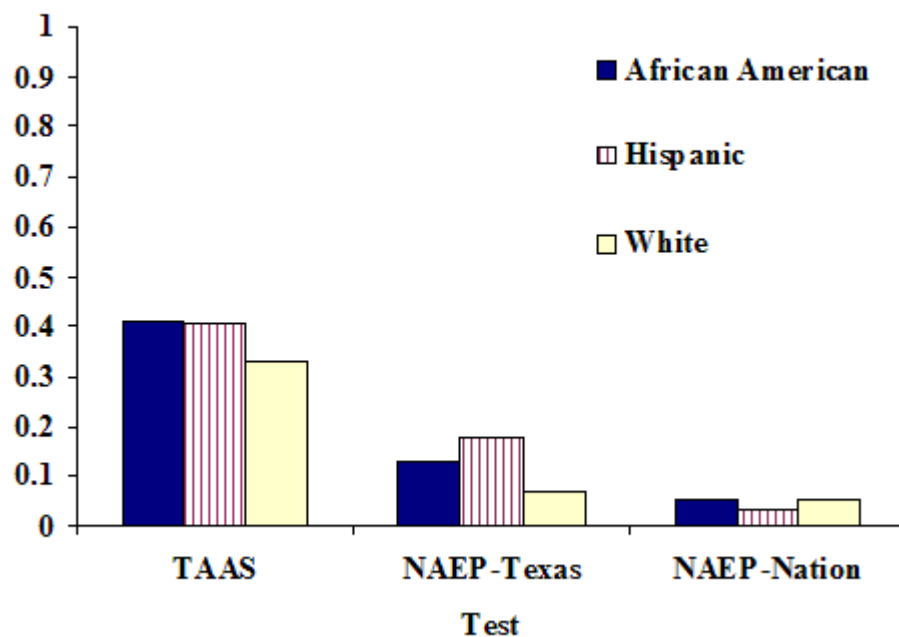
Research Question 2

Do TAAS and NAEP data show that the achievement gap between white and minority students decreased from 1996 to 2000?

 The 1996 and 2000 TAAS scores and NAEP scores for white and minority students were compared in two ways. First, mean scores for white students were compared with mean scores for African American and Hispanic students to determine if there were significant differences. Second, the distribution of test scores for white students and minority students were compared to determine if the distribution patterns became more similar over time.

TAAS—Comparison of White and Minority Scores

The comparison of mean scores for white and minority students on the TAAS presented in Table 8. Confidence intervals were used to determine the statistical significance of differences between white and minority students' mean scores in 1996 and in 2000.

- Mean scores for white students were significantly higher than African American students in both 1996 and 2000.
- The difference in mean scores for white and African American students was larger in 1996 than in 2000.
- Mean scores for white students were significantly higher than Hispanic students in both 1996 and 2000.
- The difference in mean scores for white and African American students was larger in 1996 than in 2000.
- Effect sizes for white vs. African American students was large while effect size for white vs. Hispanic students was moderate.

Table 8
Comparison of TAAS Mean TLI Score For White and Minority Students

Comparison Group	Mean Difference	Effect Size (d Statistic)
African American vs. White 1996	12.2*	.917
African American vs. White 2000	7.0*	.824
Hispanic vs. White 1996	9.4*	.682
Hispanic vs. White 2000	4.9*	.553

* p <.01

A chi-square analysis compared the score distribution for white students with the score distributions for Hispanic students and African American students. Results of the analyses are presented in Table 9.

- The score distribution for white students was significantly different from the score distributions for African American students both 1996 and 2000.

- The score distribution for white students was significantly different from the score distributions for Hispanic students both 1996 and 2000.
- Effect sizes changed very little from 1996 to 2000. This indicated that, although the distribution patterns changed, the differences between white students' scores and minority students' scores were relatively unchanged from 1996 to 2000.

Table 9
Comparison of TAAS Score Distributions For White and Minority Students

Comparison Group	Significance Level	Effect Size (Cramer's V)
African American vs. White 1996	<.01	.331
African American vs. White 2000	<.01	.333
Hispanic vs. White 1996	<.01	.313
Hispanic vs. White 2000	<.01	.286

NAEP—Comparison of White and Minority Scores

The comparison of mean scores for white and minority students on the NAEP-Texas and the National NAEP are presented in Table 10. Confidence intervals were reported to show the statistical significance of differences between white and minority students' mean scores in 1996 and 2000. For the NAEP-Texas:

- White students' mean scores were significantly higher than African American and Hispanic students' mean scores both in 1996 and in 2000.
- The differences between white and African American students' mean scores remained large and relatively unchanged from 1996 to 2000.
- The difference between white and Hispanic students' mean scores decreased from 1996 to 2000.

For the National NAEP, white students' mean scores were significantly higher than African American and Hispanic students' mean scores both in 1996 and in 2000. The differences were large and did not change appreciably from 1996 to 2000.

Table 10
Comparison of NAEP Mean Scale Scores For White and Minority Students

Test	Comparison Group	Mean Difference
NAEP-Texas	African American vs. White 1996	36[*]
	African American vs. White 2000	36[*]
	Hispanic vs. White 1996	29[*]
	Hispanic vs. White 2000	22[*]
NAEP-National	African American vs. White 1996	39[*]

	African American vs. White 2000	39 *
	Hispanic vs. White 1996	31 *
	Hispanic vs. White 2000	31 *

* p <.01

A chi-square analysis compared the score distribution for white students with the score distributions for Hispanic students and African American students. Results of the analyses are presented in Table 11.

- On the NAEP-Texas, the score distributions for Hispanic and African American students were significantly different from that of white students in both 1996 and 2000.
- On the NAEP-Texas, the effect size for the comparison of Hispanic students' and white students' scores in 2000 was smaller than in 1996. In contrast, the effect size for the comparison of African American students' and white students' scores were unchanged from 1996 to 2000.
- On the National NAEP, the score distributions for African American and Hispanic students were significantly different from the score distribution for white students in both 1996 and 2000.
- On the National NAEP, effect size was moderate and there was little change from 1996 to 2000.

Table 11
Comparison of NAEP Score Distributions For White and Minority Students

| Test	Comparison Group	Significance Level	Effect Size (d Statistic)
NAEP-Texas	African Am. vs. White 1996	<.01	.415
	African Am. vs. White 2000	<.01	.428
	Hispanic vs. White 1996	<.01	.402
	Hispanic vs. White 2000	<.01	.318
NAEP-National	African Am. vs. White 1996	<.01	.379
	African Am. vs. White 2000	<.01	.383
	Hispanic vs. White 1996	<.01	.291
	Hispanic vs. White 2000	<.01	.327

Conclusions

The TAAS results show that the difference in mean scores for white and minority students was smaller in 2000 than in 1996. This would seem to indicate that minority students were closing the achievement gap in eighth grade mathematics. However, results of the NAEP-Texas offer only partial support for this conclusion. On the NAEP-Texas,

the difference for white and African American students was unchanged from 1996 to 2000. The difference in NAEP-Texas mean scores for Hispanic and white students was smaller in 2000 than in 1996, although the difference was still large. In contrast, results from the National NAEP indicated that the difference between Hispanic and white students' mean scores actually increased slightly.

Comparison of the score distributions of white and minority students presents a similar result. Comparison of the score distributions of white and African American students yielded similar effect sized in 1996 and 2000. In contrast, the comparison of white and Hispanic students showed that the effect size decreased from 1996 to 2000 (Figure 2).

Figure 2.
Comparison of Effect Size for White vs. Minority Score Distributions
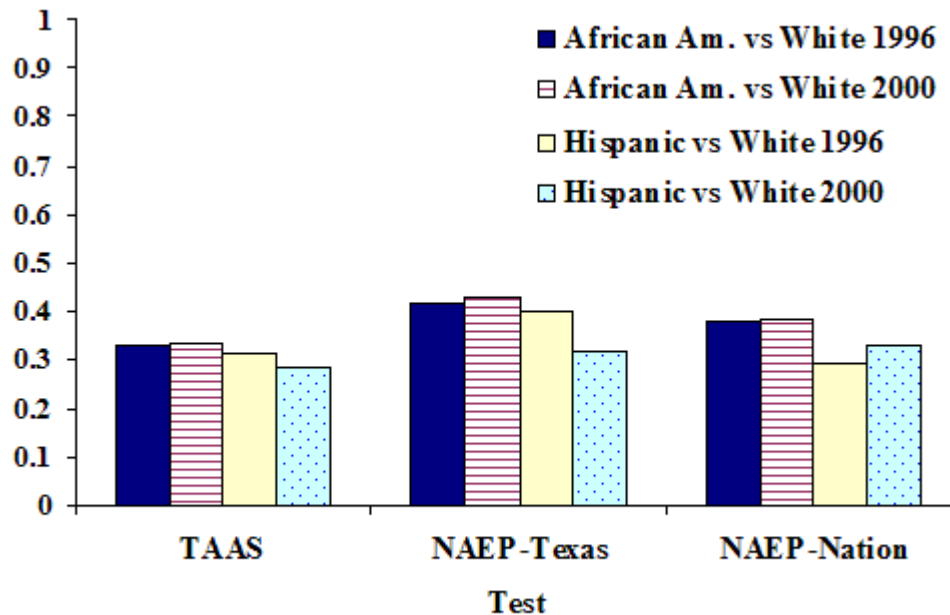
The finding that the effect size for comparisons of minority and white students is larger on the NAEP-Texas than on the TAAS show that the disparity between minority and white students is greater on the NAEP-Texas. That is, the achievement gap is more evident on the NAEP-Texas than on the TAAS both in 1996 and in 2000. However, on both tests, the achievement gap between Hispanic and white students was smaller in 2000 than in 1996. The fact that Hispanic students do not show similar gains nationally indicates that this is not part of a national trend. The disparity between the NAEP-Texas and National NAEP may be an indication that Hispanic students in Texas are beginning to close the achievement gap in eighth grade mathematics.

Research Question 3

Is there evidence that a ceiling effect artificially restricted the distribution of students' scores on the TAAS or the NAEP?

TAAS Scores

The analysis of TAAS mean score gains for each ethnic group show that white students

gained only 7.0 TLI points from 1996 to 2000 while African Americans gained and Hispanic students gained 11.5. Since the largest percentage of white students scores were in the upper 10% of the score range in both in 1996 and 2000, the gains for their highest scoring students were limited to the maximum score possible on the test. The likely result is that they were not able to show their true achievement level because the maximum score (ceiling) of the test artificially limited their scores. If this were the case, comparison of their scores with those of minority students (whose opportunity for gain was not as restricted) would create the appearance that the lower scoring students were achieving at a greater rate and therefore closing the achievement gap.

A second analysis looked at the distribution of scores for each ethnic group for evidence of a ceiling effect. This analysis, presented in Table 12, gives the percent of students in each ethnic group with TLI scores in the upper 10% of the score range (i.e. a TLI of 85 to 94). The table shows that in 1996, 34% of the white students had test scores in the upper 10% of the TAAS score range. This increased to 60% in 2000. The fact that the white students have the largest percentage of students in the upper range indicates that the score range for these students is more restricted by the maximum test score (the test ceiling).

The dramatic difference in the score distributions for white and minority students provides support for the hypothesis that a ceiling effect has restricted white student' scores to a greater degree than it has restricted Hispanic and African American students scores. If this hypothesis is correct, the result would be an artificial narrowing of the achievement gap between white and minority students' eighth grade math test scores.

Table 12
Percent of Students with TAAS TLI Scores Greater than 84

Math	TLI Scores		Cumulative Total
	85-89	90-94	
White 1996	23.2	11.0	34.2
White 2000	37.9	22.5	60.4
African American 1996	7.4	2.0	9.4
African American 2000	20.9	6.0	26.9
Hispanic 1996	10.3	3.3	13.6
Hispanic 2000	26.0	9.4	35.4

NAEP Scores

NAEP-Texas and National NAEP results showed that the mean scores for all three ethnic groups were near the middle of the test range (0 to 500). An analysis of the score distributions for the NAEP-Texas and the National NAEP (Table 13) show that student gains have been primarily at the lower range of the test (from "below basic" to "basic"), with little change in the percent of students achieving the "advanced" range. There is no evidence to support the hypothesis that there is a ceiling effect for either the NAEP-Texas or the National NAEP results.

Table 13
Percent of Students with NAEP Scores Above Basic

Ethnicity	Year	Proficient	Advanced	Cumulative Total
NAEP-Texas				
African American	1996	4%	1%	5%
	2000	6%	0%	6%
Hispanic	1996	7%	1%	8%
	2000	13%	1%	14%
White	1996	29%	4%	33%
	2000	33%	4%	37%
NAEP-National				
African American	1996	4%	0%	4%
	2000	5%	0%	5%
Hispanic	1996	7%	1%	8%
	2000	8%	1%	9%
White	1996	25%	5%	30%
	2000	23%	43%	34%

Discussion

White, African American, and Hispanic students all had large and statistically significant gains on the TAAS from 1996 to 2000. Comparison of white and minority students' scores show that white students had significantly higher TAAS scores than African Americans and Hispanics in 1996 and 2000, but that the differences were smaller in 2000. These results were consistent both for analysis of mean scores and analysis of the test distributions for each student group.

NAEP results were not consistent with the TAAS results. Hispanic students had a mean score gain from 1996 to 2000, but white and African American students did not. When white students' NAEP scores were compared to minority students' scores, the difference between Hispanic and white students' scores decreased from 1996 to 2000 but the difference between African American and white students' scores did not. These results were consistent for analysis of mean scores and test distributions for each student group.

In summary, the large student gains on the TAAS, which is a minimum skills test tailored specifically to the Texas mathematics curriculum, are only partially substantiated by the smaller gains on the NAEP, which is a more general and more difficult test of mathematics. While an explanation of the reasons for the differences in the TAAS and NAEP results is beyond the scope of this research, the authors' experience in Texas public schools suggests two likely answers. First, teaching to the TAAS is widespread and

pervasive in Texas schools. Release versions of the TAAS are available from the Texas Education Agency (along with scoring services that mimic actual TAAS reports) as are a variety of commercially developed practice and test preparation materials. It is common practice for schools to administer one or more "practice TAAS" tests in the fall and use the results to guide instruction in preparation for the state-mandated TAAS testing in the spring. Second, Texas teachers and principals are evaluated, in part, on their students' success (or lack of success) on the TAAS. These factors create very strong pressure to teach to the test. It is likely that score inflation is a significant factor in the large gains that have been consistently reported for the TAAS.

TAAS data for 2000 revealed that differences between white and minority students' scores had decreased when compared to 1996. Other studies have considered this as evidence that the achievement gap between white and minority students is being narrowed. However, analysis of the distribution of scores for each ethnic group reveal that over 60% of white students scored in the upper 10% of the test score range while about 27% of African American students and 35% of Hispanic students scored in this range.

Since a larger percentage of white students than minority students achieved the maximum score on the test (the test ceiling), white students' scores likely underestimated their true achievement level. That is, the ceiling effect has artificially restricted white students scores and created the illusion that the achievement gap has been narrowed. The presence of a ceiling effect casts doubt on the validity of claims that the achievement gap between white and minority students has been narrowed. A more reasonable interpretation of the available data is that because the test ceiling has differentially affected the scores for white, Hispanic, and African American students, TAAS results cannot be used to determine whether or not the achievement gap has been narrowed.

Analysis of the NAEP-Texas score distribution suggests that the achievement gap for African American and white students has not changed between 1996 and 2000. However, the gap between white and Hispanic students' scores did narrow, although the change was small when compared to the TAAS. Comparison of NAEP-Texas and National results indicates that this change was a Texas phenomena and was not found in the National NAEP data. This finding does indicate that Texas has been partially successful in narrowing the achievement gap between Hispanic and white students.

The results of this study have implications beyond the TAAS and the State of Texas. The national emphasis on high standards and the use of high stakes criterion-referenced tests to measure progress toward those standards have become commonplace in public education. Many states depend solely on high stakes test results for making far-reaching decisions about the content and formulation of curricula, the funding of educational initiatives, and the development of educational policy. Any state that uses a high stakes test to measure progress toward state standards must be aware of the twin dangers of test score inflation and ceiling effects. Both can lead to invalid interpretation of test scores and erroneous conclusions about student achievement. The use of comparative data such as the NAEP are vital to ensure that the test data used by state and national decision-makers presents an accurate picture of the educational achievement of their students.

Notes

1. TAAS data were ordered as a customized report of frequency distributions by ethnic group. The data set for each ethnic group consisted of a frequency count of the number of students by Texas Learning Index (TLI) score together with the mean, standard deviation, and SEM of the distribution. The TLI score is a scaled score derived specifically for the TAAS and is not comparable to other scaled scores. A complete description of the derivation of the TLI is contained in the TAAS Technical Digest available from the following Texas Education Agency web site (TEA, 2000 c).
2. All data for the NAEP-Texas and National NAEP were obtained from the following NCES reports: *The nation's report card: Mathematics 2000.* (U.S. Department of Education, 2001a) and *The nation's report card: state mathematics 2000, report for Texas* (U.S. Department of Education, 2001b)
3. The *d* statistic is the ratio of the mean difference between two groups divided by the pooled standard deviation. A value of .2, .5, or .8 is generally interpreted as small, medium, or large effect size, respectively.
4. The Cramer's V coefficient is a rescale of the phi coefficient and has a range between 0 and 1. A Cramer's V of .1, .3, or .5 is generally interpreted as small, medium, or large effect size, respectively.

References

Amrein, A. L. and Berliner, D. C. (2002). High-stakes testing, uncertainty, and student learning. *Education Policy Analysis Archives, 10(18).* Retrieved March 14, 2003, from http://epaa.asu.edu/epaa/v10n18/.

Camilli, G. (2000). Texas gains on NAEP: points of light? *Education Policy Analysis Archives, 8(42).* Retrieved March 14, 2003, from http://epaa.asu.edu/epaa/v8n42/.

Green, S. & Akey, T. (2000). *Using SPSS for windows: Analyzing and understanding data* (2nd ed.). Upper Saddle River, NJ: Prentice-Hall.

Haney, W. (2000, August). The myth of the Texas miracle in education. *Education Policy Analysis Archives, 8*(41). Retrieved March 14, 2003, from http://epaa.asu.edu/epaa/v8n41/.

Hurley, E., Chamberlain, A., Slavin, R., and Madden, N. (2001). Effects of success for all on TAAS reading scores: A Texas statewide evaluation. *Phi Delta Kappa, 82*(10), 750-756.

Jerald, C. (2000, January 13). The state of the states. *Education Week, 19*(18), 62-65.

Jerald, C. (2001). Real results, remaining challenges: The story of Texas education reform. *The Education Trust.* Retrieved March 14, 2003, from http://www.brtable.org/document.efm/532.

Klein, S., Hamilton, L., McCaffrey, B., & Stecher, B. (2000). What do test scores in Texas tell us? *Education Policy Analysis Archives 8(49).* Retrieved March 14, 2003, from http://epaa.asu.edu/epaa/v8n49/.

National Center for Educational Statistics (2001a). *The nation's report card: Mathematics*

2000 (NCES Document No. 2001-517). Washington, D.C: Office of Educational Research and Improvement, U. S. Department of Education.

National Center for Educational Statistics (2001b). *The nation's report card: state mathematics 2000, report for Texas* (NCES Document No. 2001-519 TX). Washington, DC: Office of Educational Research and Improvement, U. S. Department of Education.

Texas Education Agency (2000 a, April 20). *State's exit-level TAAS performance sets another record: 10th graders hit 90 percent mark on two sets of tests*. Austin, TX: Author.

Texas Education Agency, Office of Policy Planning and Research. (2000 b, April). *The 2000 accountability rating system for Texas public schools and school districts*. Austin, Texas: Author.

Texas Education Agency (2000 c). *TAAS Technical Digest*. Retrieved January 2, 2002, from http://www.tea.state.tx.us/student.assessment/researchers.html

Toenjes, L. and Dworkin, A. G. (2002). *Are increasing test scores in Texas really a myth, or is Haney's myth a myth?* Education Policy Analysis Archives 10(17). Retrieved March 14, 2003, from http://epaa.asu.edu/epaa/v10n17/.

About the Authors

Dr. Tom Linton is an associate professor in the Educational Leadership Doctoral Program at Texas A&M University-Corpus Christi. His experience includes 25 years in the Corpus Christi ISD Office of Research Testing, and Evaluation and with various other school districts as an outside consultant and as a project evaluator. He has also served on several advisory committees for the Student Assessment Division of the Texas Education Agency. His e-mail address is: tlinton@falcon.tamucc.edu.

Dr. Don Kester is an associate professor in the Educational Leadership Doctoral Program at Texas A&M University-Corpus Christi. His experience includes 27 years as an administrative consultant in the areas of educational program management, evaluation, research, and student assessment with the Los Angeles County Office of Education. His e-mail address is: don.kester@mail.tamucc.edu.

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

Editor: Gene V Glass, Arizona State University

Production Assistant: Chris Murrell, Arizona State University

Commentary Editor is Casey D. Cobb: casey.cobb@unh.edu .

EPAA Editorial Board

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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