

The New York Times® Reprints

This copy is for your personal, noncommercial use only. You can order presentation-ready copies for distribution to your colleagues, clients or customers [here](#) or use the "Reprints" tool that appears next to any article. Visit www.nytreprints.com for samples and additional information. [Order a reprint of this article now.](#)



October 8, 2011

Inflating the Software Report Card

By [TRIP GABRIEL](#) and [MATT RICHTEL](#)

The Web site of Carnegie Learning, a company started by scientists at Carnegie Mellon University that sells classroom software, trumpets this promise: “Revolutionary Math Curricula. Revolutionary Results.”

The pitch has sounded seductive to thousands of schools across the country for more than a decade. But a review by the [United States Department of Education](#) last year would suggest a much less alluring come-on: Undistinguished math curricula. Unproven results.

The federal review of Carnegie Learning’s flagship software, Cognitive Tutor, said the program had “no discernible effects” on the standardized test scores of high school students. A separate 2009 federal look at 10 major software products for teaching algebra as well as elementary and middle school math and reading found that nine of them, including Cognitive Tutor, “did not have statistically significant effects on test scores.”

Amid a classroom-based software boom estimated at \$2.2 billion a year, debate continues to rage over the effectiveness of technology on learning and how best to measure it. But it is hard to tell that from technology companies’ promotional materials.

Many companies ignore well-regarded independent studies that test their products’ effectiveness. Carnegie’s Web site, for example, makes no mention of the 2010 review, by the Education Department’s [What Works Clearinghouse](#), which analyzed 24 studies of Cognitive Tutor’s effectiveness but found that only four of those met high research standards. Some firms misrepresent research by cherry-picking results and promote surveys or limited case studies that lack the scientific rigor required by the clearinghouse and other authorities.

“The advertising from the companies is tremendous oversell compared to what they can actually demonstrate,” said [Grover J. Whitehurst](#), a former director of the Institute of Education Sciences, the federal agency that includes What Works.

School officials, confronted with a morass of complicated and sometimes conflicting research, often buy products based on personal impressions, marketing hype or faith in technology for its own sake.

“They want the shiny new one,” said Peter Cohen, chief executive of [Pearson School](#), a leading publisher of classroom texts and software. “They always want the latest, when other things have been proven the longest and demonstrated to get results.”

Carnegie, one of the most respected of the educational software firms, is hardly alone in overpromising or misleading. The Web site of [Houghton Mifflin Harcourt](#) says that “based on scientific research, [Destination Reading](#) is a powerful early literacy and adolescent literacy program,” but it fails to mention that it was one of the products the Department of Education found in 2009 not to have statistically significant effects on test scores.

Similarly, Pearson’s Web site cites several studies of its own to support its claim that [Waterford Early Learning](#) improves literacy, without acknowledging the same 2009 study’s conclusion that it had little impact.

And [Intel](#), in a [Web document](#) urging schools to buy computers for every student, acknowledges that “there are no longitudinal, randomized trials linking eLearning to positive learning outcomes.” Yet it nonetheless argues that research shows that technology can lead to more engaged and economically successful students, happier teachers and more involved parents.

“To compare this public relations analysis to a carefully constructed research study is laughable,” said Alex Molnar, professor of education at the National Education Policy Center at the University of Colorado. “They are selling their wares.”

Carnegie officials say 600,000 students in 44 states use its products, many taking teacher-led classes three times a week with Carnegie-provided workbooks and spending the other two class periods in computer labs using Cognitive Tutor. The full curriculum can cost nearly three times as much as a typical [textbook](#) over six years.

Officials declined to release annual revenue figures, but Carnegie Learning was acquired in August for \$75 million by the parent of the for-profit University of Phoenix. Carnegie Mellon University, which had retained ownership of the Cognitive Tutor software and licensed it to Carnegie Learning, earned an additional \$21.5 million from the sale.

Steve Ritter, a founder and the chief scientist of Carnegie Learning, said there were flaws in the What Works Clearinghouse evaluations of Cognitive Tutor and disputed the Education Department’s judgment of what makes a worthy study.

“What you want to focus on is more of the why,” he said, “and less of a horse race to find out what works and doesn’t.”

A Carnegie spokeswoman, Mary Murrin, said in a statement that the company used “the data from all studies with varying outcomes to continuously improve our programs.”

Karen Billings, a vice president of the [Software and Information Industry Association](#) — a trade group representing many education companies — said the problem was not that companies overpromise, but that schools often do not properly deploy the products or train teachers to use them. Ms. Billings’s group helped design the field trials, in 132 schools, for the landmark 2009 government study of 10 software products, which was ordered by Congress and cost \$15 million.

Then came the deflating results. The industry “became very hostile,” recalled Mr. Whitehurst, now director of education policy at the Brookings Institution. “It seems to me,” he added, “‘hypocrisy’ is the right word for loving something until the results are not what you expect.”

The Hard Sell

Shelly Allen, the math coordinator for public schools in Augusta, Ga., has seen a lot of curriculum salespeople pass through. She is wary of their sweet words and hard sell.

In June, when representatives from Carnegie Learning visited, Dr. Allen warned: “I just want everybody to know I grew up here. I graduated from here. My children go to school here. When you guys get back where you live, our kids have to still be able to reach goals we set.”

Augusta is famous for its magnolia-shaded National Golf Club, host to the Masters Tournament, but its public schools are typical of struggling urban districts. Three-quarters of the 32,000 students in the district, Richmond County, are black, and 72 percent are poor enough to qualify for the federal lunch program. The mean SAT math score last year was 443, below Georgia’s mean of 490 and the nation’s 516.

Six years ago, the district adopted Cognitive Tutor for about 3,000 students at risk of failing, paying \$101,500 annually to use it. As students work through problems, the computer analyzes their weaknesses and serves up new items until they grasp the skill and are allowed to move on. To a student, the promotional materials say, it feels “as if the software is getting to know her and supporting her like a tutor.”

So when the screen says: “You are saving to buy a bicycle. You have \$10, and each day you are able to save \$2,” the student must convert the word problem into an algebraic expression. If he is stumped, he can click on the “Hint” button.

“Define a variable for the time from now,” the software advises. Still stumped? Click “Next Hint.”

“Use x to represent the time from now.” Aha. The student types “ $2x+10$.”

The software likes this and moves on to highlight a series of questions in green, beginning with, “How many more days must you save to buy a bike that costs \$60?” Using his $2x+10$ formula, the student enters “25.”

After solving several questions of this sort and plotting them on a graph, the student would click “Skillometer” to see how he had fared. A series of forest-green bars would show that he did well labeling axes for his graph, but not so well writing the initial formula.

Moving on, Cognitive Tutor would bump him down to an easier problem: “A skier noticed that he can complete a run in about 30 minutes (half an hour).” The expression relating ski runs to time would be $2x$, with x representing hours.

“Immediate feedback,” Carnegie Learning explains on its Web site, “enables the student to self-correct and leads to more effective learning.”

Augusta officials liked the program enough that when concerns arose last winter that many 11th graders were not on track to pass a new state graduation test, the district asked to expand the software’s use to all 9,400 of its high school students. The company agreed to provide access for no additional charge — temporarily.

“As a company, it makes sense to give you the opportunity to prove it works for all students,” Anita Sprayberry, a regional sales manager, told school leaders. That way, she said, “We can talk about a bigger sale.”

Going forward, Ms. Sprayberry said, the cost would be about \$34,000 for each of the district’s 11 high schools.

In a recent interview, Dr. Allen said she was familiar with the What Works Clearinghouse, but not its 2010 finding that Cognitive Tutor did not raise test scores more than textbooks.

Though the clearinghouse is intended to help school leaders choose proven curriculum, a 2010 [Government Accountability Office survey](#) of district officials found that 58 percent of them had never heard of What Works, never mind consulted its reviews.

“Decisions are made on marketing, on politics, on personal preference,” said Robert A. Slavin, director of the Center for Research and Reform in Education at Johns Hopkins University. “An intelligent, caring principal who’d never buy a car without looking at Consumer Reports, when they plunk down serious money to buy a curriculum, they don’t even look at the evidence.”

Evaluating Curriculums

Founded in 1998 by cognitive and computer scientists along with math teachers, Carnegie Learning is proud of its academic heritage, and many education researchers consider it a model of rigor and transparency.

One founder, John R. Anderson, received the 2011 Benjamin Franklin Medal in Computer and Cognitive Science for work on how humans perceive, learn and reason. The company's Web site promises that its curriculums "provide the research-based foundation for proven results," citing "success stories" from around the country.

At Dundalk Middle School in Baltimore County, Md., for example, Carnegie Learning says that Cognitive Tutor led to an increase in the passing rate on a state assessment, to 86 percent in 2004 from 49 percent in 2002. What it does not say is that the rate remained at 85 percent last year, even though Dundalk dropped Cognitive Tutor in 2007 because of difficulties arranging lab time.

That is why many academics dismiss case studies: it is too easy for slices of data to be taken out of context, or for correlation to be confused with causation.

Instead, the gold standard of education research is a field trial in which similar groups of students are randomly assigned to classes where one uses the curriculum and the other does not.

The Carnegie Web site lists five such trials and says they all show positive results for Cognitive Tutor.

Three of these studies, however, were rejected by the What Works Clearinghouse for flaws in their design; in a fourth, the clearinghouse identified a problem with part of the study — the part that purported to show benefits. One of the rejected studies had found that users of Cognitive Tutor in 10 Miami high schools scored better on Florida state exams than a control group, but the clearinghouse found that the students being compared were not equivalent.

"The entire 'effect' of Cognitive Tutor possibly can be traced to other factors," said Mark Dynarski, a former director of the clearinghouse, "and the way in which the research was carried out does not allow one to know if this is the case."

Dr. Ritter, Carnegie's chief scientist, noted that the clearinghouse's 2010 review was limited to high schools and that a year earlier it found that Cognitive Tutor had "potentially positive effects" in middle school.

The middle school finding rested on one study, out of 14 reviewed. That study is featured prominently on the Carnegie Web site, which omits mention of two others that the Education

Department judged to be well designed but showed no benefits.

Dr. Ritter said he had excluded those studies, in Hawaii and Virginia, because the students had not used Cognitive Tutor precisely as the company intended. The researcher who did the Hawaii study, Dennis Newman, said it reflected how Cognitive Tutor was used in the real world.

Dr. Newman is also the author of research guidelines for the Software and Information Industry Association, where Dr. Ritter sits on the education research working group. One of those guidelines states, “An expectation in the scientific community is that research findings are made available regardless of the result.”

Karen Cator, a former Apple executive who directs the Office of Educational Technology at the Department of Education, said the clearinghouse reports on software should be “taken with a grain of salt” because they rely on standardized test scores. Those tests, Ms. Cator said, cannot gauge some skills that technology teaches, like collaboration, multimedia and research.

Ms. Cator’s office is developing a new framework to measure the educational value of technology, but she advised schools and districts not to wait to invest in software like Cognitive Tutor.

“They know what their students need to know and what they need to be able to do,” she said.

Real-Time Assessments

In Augusta, Dr. Allen, the math coordinator, said her district did not have the means to study the effectiveness of Cognitive Tutor formally. But she and her staff saw that low-achieving students who used it were able to join mainstream classes. And teachers appreciated the way the software transmits assessments in real time to Carnegie Learning, then kicks back a report indicating the strengths and weaknesses of each student.

Teachers “just didn’t know, skill by skill, the same type of data they are getting now,” Dr. Allen said.

On the other hand, when the new state math test was given in March, 27 percent of the district’s 11th graders did not pass, which Dr. Allen described as “something that makes us not real excited.”

At the June meeting with Carnegie Learning’s sales team, Dr. Allen said Cognitive Tutor could be worthwhile if the district, which has recently cut \$7 million from its budget and furloughed employees for nine days, could scrape together the financing. “Our negotiations are intense because we don’t have any money,” she said to laughter around the table.

In Georgia, where the state negotiates prices with publishers, an annual license for Cognitive Tutor software is \$32 per student, and the workbook, which must be replaced annually, is \$24 — for a total

of \$336 over six years, a typical lifespan of a math textbook that costs about \$120.

Ultimately, Dr. Allen's district did not have the money, so she focused on getting the most out of her staff. "Giving them the right tools and resources certainly helps," she said, "but our teachers are the ones making that difference."

Gregory W. Capelli, co-chief executive of the [Apollo Group](#), which runs the 400,000-student University of Phoenix and bought Carnegie Learning this summer, said his company first ran its own pilot project with the software and also examined independent research.

But Mr. Capelli, like others, relied at least in part on personal experience.

"My daughter, who's in eighth grade, used this product," he said.

"She would do very well" in some lessons "and not in others," Mr. Capelli said. "What I liked about it is that once she got it, it would allow her to go on to the next part of the tree."



MORE IN TECHNOLOGY (2 OF 38)

**Bits Blog: Behind the
Salesforce Cloud Bra**

[Read More »](#)