DOCUMENT RESUME

ED 382 657 TM 023 091

AUTHOR Wainer, Howard; Saka, Thomas
TITLE On the Academic Performance of Hawaii's Public School Children: I. Fourth and Eighth Grade Mathematics in 1992.
INSTITUTION Educational Testing Service, Princeton, N.J.
REPORT NO ETS-RR-94-43
PUB DATE Aug 94
NOTE 27p.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Academic Achievement; Comparative Analysis; Demography; Elementary School Students; Grade 4; Grade 8; Intermediate Grades; Junior High Schools; *Mathematics; Mathematics Tests; *Public Schools; Scoring; State Programs; *Test Results
IDENTIFIERS *Hawaii; International Assessment of Educational Progress; National Assessment of Educational Progress

ABSTRACT

This report describes the first of a series of research projects which will attempt to characterize the performance of Hawaii's public school system. The measure of performance was the 1992 National Assessment of Educational Progress (NAEP) mathematics examinations and their linked versions used in the 1991 International Assessment. The NAEP was selected because it is considered satisfactory with respect to content and form and because the psychometric model underlying its scoring yields a single scale that allows easy comparison. Based on the unstandardized results of the 1992 mathematics assessment, Hawaii was among the lowest performing states. Once results were standardized to reflect a single (national) demographic composition, Hawaii's rank among participating states increased only slightly. While the United States ranked near the bottom for the international assessment, Hawaii's students were ranked still lower. The report explicitly avoids any examination of the reasons for the low achievement of Hawaii's students. Eight figures illustrate the analysis. (Contains 9 references.) (SLD)

RR-94-43

RESEARCH REPORT

ON THE ACADEMIC PERFORMANCE OF HAWAII'S PUBLIC SCHOOL CHILDREN: I. FOURTH AND EIGHTH GRADE MATHEMATICS IN 1992

Howard Wainer
Thomas Saka

Educational Testing Service
Princeton, New Jersey
August 1994

On the Academic Performance of Hawaii's Public School Children: I. Fourth and Eighth Grade Mathematics in 1992

Howard Wainer[§]
Educational Testing Service
and
Thomas Saka
Hawaiian State Department of Education

Introduction

"...teaching is validated by the transformation of the minds and persons of the intended audience."

Bressler, 1991

The most critical measure of any educational system is the performance of its students. But what yardstick should be used to accomplish this measure? The fact that modern education has many goals suggests that we must measure the extent of its success in a variety of ways. This report describes the first of a series of researches that will attempt to characterize the performance of Hawaii's public school system. We will do this through comparative and absolute measures, the primary instrument of which will be the data gathered during the course of the National Assessment of Educational Progress (NAEP).

NAEP is a congressionally mandated survey of the educational achievement of American students and of changes in that achievement across time. Although this survey has been operational for nearly 25 years, it was only in 1988 that Congress authorized adding state level surveys to the national assessment. This was begun on a trial basis with states participating on a voluntary basis. In 1990 37 states, two territories and the District of Columbia participated in the first Eighth Grade Math Assessment. In 1992 seven more states joined the state assessment yielding 44 jurisdictions. The 1992 assessment was expanded to also include the Fourth grade. In this report we shall focus attention only upon the 41 states in the assessment. Guam, the Virgin Islands and the District of Columbia will be explicitly excluded because they are sufficiently different from the states in their size, character and composition so as to distort most comparisons.

The assessment methodology is technically sophisticated. Through the use of linking items and item response theory, all students participating in the assessment can be placed on the same numerical scale. Measuring students' growth is thus straightforward. Subtracting 4th grade scores from 8th grade scores is the growth obtained. Consequently the expansion of the assessment to the fourth grade provides us with two important bits of information. First, is a measure of how much mathematics Fourth graders know. Second, a measure of how much mathematics is learned between 4th and 8th grade. Note that having a measure of the gain obtained (about 49 NAEP points on average) helps us to

§ This research was supported by the ETS Senior Scientist Award to the first author from the Trustees of the Educational Testing Service. We are pleased to acknowledge this help. Furthermore we are grateful for the advice and help of Brent Bridgeman, Philip Leung and John Mazzeo.

interpret the NAEP scale. It tells us that if one state trails another by about 12 points this is about the same as the average gain in one year of school. Thus, when we compare Mississippi's mean 8th grade NAEP score of 246 to Hawaii's score of 257, we can interpret the 11 point difference as indicating that the average Mississippi 8th grader performs about the same in mathematics as the average Hawaii 7th grader would have. This helps give additional meaning to the numerical scale.

More meaning still for the eighth grade math assessment is yielded by comparing performance on it with performance of 13 year old students in the 1991 International Assessment. Because the NAEP Math assessment was coordinated with the International Assessment both sets of scores can, with reasonable confidence, be placed on the same scale. This was accomplished by having a common sample of examinees for both assessments. As we shall see, the performance of Hawaii's students does not compare favorably with those from the developed nations.

The Mathematics assessment contained tasks for the students drawn from the framework provided by the *Curriculum and Evaluation Standards for School Mathematics*, developed by the National Council of Teachers of Mathematics. The content and the structure of the assessment has been widely praised as being representative of the best that current knowledge and technology allows. A full description of the 1992 Mathematics Assessment is found in *NAEP 1992: Mathematics Report Card for the Nation and the States* (Mullis, Dossey, Owen, & Phillips; 1993).

The NAEP State Assessment Sample

Within each state 100 public schools[1] are carefully selected to be representative of all public schools in that state. Within each school at least 30 students are chosen at random to be tested (in larger schools this number can be as large as 90). In Hawaii all public schools were sampled. Students (usually of foreign birth) whose English language proficiency is deemed to be insufficient to deal with the test, are excluded from the sampling frame.

The Results

All results are reported on a uniform scale that can meaningfully characterize the performance of students over a very wide range of proficiency. This scale can be used in a normative manner, for example comparing one state with another, or one state with the nation as a whole. Or it can be used as an absolute measure, since expert judges have indicated what score levels constitute specific proficiencies. These proficiencies are denoted Basic, Proficient and Advanced and what is required to perform at each of these levels obviously increases as the student progresses through school, but are always referred back to the five NAEP content areas. These are: (1) numbers and operations, (2) measurement, (3) geometry, (4) data analysis, statistics, and probability, and (5) algebra and functions.

[1] In small states, where the number of public schools is fewer than 100, all schools were sampled.

For example, a score of 211 is characterized as "Basic Level" fourth grade performance. "Basic Level" is defined as *"showing some evidence of understanding the mathematical concepts and procedures of the five NAEP content areas."* The second level is called "Proficient" and is located at score 248 and reflects being able to *"consistently apply integrated procedural knowledge and conceptual understanding to problem solving in the five NAEP areas."* The highest performance level is termed "Advanced", is located at score 280 and reflects the ability to *"apply integrated procedural knowledge and conceptual understanding to complex and nonroutine real-world problem solving in the five NAEP areas."*

The mean performance of all participating states for the 8th grade assessment is shown in Figure 1.

Insert Figure 1 Here

The results shown in Figure 1, while accurately reflecting the actual mean performance within each state, may not be appropriate for certain kinds of state-by-state comparisons. The student populations of each state differ in their demographic make-up. As such, some states face more difficult challenges in educating their populations than others. One obvious example of such a situation occurs in states like California, Hawaii and Florida that have large immigrant populations whose children, even if they do not participate directly in the assessment, require a larger share of instructional resources than native English speakers. In addition, the various subpopulations of students in each state often perform very differently from one another. For example, in Figure 2 are displayed the mean performance of students in different parts of the country broken down by race/ethnicity.

Figure 2 has two important messages:

1. There are very large difference in performance by ethnic group. These differences are much larger than the geographic variation observed.

2. Hawaii's students perform worse than the national average and all regional averages for all groups. This does not appear to be due to Hawaii's unusual demographic mixture, for although it is true that Hawaii's Asian/Pacific Islander and Hispanic students do worse than White students, they also do worse than Asian/Pacific Islander and Hispanic students in any region. Part of this effect may be due to the over-broad classification 'Asian/Pacific Islander,' a topic that we shall return to later.

Insert Figure 2 Here

In addition to the widely different performances of the various demographic subgroups, the distribution of these groups is not uniform across all states. A brief summary of these distributions is shown in Table 1. As is evident, Hawaii's racial/ethnic distribution is very different than that of the nation as a whole. Hawaii's demographics are the most deviant of all of the states in the assessment.

Table 1

NAEP 1992 Trial State Assessment
Percentage Race/Ethnic Representation in NJ
Compared to that in other parts of the Nation

	Race/Ethnicity				
Mathematics	*White*	*Black*	*Hispanic*	*Asian*	*Other*
NATION	**69**	**17**	**10**	**2**	**2**
Northeast	68	20	9	2	1
Southeast	63	29	5	1	2
Central	79	11	7	1	2
West	65	11	16	5	3
Hawaii	**17**	**3**	**11**	**66**	**3**

Table 1. The national and regional racial/ethnic distribution.

If we wish to use such data to draw inferences about the relative efficacy of a state's schools, we must statistically adjust for the demographic differences. Why is it helpful to make such adjustments? It is beyond the goals of this report to investigate fully why there are differences in performance by demographic group, although there is a rich literature of fact and conjecture that attempts to do so[2]. However, to understand why we need to make a statistical adjustment it is important to provide some explanation. To do so requires that we draw the important distinction between education and schooling. The school is only one agency among many -- family, church, neighborhood, mass media -- that provides children with windows on the world. Mass schooling was invented because families could no longer perform essential educational functions. "Once upon a time schools could proceed on the only partly fictitious assumption that, in their efforts to teach children, they were supported by relatively stable families, and by neighborhoods that enforced elementary standards of civility." [3] Not only is this much less true now than in the past, but also it is less true within some demographic groups than others. NAEP

[2]The Coleman report (Coleman et al, 1966) remains the most encyclopedic of such investigations, summarizing, as it does, the performance of more than 645,000 children in 4,000 public schools. It arrives at the not surprising conclusion that family and economic variables drive educational achievement.
[3]This quote and much of the surrounding logic comes from Marvin Bressler's delightful and wise 1992 essay, "A teacher reflects."

measures education and not just schooling, but inferences about the differences among states are explicitly about schools. To add some validity to those inferences we must try, statistically, to place all states on a level playing field with respect to their children's nonschool educational opportunities. Adjusting for differences in demographic groupings is a crude beginning.

What follows is a methodology that recognizes that such differences exist, and a statistical technology that attempts to partition state differences due to demographics from those due to differences in school performance.

Interpretable comparisons through statistical standardization

The between-state comparisons that are implicit in Figure 1 can yield misleading inferences if one is not acutely aware of the differences in the demographic make-up of all of the constituent units. This is of such a complex nature that it is impossible to keep things straight without some formal adjustment. One accepted way to accomplish this adjustment is termed 'standardization' (Mosteller & Tukey, 1977, p. 223). The basic idea of standardization is to choose some specific demographic proportions as the standard and then estimate each state's mean proficiency on that specific mixture. In this instance it is sensible to choose the configuration of the entire United States as the standard mixture. Thus the estimated score for each state will be the answer to the question, "What would the national average be if all children went to school in this state?"

How is this adjustment accomplished? It is very simple in theory, although sometimes, because of peculiarities in sampling weights, a bit tricky to execute. We take the mean score obtained in a state for a particular subgroup and multiply it by that subgroup's proportional representation in the standard (national) mixture. Do this for all subgroups and the resulting score is the adjusted one. So far we have reported Hawaii's scores for three racial/ethnic groups. As we have seen, because Hawaii's demographics are so unlike the national standard this sort of adjustment could have a large effect. As it will turn out, it does not. Is it sufficient to adjust simply on the basis of this one demographic variable? No, although if we adjust on too many variables, and so include some irrelevant ones, no damage will be done, since irrelevant variables will typically not show differences in performance. In this paper we adjust on three variables. These are:

1. Race/ethnicity - Five categories: White, African-American, Hispanic, Asian/Pacific Islander, Other

2. Type of community - Four categories: Extreme Rural, Advantaged Urban, Disadvantaged Urban, Other.

3. Limited English Proficiency - Two categories: Yes, No

This resulted in dividing each state up into forty pieces corresponding to the forty possible combinations (5×4×2) and calculating the mean proficiency within each of those 40 groups. These 40 means were then weighted by their representation in the entire nation

yielding a standardized score for each state. The results of this standardization are shown in Figure 3.

Insert Figure 3 Here

After standardization to national demographic norms we find that although Hawaii's mean score has not changed, two other states, with more homogeneous student populations (West Virginia and Tennessee) that had previously been slightly higher are now ranked equal to or below Hawaii.

What is the point of standardization? There are many reasons. So far we have mentioned just one -- making comparisons between states on the basis of their children's performance on the same tasks and not on the differences in the demographic structure of their population. A second, and oftentimes more important use of standardized scores is in easing the difficulties in making inferences about changes that occur within a state across time. When changes do occur the standardized scores assure that the change reflects changes in the students' performance and not changes in the demographic structure of the state. We expect that as time goes on this will be the aspect of greatest value of the standardization.[4]

A natural question to ask is, "At what age do the differences observed among the states manifest themselves?" If we see the same difference between two states in 4th grade as we do in 8th, it implies that the lower scoring state needs to place more emphasis on learning in lower grades. If the difference observed in 4th grade grows proportionally larger in 8th it means that the deficit is spread throughout the years of school and a more systemic change is needed for improvement. Trying to make inferences of this sort based on just two time points is risky, but is certainly instructive. Shown in Figure 4 are the standardized scores for the 1992 4th grade math assessment. A comparison with the 8th grade rankings shown in Figure 3 indicates that the positions established in 4th grade are maintained and the differences observed between states increase. The range of 24 NAEP points observed between the relatively extreme states of North Dakota and Mississippi in 8th grade was 13 NAEP points in 4th grade. One way to interpret this is that the average Mississippi 4th grader was a year behind the average North Dakota 4th grader in math, and by the time they both reached 8th grade this deficit had increased to two years.

[4]A caveat: Big changes as a result of a statistical adjustment tell us that great care must be exercised in making inferences. A careful comparison of Figure 1 and Figure 3 reveals that most states do not change very much. This is evidence that the standardization is generally behaving as it ought, for if one disagrees with the structure of the adjustment one can still be content that it isn't changing anything drastically. A notable exception to this would be the District of Columbia, whose small size and atypical demographic structure would combine to yield an enormous shift. Inferences about the meaning of its standardized location ought not be the same as those drawn about the states. For the more important purpose of tracking changes in a jurisdiction's performance over time, it is probably prudent to develop a special standardization for each of the four most unusual jurisdictions (DC, Hawaii, Guam, Virgin Islands).

Insert Figure 4 Here

By subtracting the scores shown in Figure 4 from those in Figure 3 we obtain estimates of the average growth exhibited in each state. This result is shown in Figure 5 below. All states' scores are standardized to the demographic structure of the nation as a whole. Thus were these results longitudinal rather than cross-sectional, we would be able to interpret the changes as due entirely to growth and not demographic changes. As they are now constituted these changes in scores are due to differences in performance and not to demographic differences in the two grades. As we can readily see, while Hawaii's 4th grade children are already performing well below the national average the gains shown between 4th and 8th grade are also smaller than average. This makes low performance of Hawaii's 8th graders inexorable.

Insert Figure 5 Here

International Comparisons

As mentioned previously, the 1991 International Assessment contained enough NAEP items to allow accurate comparisons. The most newsworthy result was that the United States finished near the bottom in this assessment, finishing ahead of Jordan but behind all of the participating developed nations. This was (properly) viewed with alarm. The 1992 assessment showed some improvement, with the United States climbing past Spain in the rankings to settle in comfortably next to Slovenia. However, as we have seen in the preceding figures, there is tremendous variation within the United States. Shown in Figure 6, are the results of this assessment augmented by the inclusion of the individual states (standardized to national demographics). As is evident, Hawaii's 8th grade students' performance was better than those of Jordan's, but on average ranked about a year behind Spain's and two and a half year's behind those of Korea and Taiwan. Further details of the International Assessment can be found in Salganik, Phelps, Bianchi, Nohara, & Smith (1993).

Insert Figure 6 Here

Thus we see rather dramatically that because of the great diversity within the United States looking at just an overall figure for the entire country provides an incomplete and, for some purposes, misleading picture. Because Hawaii's score is standardized to the demographic structure of the entire nation one can interpret this result as what the nation's

location would have been if all of the states' school systems performed as well as Hawaii's.

Within state variation

We have seen that the variation among states (roughly 30 NAEP points from highest to lowest) makes interpretation of a national mean of limited value. In the same way, the variation within states dwarfs the variation between them. In most states the average score obtained by the lowest 10% of the students is more than 90 NAEP points lower than the score obtained by the top 10%[5]. 90 points is an enormous gulf. Before trying to understand the reasons for this great disparity (with an eye toward developing strategies for ameliorating it) it will be useful to continue this series of comparisons for one important segment of the population -- the very top.

In Figure 7 is a comparison of the performance of the top 5% in the 1992 8th grade math assessment with the top 5% of the various OECD countries. We see immediately that Hawaii's top 8th grade students compare favorably with their European counterparts although still trail Taiwan and Korea's top students. The United States as a whole has also improved relative to the other countries, but still lags the other developed nations by from 3 to 12 months.

Insert Figure 7 Here

But is this the whole story? In Hawaii the ethnic diversity is enormous. Shown in Table 2 are the results from the Spring 1993 statewide assessment of 8th grade mathematics that used the Stanford Achievement Test. The table shows the extent of Hawaii's ethnic diversity. It also shows that collapsing so much of this diversity (those shown in *italics*) into a single classification of "Asian/Pacific Islander" conceals important information. The last column of the table shows the Stanford scores transformed into the same scale as NAEP.[6]

---

[5]In all but one of the OECD countries this gulf between the 10th percentile and the 90th is somewhat smaller, about 70 points. Taiwan is the lone exception a difference of 96 points.

[6]This was achieved simply by linearly transforming the Stanford scores to have the same mean and variance as the within Hawaii NAEP scores (NAEP score = 174 + 1.74 Stanford Score). While we would not swear that this is absolutely the same as what would be obtained if we had the appropriate ethnic information from NAEP, we believe that it is close enough for our immediate purposes. Some confirmation to this claim is obtained by noting that those Hawaiian ethnic group scores that are available from NAEP (White and the Asian/Pacific aggregate) are predicted accurately.

Table 2

Ethnicity	N	%	Stanford Math Score	Math scaled in NAEP Units
Part-Hawaiian	2,272	20.1%	39.4	243
Filipino	2,232	19.8%	43.7	250
White	1,992	17.7%	51.1	263
Japanese	1,663	14.7%	62.4	283
Other	946	8.4%	46.2	255
Chinese	369	3.3%	64.0	286
Hawaiian	323	2.9%	38.3	241
Portuguese	315	2.8%	38.2	241
Samoan	298	2.6%	35.2	235
Spanish	262	2.3%	39.2	243
Black	242	2.1%	42.6	248
Korean	223	2.0%	63.3	284
Indo-Chinese	107	0.9%	53.7	268
American Indian	38	0.3%	42.0	247
Total/Mean	**11,282**	**100%**	**47.6**	**257**

Many interesting hypotheses are suggested when we combine the rankings of the Hawaiian ethnic groups associated with these scores with the rankings of the countries in the International Assessment previously shown in Figure 6. The result is shown in Figure 8. In this figure we see the amazing concordance between the performance of Hawaiian students of Chinese background with the performance of students from Taiwan; Korean students show a similar agreement. This remarkable similarity suggests a deeper look can provide us with important insights into what are the causal determinants of outstanding performance. We should start by seeking answers to these questions:

> What is there about the educational environment of Hawaii's students of Chinese and Korean background that is like the educational situation in China and Korea?
>
> How is this situation different from that of Hawaiian students of Samoan or Hawaiian background?

It seems to us that the distinction between education and schooling, made earlier, will probably ought usefully be included in this discussion.

Insert Figure 8 Here

Obiter dictum

The demographic standards, on which the score standardization is based, was developed for the country as a whole. They seem especially unsuitable for Hawaii. We can get a rough idea of at least one component of an alternative classification scheme from the data shown in Table 2. We can establish these as the 14 principal ethnic groups within Hawaii, and set this proportional representation within the population as the standard. We can thus standardize all future test results to this population mix and thence any changes observed will be due to changes in students' performance and not changes in the demographic mixture that comprises modern Hawaii.

Summary & Conclusions

This report is the first of a series that examines of the performance of Hawaii's school children relative to other children within the United States and world-wide. The measure of performance used was the 1992 National Assessment of Educational Progress mathematics exams and their linked versions used in the 1991 International Assessment. This is done in the ardent belief that the efficacy of schools must be measured by the performance of their students. We chose NAEP for several reasons, three of which were:

1. It is composed of test items that satisfy the best of current wisdom with respect to both their content and their form.

2. The psychometric model underlying the scoring of NAEP yields a single scale on which not only can the fourth grade and eighth graders be characterized, but also the 13 year olds from the OECD countries from around the world.

3. The students sampled by the NAEP are drawn in a principled way from the populations of interest. This in sharp contrast to the sorts of self-selected samples that are represented by state means of such college admission tests as the SAT and the ACT. It is well known that trying to draw inferences of useful accuracy from such self-selected samples is impossible (Wainer, 1986a, b; 1989a, b).

We concur with prevailing expert opinion that of all broad-based tests NAEP provides the most honest and accurate estimates of the performance of the students over the broad range of jurisdictions sampled.

We found that, based on the unstandardized results of the 1992 Mathematics Assessment, Hawaii was among the lowest performing states. Once these results were standardized to reflect a single (national) demographic composition Hawaii's rank among the participating states increased only slightly.

The United States finished near the bottom when the performance of it's students was compared with that of the students in the other 14 participating OECD nations in the

1991 International Assessment. Hawaii's students were ranked lower still on the same assessment when their performance was placed on the same scale. However Hawaii's best students, it's top 5%, when compared with the performance of the top 5% of all other OECD nations, ranked eighth; six months ahead of the United States as a whole.

Within Hawaii we found that there were enormous differences in performance by ethnic group. The highest scoring groups (Chinese, Korean and Japanese) did as well as any OECD country. The lowest scoring groups did worse than any OECD country.

This report has explicitly avoided any discussion of possible reasons for the results we report. We have been content, as a start, to merely report the performance of Hawaii's 4th and 8th graders in an important mathematics assessment, and to place these results into a context to make them more understandable. We plan, in subsequent reports, to begin to examine plausible causal variables.

References

Bressler, M. (1991). Reflections on teaching. In *Teaching at Princeton.* Princeton, NJ: Princeton University.

Bressler, M. (1992). A teacher reflects. *Princeton Alumni Weekly, 93(5),* 11-14.

Mosteller, F., & Tukey, J. W. (1977). *Data analysis and regression.* Reading, MA: Addison-Wesley.

Mullis, I. V. S., Dossey, J. A., Owen, E. H. & Phillips, G. W. (1993). *NAEP 1992: Mathematics Report Card for the Nation and the States.* Report No. 23-ST02. Washington, DC: National Center for Education Statistics.

Salganik, L. H., Phelps, R. P., Bianchi, L., Nohara, D., & Smith, T. M. (1993). *Education in States and Nations: Indicators Comparing U.S. States with the OECD Countries in 1988.* NCES Report No. 93-237. Washington, DC: National Center for Education Statistics.

Wainer, H. (1986a). *Drawing inferences from self-selected samples.* New York: Springer-Verlag.

Wainer, H. (1986b). Five pitfalls encountered while trying to compare states on their SAT scores. *Journal of Educational Measurement, 23,* 69-81.

Wainer, H. (1989a). Eelworms, bulletholes & Geraldine Ferraro: Some problems with statistical adjustment and some solutions. *Journal of Educational Statistics, 14,* 121-140 (with discussions). Reprinted in Shaffer, J. P. (Ed.) (1992). *The role of models in nonexperimental social science* (pps. 129-148). Washington, D.C.: American Educational Research Association & American Statistical Association.

Wainer, H. (1989b). Responsum. *Journal of Educational Statistics, 14,* 187-200. Reprinted in Shaffer, J. P. (Ed.) (1992). *The role of models in nonexperimental social science* (pps. 195-207). Washington, D.C.: American Educational Research Association & American Statistical Association.

NAEP 1992 Mathematics Assessment

Overall Proficiency-8th Grade Mathematics
(unstandardized)

283	Iowa	North Dakota			
282	Minnesota				
281					
280					
279					
278	Maine	New Hampshire			
277	Nebraska	Wisconsin			
276					
275					
274	Idaho	Wyoming	Utah		
273	Connecticut				
272	Colorado	Massachusetts			
271	New Jersey	Pennsylvania			
270	Missouri				
269	Indiana				
268					
267	Michigan	Oklahoma	Virginia	Ohio	
266	**NATION**	New York			
265	Arizona	Rhode Island			
264	Maryland	Texas			
263					
262	Delaware				
261	Kentucky				
260	California	South Carolina			
259	Florida	New Mexico	Georgia		
258	West Virginia	Tennessee	North Carolina		
257	**Hawaii**				
256					
255	Arkansas				
254					
253					
252					
251	Alabama				
250					
249	Louisiana				
248					
247					
246	Mississippi				

Figure 1. A stem & leaf display of the 1992 NAEP State Assessment in 8th Grade Mathematics. These results are the raw (unstandardized) means from each state. Hawaii ranks 37th among all participants.

NAEP 1992 Trial State Assessment

Subgroup comparisons of Hawaii with other parts of the Nation

	Race/Ethnicity		
Grade 8 Mathematics	*White*	*Asian/Pacific Islander*	*Hispanic*
287		**NATION**	
286		West	
285			
284			
283			
282			
281			
280	Central		
279	Northeast		
278			
277	West		
276	**NATION**		
275			
274			
273			
272			
271			
270			
269	Southeast		
268			
267			
266			
265	**Hawaii**		
264			
263			
262			
261			
260			
259		**Hawaii**	
•	•	•	•
•	•	•	•
•	•	•	•
246			Central & West
245			**NATION**
244			
243			
242			
241			Northeast
240			
239			Southeast
238			**Hawaii**

Figure 2. A stem & leaf depiction comparing the performance Hawaii's students, broken down by race/ethnicity, with similar groups from all other parts of the country. Samples of "Asian/Pacific Islanders" were insufficient to obtain accurate estimates for any other regions than the West. There were too few "African-Americans" in Hawaii to provide accurate estimates

NAEP 1992 Trial state Assessment
(Standardized for demographic differences)

Grade 8 Mathematics					
278	North Dakota				
277					
276					
275	Iowa	Minnesota			
274	New Jersey	Maine	New Hampshire		
273	Idaho				
272	Connecticut				
271	Massachusetts	Wisconsin			
270	Nebraska	Wyoming	Utah		
269	Texas				
268	Colorado	Pennsylvania	New York	Virginia	Missouri
267	Arizona	California	Maryland	Indiana	Michigan
266	**NATION**				
265	South Carolina	Oklahoma	Ohio		
264	Delaware	New Mexico	Florida		
263	Rhode Island				
262	Georgia				
261					
260	Kentucky				
259	North Carolina				
258					
257	**Hawaii**	Tennessee			
256					
255	Alabama	Louisiana	Arkansas	West Virginia	
254	Mississippi				

Figure 3. After standardization Hawaii ranks 35th among all participating states in the 1992 8th grade mathematics assessment.

NAEP 1992 Trial state Assessment
(Standardized for demographic differences)

Grade 4 Mathematics	
227	New Hampshire
226	Maine
225	
224	Connecticut
223	New Jersey Iowa Wisconsin
222	North Dakota Pennsylvania
221	Minnesota Texas Wyoming Virginia Massachusetts
220	Nebraska Missouri New York
219	Maryland Georgia
218	Colorado Idaho Indiana Michigan Delaware Oklahoma
217	**NATION** Utah Ohio Arizona
216	South Carolina
215	New Mexico North Carolina
214	Rhode Island Florida
213	Kentucky
212	California West Virginia
211	**Hawaii**
210	Tennessee Alabama Arkansas Louisiana
209	Mississippi

Figure 4. The standardized scores for the 41 states in the 1992 4th grade math assessment.

Gain in Mathematics Proficiency from 4th to 8th grade

(Scores are standardized to entire US population)

56	North Dakota
55	California Idaho
54	Minnesota
53	Utah
52	Iowa
51	New Jersey
50	Arizona Colorado Florida Massachusetts Nebraska
49	**NATION** Indiana Michigan New Mexico Rhode Island Wyoming South Carolina Ohio Texas
48	Connecticut Maine Wisconsin New York Maryland Missouri
47	Kentucky Oklahoma Virginia Tennessee New Hampshire
46	Delaware **Hawaii** Pennsylvania
45	Alabama Arkansas Louisiana Mississippi
44	North Carolina
43	Georgia West Virginia

Figure 5. Standardized estimates of change in mathematics performance seen by state between 4th and 8th grade in the 1992 assessment. Hawaii' ranked 33rd among states in gain .

International 1991 Mathematics Assessment
(Predicted Proficiency for 13 year olds)

285	Taiwan	
284		
283	Korea	
282		
281		
280		
279	Soviet Union	Switzerland
278		
277	Hungary	
276		
275		
274		
273	France	
272	Italy	Israel
271		
270	Canada	
269	Ireland	Scotland
268		
267		
266	Slovenia	
265		
264		
263	Spain	
262	**United States**	
•		
256	**HAWAII**	
•		
246	Jordan	

Figure 6. Placing Hawaii explicitly into the 1991 International Assessment shows that its students' performance trailed behind that of all the developed nations.

International 1991 Mathematics Assessment
(Predicted Proficiency for 95th %ile of 13 year olds)

345	Taiwan
•	
•	
•	
335	Korea
334	
333	
332	
331	
330	
329	
328	
327	
326	Hungary
325	
324	Soviet Union
323	
322	Switzerland
321	
320	
319	France
318	
317	**Hawaii** Israel Italy
316	Ireland
315	Scotland Canada
314	
313	
312	**United States**
311	Slovenia
310	
309	
308	
307	
306	Spain
•	
•	
•	
296	Jordan

Figure 7. Hawaii's top students rank seventh in the world in the 8th grade math assessment.

International 1991 Mathematics Assessment (Predicted Proficiency for 13 year olds)		*Estimated 8th Grade NAEP Math Scores for Hawaii's Principal Ethnic Groups*
Taiwan	286	Chinese
Korea	283	Korean, Japanese
Soviet Union, Switzerland	280	
Hungary	277	
France	274	
Italy, Israel, Canada	271	
Ireland, Scotland	268	Indo-Chinese
Slovenia, **United States**	265	
Spain	262	White
	259	
HAWAII	256	Other
	253	
	250	Filipino
Jordan	247	American-Indian, Black
	244	Part-Hawaiian, Spanish
	241	Portuguese, Hawaiian
	238	
	235	Samoan

Figure 8. The performance of Hawaii's 8th grade students in mathematics, when disaggregated by ethnic group and compared with the results of the International Assessment, shows some remarkable similarities.

Figure Captions

Figure 1. A stem & leaf display of the 1992 NAEP State Assessment in 8th Grade Mathematics. These results are the raw (unstandardized) means from each state. Hawaii ranks 37th among all participants.

Figure 2. A stem & leaf depiction comparing the performance Hawaii's students, broken down by race/ethnicity, with similar groups from all other parts of the country. Samples of "Asian/Pacific Islanders" were insufficient to obtain accurate estimates for any other regions than the West. There were too few "African-Americans" in Hawaii to provide accurate estimates.

Figure 3. After standardization Hawaii ranks 35th among all participating states in the 1992 8th grade mathematics assessment.

Figure 4. The standardized scores for the 41 states in the 1992 4th grade math assessment.

Figure 5. Standardized estimates of change in mathematics performance seen by state between 4th and 8th grade in the 1992 assessment. Hawaii ranked 33rd among states in gain

Figure 6. Placing Hawaii explicitly into the 1991 International Assessment shows that its students' performance trailed behind that of all the developed nations.

Figure 7. Hawaii's top students rank seventh in the world in the 8th grade math assessment.

Figure 8. The performance of Hawaii's 8th grade students in mathematics, when disaggregated by ethnic group and compared with the results of the International Assessment, shows some remarkable similarities.