As the recession crimps education budgets, states are beginning to pare the number of standardized tests they give, particularly those that no longer factor into state or federal accountability decisions.

At the district level, though, it’s a different story. Despite pressure not to cut staffing and programs, many districts are preserving local "interim" or "benchmark" tests meant to gauge how students are progressing over the course of the year—even though such assessments are not required under the federal No Child Left Behind Act or generally by state officials.

The trend provides insights into how the landscape of educational testing has shifted over the past decade to the more frequent assessment of students. In larger districts such as Los Angeles, it is also keying up debates about the benefits of benchmark tests and whether they can and should be improved.

"I think there’s an expectation that both states and local districts will have some money that can allow them to continue to use, or begin to use, interim assessments," said Stuart R. Kahl, the president and chief executive officer of Measured Progress, a test-maker based in Dover, N.H.

Critics, for their part, lament that the popularity of such testing doesn’t seem to be waning. "It seems that in far too many places, testing has become sacrosanct, even as educators and parents decry its overuse," said Monty Neill, the deputy director of FairTest, a Cambridge, Mass., group that advocates less reliance on standardized testing.

Reductions From Above

Some standardized testing is indeed protected from the vagaries of the market: The 2002 No Child Left Behind law requires annual testing in reading and math in grades 3-8 and once in high school.

So states are cutting back on related bells and whistles and on other standardized tests that aren’t mandated by the federal law.

Florida, which has budgeted for about $18 million in assessment cuts, will do away with certificates of achievement for high-scoring students, paper score reports, and the state’s computer-based-testing option, for instance.

But it will reap the bulk of its savings, $11.5 million, by jettisoning norm-referenced testing.

Norm-referenced tests, which compare the general performance of individual students against that of their peers, were once common. But the NCLB law caused a shift toward criterion-referenced tests, which measure student mastery of specific curricular objectives.

Faced with its cutbacks, Georgia indicated it would no longer cover more than two years of norm-referenced testing for districts, and they have begun scaling back. The Marietta, Ga., district, which serves 8,000 students, has decided to maintain its norm-referenced test in 2nd grade to determine eligibility for magnet programs, and in the 4th
and 7th grades as preparation for upcoming gateway exams. But it will save about $42,000 by forgoing tests in other grades, according to Debra Pickett, the assistant superintendent for curriculum and instruction.

The North Carolina Senate’s version of the pending state budget bill for fiscal 2010 would cut about $3.6 million from nonfederally required assessments, including nationally norm-referenced tests.

Some assessment experts say such cuts are well founded.

Assessment Glossary

Summative Assessment:
Sometimes referred to as “assessments of learning,” summative tests are typically administered near the end of the year. They are meant to give a picture of students’ mastery of particular curricular objectives. The tests states use to meet state and federal accountability requirements, as well as for high school end-of-course testing, are examples of summative assessments.

Benchmark Assessment:
Also called “interim” or “periodic” assessments, these tests are shorter, standardized forms that cover a limited set of objectives within a specific time frame such as six weeks. Districts administer them for a variety of purposes, including to diagnose problems, evaluate the efficacy of particular instructional approaches, and predict performance on end-of-year summative tests. The data from these tests, like the results of summative assessments, can be aggregated and reported beyond the classroom level. Critics say these tests, which are often purchased from commercial vendors, are frequently confused with formative assessments.

Formative Assessment:
Known as “assessments for learning,” these exercises are not used for high-stakes purposes or reporting, but are short measures embedded in lessons as part of instruction. They give real-time, immediate feedback to teachers about gaps in student learning relative to a discrete instructional goal so that teachers can vary their teaching approaches. They change depending on individual students’ needs.

SOURCES: Center for Assessment; Education Week

“If the purpose is to determine how states are doing nationally, then they should be using [the National Assessment of Educational Progress], because it is a much better assessment than any NRT,” said Scott Marion, the associate director of the Dover, N.H.-based Center for Assessment, an assessment consulting group.

But others worry about related cuts that don’t always show up on states’ line-item budgets.

Thomas Toch, a co-director of Education Sector, a Washington-based think tank, suggested that as states pare their programs, they might also further reduce the number of assessment employees, even though most states’ testing capacities are stretched. Still others may choose to eliminate constructed-response items on tests in favor of cheaper multiple-choice formats, he said. “You see states cutting back on their testing capacity at the very time the [U.S.] secretary of education is calling for more-rigorous standards and more-rigorous assessments leading up to the reauthorization of NCLB,” said Mr. Toch.

A handful of states have devised creative ways to save elements such as constructed-response, which requires students to supply short written answers, or longer essay sections. Florida, for instance, discarded the multiple-choice section of its writing exam but kept the constructed-response section.
Maine announced in December that it would move to join the New England Common Assessment Program, a multistate testing partnership. "A good chunk of testing costs are fixed costs, like test development, analysis, and reporting, that can be shared equally and yield dramatic savings for the participants," Mr. Kahl said. "We’re going to see a lot more people thinking about it, investigating it."

Cleaving to Benchmarks

Districts seem to be preserving their own programs—particularly "periodic," interim, or benchmark tests that have sprung up in the wake of the NCLB law. Such tests are meant to determine whether students have mastered a set amount of material, and in some cases, to predict whether a student is on track to pass end-of-year tests used by states for accountability purposes.

An Education Week review of news reports found no evidence of reductions in district practices. FairTest officials reported a few local instances, but no hard trend. There are indications that major test publishers could benefit from the up to $100 billion allocated for education in the federal economic-stimulus bill enacted in February. "We’re getting market feedback that the federal stimulus funds will be a plus for testing," Harold McGraw III, the president and chief executive officer of the McGraw-Hill Companies Inc., one of the major test publishers, told investors during an April 28 phone call. Some experts attribute the relative stability of the benchmark-assessment market to the pressure schools are under to raise students’ test scores. "They perceive it to be important to student success on the NCLB tests," Mr. Toch said. "The stakes are so high ... they are continuing to invest in those assessments."

Properly administered, such tests ensure that students have mastered concepts essential for them to move on to new topics and give teachers data for diagnosing which instructional practices need improvement, said Holly Fisackerly, a former Aldine, Texas, principal who is now a program manager at the National Center for Educational Achievement. The center, owned by the Iowa City, Iowa-based test-maker ACT Inc., works with districts to use the tests effectively.

L.A. Tests Debated

But opponents say the interim checkups merely encourage the practice of "teaching to the test."

In Los Angeles, the teachers’ union has encouraged its members to boycott such tests. It disputes the district’s official estimate of $7 million a year spent on those tests. United Teachers Los Angeles estimates the district spends upwards of $100 million when all personnel expenses associated with the tests are factored in. "The party line is that ... [testing] leads to higher test scores. Well if you test someone to death, they’re going to get it eventually," said A.J. Duffy, the union’s president. But Ramon C. Cortines, the superintendent of the 700,000-student district, is standing firm in the face of such protests. All but three of the 35 schools in “program improvement” for repeatedly failing to meet state achievement-growth targets made gains last year, a feat the schools attribute to the use of the benchmark tests, he said. "I look at schools that are fighting me as it relates to periodic assessment, and I do not see that kind of growth," he said.

Mr. Cortines and Mr. Duffy have agreed on some issues, namely that the tests aren’t always well aligned with California standards and that the results sometimes arrive too late to be useful to teachers. They have agreed to discuss additional amendments to the interim-testing program.
But Mr. Cortines added: "I've made it clear that I would not eliminate it. ... It's hard to argue, from my standpoint, with the success that many schools and classes are having [with benchmark assessments]. How do I say no to that?"

Nationwide, such arguments are resulting in a difficult balancing act for administrators. Ms. Pickett, the Marietta, Ga., administrator, said her district also has chosen to preserve benchmark testing for feedback purposes even in the tough fiscal times. Her district, though, has created its assessments in-house.

Every summer, Marietta teachers and coaches are given an opportunity to revise the tests to ensure alignment with the curriculum, an exercise that has helped some teachers—though not all—take ownership of them, she said. "It's a mixed bag," Ms. Pickett said. "Some teachers feel like it is a tool, and some feel that they do enough [classroom] assessment during the day that they don't need it."

Library Intern Jessica Cain contributed research to this report.