says Tucker. But Freilich is hesitant to commit.

CCSP is leading an interagency effort to finalize a plan that would for the first time provide a set of government priorities for Earth observations, says Tucker. Scientists say that’s the first step needed to fill the gaps.

In some cases, international partners might help, he says. Japan’s recently launched Greenhouse Gases Observing Satellite, for example, detects CO₂ starting at the ground level, like OCO. Although its resolution is lower, it well complements an existing NASA satellite that only picks up CO₂ readings at an altitude of 5 km. “We’re going to see what we can continue to do with our Japanese partners to fill the gaps,” says NASA’s Stacey Boland.

The effort to better coordinate climate-related monitoring will no doubt be strengthened by the NRC report released last week and presented on 26 February to presidential science adviser John Holdren. Holdren has extolled the virtues of “observation and scientific study of the condition of our home planet’s land, vegetation, oceans, and atmosphere.” Climate researchers are hoping those words, delivered at his confirmation hearing to be director of the White House Office of Science and Technology Policy, will soon translate into support for filling the holes in the nation’s current Earth-monitoring system.

—ELI KINTISCH

**EDUCATION**

**Study Questions Value of School Software for Students**

U.S. students using educational software do no better learning primary school math and first-year algebra than their counterparts who follow a traditional curriculum. That’s the conclusion of a new federally funded study that is loaded with caveats about what it means for students, educators, and the companies that make the software.

The $14.5 million study, funded by the U.S. Department of Education and conducted by Mathematica Policy Research Inc. in Princeton, New Jersey, was designed to find out whether students are benefiting from the growing use of educational software. Preliminary results released in April 2007 (ies.ed.gov/ncee/pubs/20074005/index.asp) suggested that the answer, for first- and fourth-grade students in reading and for sixth- and ninth-grade students in math, was no. But those results were widely criticized by educators and software makers for lumping together the outcomes from many different products and for testing their impact on student achievement too early, in the first year the teacher had used the material.

The second-year results, posted 17 February (ies.ed.gov/ncee/pubs/20094041/index.asp), address those concerns by reporting results for individual software packages and by testing a new cohort of students whose teachers already had a year of using the software under their belts.

What it found is that none of the four math products tested produced a statistically significant difference in achievement between control and treatment groups over the 2-year study, which began in the fall of 2004. (One of the six reading packages registered a meaningful jump in test scores by fourth graders.) Students of teachers using Cognitive Tutor, a computer-based curriculum for algebra I students developed by Carnegie Mellon University researchers (Science, 2 January, p. 64), for a second year showed a meaningful improvement in test scores. But the software had no overall impact on student achievement when data from both student cohorts were combined.

Education research is rarely definitive, however, and this study is no exception. For one thing, the students and teachers were not randomly assigned to the products, explains Mathematica’s Mark Dynarski. That rules out comparing one software package with another. It also means that the results shouldn’t be generalized to different student populations. “Everyone began at different starting points. So we can only say how this product works at a particular school,” he says.

Another confounding factor is that researchers didn’t have enough money to continue observing teachers in their classrooms or surveying them about how they used the product, something that was done in the first cohort. That makes it difficult, for example, to interpret large fluctuations from one year to the next in the amount of time spent using particular software.

Does an increase for algebra I students, for example, mean that teachers found the product more helpful once they became comfortable with it? That seems a logical explanation to Cognitive Tutor’s Steve Ritter, who points to his company’s emphasis on supporting teachers. “For us, usage went up, and second-year teachers achieved statistically significant improvements,” Ritter says. “That’s worth talking about.”

In contrast, does a decline in usage mean that teachers found that it wasn’t right for their students, who are disproportionately from large, poor, and urban schools? Or was it simply due to a dearth of computer facilities, inadequate tech support, or scheduling problems? And what effect does an extremely mobile teacher corps—only one in four taught the same grade at the same school for both years—have on student outcomes?

Dynarski cautions against overinterpretation of the data. He notes that Congress wanted to know whether education software improves student learning, not whether individual products can help certain populations taking certain subjects. “There might be a lot of positive things going on in the classroom, such as greater fluency with computers,” he says. “But it’s not leading to the kind of results that people want, which are better test scores.”

—JEFFREY MERVIS