

National Assessment of Educational Progress
The Nation's Report Card: Science 2009
Trial Urban District Assessment

Jack Buckley
Commissioner

National Center for Education Statistics

February 24, 2011

Good morning. I am here today to release the results of the NAEP Science 2009 Trial Urban District Assessment—or as we call it, Science TUDA.

Last month the results of the 2009 science assessment for the nation and the states were released. Today I am releasing the science results for public school students in selected urban school districts. The TUDA program is a collaboration involving the National Center for Education Statistics, or NCES, along with the National Assessment Governing Board and the Council of the Great City Schools.

TUDA is not yet a permanent part of NAEP and participation in TUDA is voluntary. In 2009, 17 large urban school districts from around the country were invited to participate, and all 17 agreed to do so.

Because the NAEP assessments are the same for the nation, the states, and the urban districts, NAEP serves as a common yardstick for comparison.

The list of participating districts has increased from 10 to 17 since the first science TUDA assessment in 2005. In that year the participating districts included Los Angeles, San Diego, Austin, Houston, Atlanta, Charlotte-Mecklenburg, Chicago, Cleveland, New York City, and Boston. In 2009, we added 7 districts—Fresno, Jefferson County, Kentucky, which includes the city of Louisville, as well as Milwaukee, Detroit, Baltimore City, Philadelphia, and Miami-Dade in Florida.

The District of Columbia participates in the TUDA reading and mathematics assessments. To avoid having to require students to participate in more than one NAEP assessment in a single year, the District of Columbia did not participate in the 2009 Science TUDA.

First, a comment about the demographic composition of these urban districts. The student populations of our large urban districts tend to vary both from the national average and from one another. For example, in our 17 districts the percentages of Black students in grade 4 range from a low of 7 percent in Los Angeles to a high of 88 percent in Baltimore City. This compares to an overall national average of 16 percent. Nationally, about 22 percent of fourth-graders are Hispanic, while the district percentages range from 2 percent in Baltimore City to 77 percent in Los Angeles.

About 10 percent of students nationally are English language learners. The range in the districts is from 1 percent in Baltimore City to 40 percent in Los Angeles.

Students in lower-income families—those eligible to participate in the National School Lunch Program—amount to 48 percent of fourth-graders nationally. The range in the districts is from 47 percent to 100 percent in Charlotte and Cleveland, respectively.

For comparison, we also compute the population percentages for all of the nation's large city public school students—those attending schools in cities with a population of at least 250,000. The large city percentages for the characteristics I just described are uniformly higher than the national percentages—in 2009, for example, Black students constituted 29 percent of students in public schools in the large cities, and 42 percent were Hispanic students. These large city percentages include students in both the TUDA districts and other large cities as well.

The 2009 science assessment was based on a new framework, which emphasizes four science practices—identifying science principles, using science principles, using scientific inquiry, and using technological design—that describe *how* students use their scientific knowledge by measuring what they *are able to do* with the science content. The framework also provides an increased focus on the understanding of science principles and asks students to answer questions that cut across the three science content areas, which I'll describe in a moment. Because this assessment is based on a new framework, the 2009 results cannot be compared to the previous science TUDA assessment in 2005.

The science content areas in 2009 were physical science, life science, and Earth and space sciences. At grade 4, equal emphasis is placed on all three content areas. At grade 8, the emphasis on Earth and space sciences was increased to 40 percent, while the emphasis on the other two content areas was reduced to 30 percent each.

For science TUDA in 2009, we have results for public school students at grades 4 and 8. The sample sizes ranged from about 900 to 2,200 students per district per grade. The variation in sample size among districts was large because the districts themselves differ

dramatically in size. For all of our districts, regardless of size, we have samples that provide reliable results.

We report student performance in two ways: average scale scores and percentages at or above achievement levels. NAEP scale scores indicate what students know and can do and are reported on a scale ranging from zero to 300 for science. The achievement levels, which were developed by the National Assessment Governing Board, set standards for what students should know and be able to do. For each subject and for each grade, the Governing Board has established standards for *Basic*, *Proficient*, and *Advanced* performance. Ultimately, the goal is to have all students performing at or above the *Proficient* level.

In this presentation, I will, for the most part, be comparing the districts' performance to the average performance for public school students nationally and for public school students in large cities. When comparing scores and other NAEP results, I will only discuss differences that are statistically significant.

Results for Grade 4

Students at grade 4 answered both multiple-choice and constructed-response questions. The latter questions require students to supply a written response—sometimes a word and sometimes several paragraphs.

The national average score at grade 4 was 149, higher than the large city average of 135 and higher than the average for most of the 17 districts. Austin, Charlotte, and Jefferson

County (KY) had average scores that were not significantly different from the national average. The other 14 districts scored below the national average.

Six districts—Austin, Boston, Charlotte, Jefferson County (KY), Miami-Dade, and San Diego—had scores that were higher than the large city average. Eight districts—Baltimore City, Chicago, Cleveland, Detroit, Fresno, Los Angeles, Milwaukee, and Philadelphia—had scores that were lower than the large city average.

Comparing the percentages at or above *Basic* in the districts to the percentage for large cities, six districts were higher than the large city percentage—the same six that had higher average scores than the large city average. Percentages in two districts—New York City and Houston—were not significantly different from the large city percentage. The remaining nine districts had lower percentages of students at or above *Basic* than did large cities as a whole.

It is important to keep in mind that all of the districts had some students performing at or above *Proficient*, ranging from 33 percent in Charlotte and Jefferson County (KY) to 4 percent in Cleveland and Detroit.

It is useful to go beyond these overall results and examine the performance in the districts of students in various racial/ethnic groups. Nationally, White students had an average score of 162. White students in six districts had an average score higher than 162. In five districts, White students had average scores comparable to the national average, and five districts had lower scores. For one district—Detroit—there were not enough White students to provide reliable results.

For Black students, two districts had average scores that were higher than the national average of 127 for Black students. In seven more districts, the average score was comparable. And in eight districts, it was lower.

For Hispanic students, the average score in one district was higher than the national Hispanic average score of 130. In another ten districts, average scores were comparable, and in the remaining four, they were lower. In two districts there were not enough Hispanic students to provide reliable results.

For Asian/Pacific Islander students, there was no district in which the average score was higher than the national average for those students—160 points. In seven districts, it was comparable. In two it was lower, and in the remaining eight there were not enough students to provide reliable results for this group.

We also compared average scores for each district to the large city average scores for students eligible for free/reduced-price school lunches, and for those who were not eligible. In Austin, Charlotte, and Jefferson County (KY), scores for these two groups were higher than the comparable large city averages. In Baltimore City, Detroit, Los Angeles, Milwaukee, and Philadelphia, scores were lower for these two groups.

Now let's look at one of the science questions used in the 2009 grade 4 assessment. This question, from the Earth and space sciences content area, asked students to choose the type of grocery bag that is best to use to protect the environment. Students could choose from among plastic, paper, and cloth bags.

To receive a "Complete" score on this question, students had to choose one type of grocery bag and correctly explain why using this type of bag best helps protect the

environment. Students could choose any one of the three types of bags and could either argue that the type chosen was superior to the other two or else that the other two types were inferior. For example, one student chose paper and argued that paper grocery bags are best because they are made from trees and “become part of the ground” rapidly, unlike plastic or cloth.

Nationally, 54 percent of students received a score of “Complete.” For large cities, the percentage was 47 percent. In 8 of the 17 districts, at least 45 percent of students received a “Complete” score on this question. In the remaining districts, the percentage was less than 45 percent.

Additional sample questions, covering all three science content areas, are available in the Report Card.

The Report Card also contains a question or item map, showing the NAEP grade 4 science scale running vertically from zero to 300. The questions are listed alongside their corresponding point on the scale, ranked in ascending order of difficulty. Questions from the three content areas are grouped into four achievement-level categories: below *Basic*, *Basic*, *Proficient*, and *Advanced*.

Students scoring at a given point on the scale would be likely to answer correctly the questions at or below that scale point. They would have more difficulty with questions above that point. By looking at the questions that rank at or below the scale score obtained by students in a given district, we can see the kinds of questions students in a district would be likely to answer correctly, on average.

For example, the average score for students in Boston was 139. Students scoring at 139 would be likely to answer correctly questions such as “Recognize an example of a change of state,” “Identify the data on a motion chart,” “Identify the best tool to measure rainfall,” and “Place stages of a life cycle in correct order.”

Grade 8 Results

Now we’re going to look at student performance results for grade 8. Again, I’ll be comparing the scores for the 17 districts to both the national and large city averages.

The national average at grade 8 was 149, higher than the large city average of 134 and higher than all but 1 of the 17 districts. The one district whose score was not significantly different from the national average was Austin. Five districts—Austin, Charlotte, Houston, Jefferson County, and Miami-Dade—had scores that were higher than the large city average. Eleven districts had scores that were lower than the large city average, and lower than the national average as well.

In terms of the NAEP achievement levels—*Basic*, *Proficient*, and *Advanced*—the percentage of students in large cities who scored at or above *Basic* was 44 percent, including 27 percent at *Basic*, 16 percent at *Proficient*, and 1 percent at *Advanced*.

Five districts had higher percentages of their students scoring at or above *Basic* than the large city percentage—the same five with scores higher than the large city average. In one district—San Diego—the percentage was not significantly different from the large

city percentage. The remaining 11 districts had lower percentages of students scoring at or above *Basic* than in large cities as a whole.

We asked grade 8 students about the content of their science classes—whether they were taught materials relating to one, or two, or perhaps all three of the NAEP science content areas. Using their responses, we are able to report the percentages of students receiving instruction in these areas. There is substantial variation among the 17 urban districts.

Nationally, 42 percent of students said they were taught life science at grade 8. Among the districts, the percentages ranged from 31 percent in San Diego to 92 percent in Detroit. For physical science, the national percentage was 81 percent, while the district percentages ranged from 61 percent in New York City to 94 percent in Atlanta. For Earth and space sciences, the national percentage was 56 percent, while the district percentages ranged from 28 percent in Atlanta to 90 percent in Austin.

Now we'll look at one of the questions from the grade 8 assessment. This question is from the life science content area. Students were asked to examine a diagram of a food web. Each arrow shown in the web points from the organism that is consumed to the organism that consumes it.

Students were asked to pick the answer that best explains why decomposers are an important part of the web. The correct answer was “They make nutrients available to plants.”

Nationally, 64 percent of public school students gave the correct answer. For large cities, it was 59 percent. In Atlanta, Austin, Boston, Charlotte, Cleveland, Houston, Jefferson County (KY), Miami-Dade, Milwaukee, New York City, and San Diego, more than 50

percent of students answered this question correctly. In the remaining districts, the percentage was 50 percent or less.

The Report Card contains an item map for grade 8 science. As with the grade 4 item map, questions from each content area are grouped into the four achievement-level categories, below *Basic*, *Basic*, *Proficient*, and *Advanced*.

Again, by looking at the questions that rank at or below the scale score obtained by students in a given district, we can see the kinds of questions students in a district would be likely to answer correctly, on average.

For example, the average score for students in Jefferson County (KY) was 145, placing them above the cut point of the *Basic* range, 141. An average student in Jefferson County (KY) would be likely to answer correctly the questions shown in the below *Basic* range.

District Profiles

Much of the 2009 TUDA Science Report Card is devoted to individual profiles of each of the 17 districts. A variety of information is available through these profiles. For example, there are scores for the district and the state in which the district is located. In addition, there are scores for lower-income students, for both the district and the nation. Each district's scores for the major racial/ethnic groups are shown as well. Finally, the profiles show achievement-level percentages in the district, compared to both large cities and the nation.

Conclusion

That completes my overview of results from the Science 2009 TUDA assessment. There is much more information in the Report Card. In addition, the website

(<http://nationsreportcard.gov>) will give you extensive information on the performance of students in each district, access to released assessment questions through NAEP's Questions Tool, and use of the NAEP Data Explorer, our online data analysis tool.

In closing, I would like to offer my sincere thanks to the students, schools, and school districts that participated in these assessments. Thank you very much.