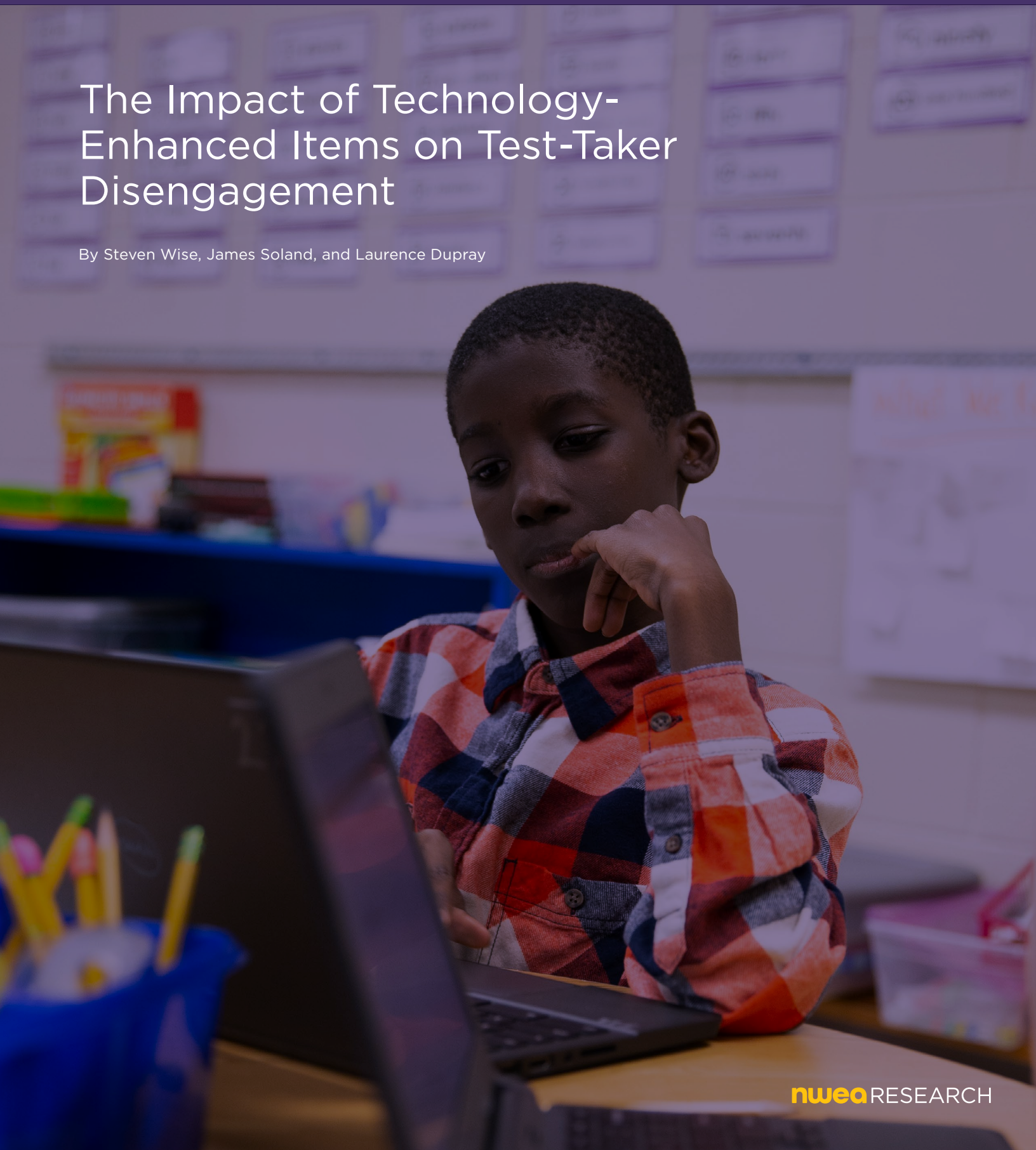
COLLABORATIVE FOR STUDENT GROWTH

BRIEF

The Impact of Technology-Enhanced Items on Test-Taker Disengagement

By Steven Wise, James Soland, and Laurence Dupray



nwea RESEARCH

KEY FINDINGS

- **Technology-enhanced items (TEIs) consistently showed the lowest rapid guessing rate, suggesting that their use may help mitigate the problem of disengaged test taking.** In math, reading, and science, across all grades studied, rapid guesses were more than 10 times more common on multiple choice items than on TEIs.
- **Rapid guessing rates did not differ much between the two types of TEIs examined.** Engagement was higher on both gap match and selectable text TEIs than on multiple select items, which, in turn, had higher engagement than multiple choice items.

Student engagement during an assessment matters: when test takers do not answer questions effortfully, the results do not provide accurate information on what they actually know and can do. Research has shown that low effort and rapid guessing in assessments often occurs and can substantially bias estimates of academic achievement.[i,ii] Recognizing the many problems this can cause, researchers are investigating ways to improve student test engagement.[iii,iv]

Incorporating technology-enhanced items (TEIs) rather than relying solely on multiple choice questions is often suggested as one way to make assessments more motivating and engaging for test takers.

However, research on whether the use of TEIs actually improves student test engagement is limited. Findings of early studies suggest two competing engagement dynamics may be at play: To the extent a TEI is more novel and interesting to interact with than a multiple choice item, test engagement may increase. However, if the TEI appears to be more mentally taxing than a multiple choice question, engagement may actually decrease if the perceived time and energy to complete the TEI are higher than what the student is willing to expend on the question.

To date, most studies on how TEIs impact student effort also have relied on self-reported measures of engagement, typically surveys of test-takers, and most studies were not in an operational testing setting. All of this means that these early results may differ from how test takers actually behave in the real world.
This study examined direct measures of test-taker behavior during operational test events to compare student engagement on three item types: multiple choice, multiple select, and TEIs. Rather than relying on self-reported measures of engagement, the study identified rapid guessing, a type of disengaged test-taking behavior in which a student answers a question so quickly that they could not have understood its content, in these different item types.

Item type	Response action provided by students
Multiple choice	Students select one response from multiple options.
Multiple select	Students select two or more responses from multiple options.
Gap match	Students select an option or options in an area called the toolbar and move or drag these options to designated containers on the screen. The options consist of words, phrases, symbols, graphic elements, or numbers.
Selectable text	Students select a response from within a piece of text or a table of information (e.g., word, section of a passage, number, symbol, or equation), which highlights the selected text.

Item types included in this study. TEIs included gap match and selectable text Items.

Technology-enhanced items (TEIs): computer-based innovative item formats in which test takers engage with questions in different ways, such as dragging and dropping content, highlighting text, or selecting multiple responses. TEIs are thought to measure aspects of student's knowledge of content that cannot be measured using multiple choice items and to improve engagement and motivation among examinees.

The study used item, response, and response time data from MAP® Growth™ adaptive assessments, including data from about three million students in grades 6 through 12 in math and reading, and about 600,000 students in grades 4 through 12 in science (over 138 million, 100 million, and 24 million item responses in math, reading, and science, respectively) to identify and quantify rapid guesses to different item types.

This study addressed three questions:

1. Do TEIs and traditional multiple choice items differ in the levels of engagement they receive?
2. Do TEI engagement levels vary across type of TEI?
3. Does the relative engagement of TEIs and multiple choice items vary across grade?

TEIs had higher test-taker engagement than multiple choice items.

Students showed lower rates of rapid guessing behavior on TEIs than on multiple choice items in all subjects studied. In math assessments, rapid guessed responses were seen in just over two percent of multiple choice items. In contrast, multiple select items showed rapid guesses at less than half that rate, and TEIs showed the lowest rates of rapid guessing, at about ten percent of the rate seen for multiple choice items (0.2 percent for gap match items and less than 0.1 percent for selectable text items).

The findings for the reading and science item responses were similar to those observed for math, with substantially higher engagement on TEI items than on multiple choice items: rapid guesses in science were seen in 3.4 percent of multiple choice items, 0.4 percent of multiple select items, and 0.1 percent of TEIs. The results were similar in reading, though selectable text responses exhibited a higher rapid guessing percentage in this subject, nearly matching the disengaged response rates of multiple select items: rapid guesses were seen in 4.9 percent of multiple choice, 1.7 percent of multiple select, 0.2 percent of gap match, and 1.4 percent of selectable text reading items.

Accuracy rates were consistent with those expected for this adaptive assessment. For multiple choice questions, for engaged responses, 50% of responses were correct; for rapid guesses, 24%, close to that expected by random chance. For TEIs and multiple select items, accuracy rates were lower, as expected, since a greater range of responses is possible. While this could have a distortive effect on scores, the much lower rapid guessing rates of TEIs seem likely to have a stronger positive effect on score validity than their increased distortive effect per rapid guess.

Both gap match and selectable text TEIs improved test engagement; improvements were seen in all grades and subjects studied.

Rapid guessing rates did not differ much between gap match and selectable text items, and neither type showed a lower rate across all three subject areas. While there was some variation in percentage of rapid guesses by grades, the consistent pattern showed that the highest rate of rapid guessing occurred with multiple choice items, followed by multiple select, with the lowest rates consistently observed for TEIs.

RECOMMENDATIONS

Educators and test makers should consider including TEIs on assessments to improve student test engagement and test validity.

If tests are to provide accurate information on what students know and can do, test takers need to be engaged. The results of this study show that TEIs received a much lower rate of rapid guesses than multiple choice or multiple select items, supporting the claim that TEIs are more engaging for test takers. Including TEIs, then, may both provide innovative ways to measure students' knowledge, skills, and abilities on important content domains, and also improve test engagement and score validity. TEIs can be a useful tool, then, for educators and test makers.

More research would be helpful to examine whether more complex TEIs, like simulations or those requiring more sophisticated interactions, also improve engagement. It is possible that if the perceived response effort for more complex TEIs increases too much, engagement could decrease. It is also not clear how engaging an assessment consisting only of TEIs would be: multiple choice items made up a high percentage of the items in the test events studied, so TEIs may have benefited from their novelty.

The pursuit of valid test scores requires consideration both of the types of items administered and the contexts in which test events occur. Measurement practitioners should be mindful of the ways in which each of these factors can influence test taking engagement.

i. Soland, J. (2018). Are achievement gap estimates biased by differential student test effort? Putting an important policy metric to the test. *Teachers College Record*, 120(12), 1–26.

ii. Rios, J. A., Guo, H., Mao, L., & Liu, O. L. (2017). Evaluating the impact of careless responding on aggregated-scores: To filter unmotivated examinees or not? *International Journal of Testing*, 17, 74–104.

iii. Finney, S. J., Sundre, D. L., Swain, M. S., & Williams, L. M. (2016). The validity of value-added estimates from low-stakes testing contexts: The impact of change in test-taking motivation and test consequences. *Educational Assessment*, 21, 60–87.

iv. Wise, S. L., Kuhfeld, M., & Soland, J. (2019). The effects of effort monitoring with proctor notification on test-taking engagement, test performance, and validity. *Applied Measurement in Education*, 32(2), 183–192.

This brief describes research documented in:

Wise, S., Soland, J., and Dupray, L. (2019). The Impact of Technology-Enhanced Items on Test-Taker Disengagement. (Collaborative for Student Growth Working Paper).

Suggested citation:

Wise, S., Soland, J., and Dupray, L. (2019). The Impact of Technology-Enhanced Items on Test-Taker Disengagement. (Collaborative for Student Growth Research Brief).

ABOUT THE AUTHORS

Dr. Steven Wise is a senior research fellow at the Collaborative for Student Growth at NWEA, serves on the editorial board of several academic journals, and has provided psychometric consultation to organizations including state departments of education in Maryland, Virginia, and Nebraska; the National Assessment Governing Board; and the GED Testing Service. Dr. Wise's current research focus is primarily on practical methods for effectively dealing with the measurement problems posed by low examinee engagement on achievement tests. He holds a PhD in educational psychology, measurement and statistics from the University of Illinois at Urbana-Champaign.

Dr. James Soland is a senior research scientist at the Collaborative for Student Growth at NWEA and is an assistant professor at the Curry School of Education at the University of Virginia. His research focuses on assessment and evaluation policy and practice, with particular emphasis on measuring social-emotional learning, test engagement, and estimating teacher and school effectiveness. Dr. Soland completed a PhD in educational psychology at Stanford University with a concentration in measurement and policy.

Dr. Laurence Dupray is the director of assessment products for MAP Growth and MAP® Skills™ at NWEA. She has been working in assessment for over 10 years and focuses on creating high-quality authentic assessments that support evidence-based instructional practices. She holds a PhD in physical organic chemistry from the University of North Carolina at Chapel Hill.

ABOUT THE
COLLABORATIVE FOR STUDENT GROWTH

The Collaborative for Student Growth at NWEA® is devoted to transforming education research through advancements in assessment, growth measurement, and the availability of longitudinal data. The work of our researchers spans a range of educational measurement and policy issues including achievement gaps, assessment engagement, social-emotional learning, and innovations in how we measure student learning. Core to our mission is partnering with researchers from universities, think tanks, grant-funding agencies, and other stakeholders to expand the insights drawn from our student growth database—one of the most extensive in the world.

nwea

NWEA is a not-for-profit organization that supports students and educators worldwide by providing assessment solutions, insightful reports, professional learning offerings, and research services. Visit NWEA.org to find out how NWEA can partner with you to help all kids learn.

OCT19 | KAP4375