One Standardized Tests Provider Looks to Gaming and Personalized Learning to Innovate Exams

When the Opt-Out movement gained traction in 2015, more than 20 percent of New York students (about 200,000) in grades 3 through 8 declined to take state standardized exams, a statistic that raised questions about the future of such testing.

The Educational Testing Service (ETS), a New Jersey-based nonprofit, is one of the largest high-stakes testing providers in the United States. Their assessments are currently used by more than 8.5 million students across two states.

After years of watching debates about their exams unfold, the testing provider says they are finally ready to make public statements about changes they hope will address user concerns.

"We want to be more engaged with the conversation around what needs are not being met by the tests we have," says Joanna Gorin, vice president of research at ETS. "We are hearing from educators that oftentimes our assessments seem so detached from what is going on in the classroom."

Aligning tests with the ways students learn in the classroom is one of Gorin's goals. Her research team has been evaluating ways to consider aspects of personalized learning and gaming into their tests, a project she describes as difficult but necessary.

Help or Hindrance?

Proponents of standardized testing note that the exams have been instrumental in bringing accountability to the education system. They say testing allows for policymakers to view and compare performance data that can be instrumental in identifying gaps in achievement or underserved student populations.

In 2015, several civil rights groups, including the League of United Latin American Citizens, the Disability Rights Education and Defense Fund and the National Association for the Advancement of Colored People signed a statement coming out against the Opt-Out movement in support of testing. "For the civil rights community, data provide the power to advocate for greater equality under the law," the statement reads, "...anti-testing efforts that appear to be growing in states across the nation, like in Colorado and New York, would sabotage important data and rob us of the right to know how our students are faring."

Those opposed to high-stakes exams, however, note the limitations of such tools for accurately capturing learning. One educator from Atlanta, Farhat Ahmed, noted in an opinion piece published by EdSurge that "[standardized test] scores [rarely] correlated to the amount of growth each learner had experienced during our time working together." Other educators complain that high stakes tests hamper creativity in instructional approaches and place undue stress on students. Gorin says she has been speaking with such educators and that is why she is advocating for changes—with some limits.

“One of the things we are working on is, how do we bring more of the rich types of activities and interactions that happen in classrooms to the types of things students are doing on tests,” says Gorin. “That has a lot to do with how we design tests with technology to be more interactive and responsive to students in the same way that another student or teacher would be in the classroom.”

Pushing the Envelope

Gorin says adaptive tests have the possibility of being more engaging for students. It can also paint a picture of how students come to particular answers, so educators know why a student missed a question. This ability to know why a student got an exam question wrong can empower educators to help students learn from their mistakes—an aspect of personalized learning that Gorin says can be applied to standardized tests.

“Those principles are not unique to personalized learning, but I don’t think people associate it with standardized testing because we haven’t had the capabilities and technologies to bring it in there. That is what we are working towards,” Gorin explains.

To study these types of innovative approaches to testing, ETS has partnered with groups such as Game Lab to design activities that can both measure and help develop competencies and skills. One Pokemon-style game they have prototyped together encourages kids to build argumentation skills by asking students to design a robot and make it stronger by building stronger arguments—adding claims and counter-claims to text on a screen.

Such activities may one day find their way to the tests ETS produces. “We are looking push as far as we can in terms of what principles and elements in games, simulations and other types of interactive performance activities we can bring into standardized testing without losing the important quality metrics that we need,” says Gorin.

Yet she is quick to note a raft of concerns surrounding fairness, equity, bias, budgets, structure, logistics and policy, which increase the stakes of any change they choose to implement. This could mean it will be a long time before students see any innovations to their state exams.

Unstandardized Testing

In the meantime, Gorin hopes educators begin the process of breaking down exactly why they are using approaches like personalized learning—whether it’s for the feedback loop or increased student agency. Once they do, ETS plans to circle back with them and use that information to amend specific aspects of its tests.

Gorin does, however, issue a word of caution about the limitations of trying to apply personalized learning ideas to standardized testing given the opposing goals of the two practices.

“When we think about what accountability tests are used for, it is to make comparisons," she says. "It is to say, ‘are students with disabilities being left behind because we see that

they are not being given opportunities, or not seeing the growth that other students are?' If you don't have a way of making these types of comparisons, you have a real civil rights issue."

Personalized learning, she adds, is more "contextualized to the individual," and thus by nature unstandardized. "It makes those comparisons difficult."

In spite of the difficulties and limitations in changing standardized testing, Gorin notes that science exams may be the first to see change. She says that this is possible because Next Generation Science Standards, a widely adopted set of rules detailing science instructional methods, call for instruction and assessment to be focused on the integration of core ideas, cross-cutting skills and the practice of doing science.

"If you really want to measure the practice of doing science, you have to have your test make them do science," says Gorin. Her office has researched methods such as hands-on tasks, classroom-based activities and simulations to measure science proficiency. "Those are harder to standardize...but because it is so aligned with what people want to measure, we feel that's where there will be the greatest openness to change."