

The Impact of Proctor Notification When Students Disengage

By Steven Wise, Megan Kuhfeld, and Jim Soland

KEY FINDINGS

- When students exhibit rapid-guessing behavior during an achievement test, notifying the test proctor was associated with restored engagement and improved test performance.
- These results suggest that when students rapid-guess, they are choosing to not effortfully answer a test item, rather than quickly answering because they don't know the answer.
- Engagement-monitoring testing methods, such as proctor notification, can improve the validity of test scores as indicators of what students know and can do.

When we administer academic assessments, we want to be confident that the resulting scores validly indicate what the students know and can do. This requires a well-designed test, but also assumes that students gave good effort: a valid score requires an engaged test taker. Unfortunately, students sometimes disengage during a test, and do not always try their best, especially when they perceive minimal personal consequences associated with their test performance.

Multiple approaches have been taken to address disengaged test taking. One is to measure test-taking engagement and to note instances of disengagement on score reports to help educators understand how to interpret scores. A second approach is through data management, by either filtering disengaged test takers from datasets, or by statistically adjusting scores to account for disengagement effects. In both approaches, however, remedies are applied after a test event has completed and disengagement has occurred.

When computer-based tests (CBTs) are used, an additional, more proactive approach becomes possible. CBTs can collect response time data, which can be used to identify disengaged test taking. Specifically, for multiple-choice items, rapid guesses have been shown to indicate disengaged test taking. A CBT can monitor item response time throughout a test event, detect in real time when rapid guessing had begun to occur, and then intervene in some way.

Rapid guessing: a disengaged test taking behavior in which a student responds to a test item so quickly that they could not have understood its content. This contrasts with solution behavior, where students try their best on test items.

This study examined the impact of proctor notification, an innovative effort-monitoring feature recently implemented in MAP® Growth™ (an adaptive assessment from NWEA®), on student engagement and performance. When disengagement is detected, the test sends a notification directly to the test proctor, who is then encouraged to personally intervene with the test taker to restore engagement. This intervention is intended to curtail subsequent rapid guessing, resulting in more valid scores.

Using data from over a quarter of a million test events associated with three school districts who use MAP Growth in reading and math, the study addressed three questions: for students who deserved proctor notification, was its availability related to:

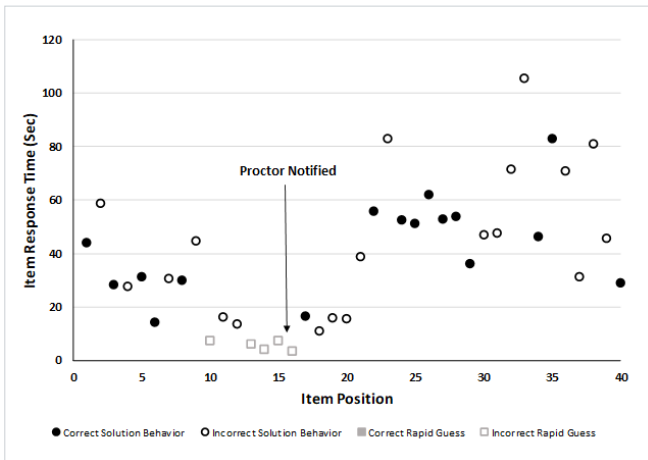
- higher test-taking engagement
- higher test performance
- scores with higher validity

The group of students who had triggered proctor notifications during a fall testing term were compared with those whose test-taking behavior would have triggered a notification the prior year (before this feature was available) to evaluate the impact of the notifications.

Proctor notification increased engagement, performance, and validity

The study results found support for all three questions:

- In reading, post-notification engagement increased by .34 standard deviations (SDs), while in math it was unchanged.
- In reading, the time spent on test items post-notification showed a .54 SD increase. In math, there was a .25 SD increase.
- Post-notification test performance increased by .38 SDs in reading, and by .29 SDs in math.
- Convergent-related validation correlations of test scores were higher by a sizable amount in the year proctor notification was available.



Response time graph for a MAP Growth test event showing gradual decreases in response time and accuracy. After proctor notification, item response time and response accuracy rate increased.

Effort monitoring is one important tool to improve our ability to measure what students know and can do

Although the field of measurement has long recognized the threat to validity posed by disengaged test taking, the past 20 years have seen a surge in research directed at understanding and managing disengagement. This is due in large part to the development of newer, CBT-based measures for detecting rapid-guessing behavior. These types of measures are important to our efforts to proactively reduce the effects of disengagement.

Research on effort monitoring, including proctor notification, provides important insights about student disengagement. It demonstrates that disengagement is a test-taking state that can be altered through intervention. In addition, the increases in test performance associated with effort monitoring-based interventions support the conclusion that when students rapid-guess, they are opting out of engaging with items they are capable of answering correctly, not simply avoiding items they believe they cannot answer. In this sense, effort monitoring advances our ability to measure what students know and can do.

It is important to keep in mind, however, that effort

monitoring is not a panacea. While it has been shown to reduce the amount of disengagement that occurs, it often does not eliminate it completely. Proctor notification should be viewed as an additional tool available to test givers that can curtail disengagement. There are other tools that test givers should be mindful of as well. The instructions given to students, the time of day testing occurs, the physical setting in which testing is done, and the general support for the assessment displayed by test givers to the students are all factors that can enhance test taking engagement.

Less is known about why students disengage in the first place. Test disengagement can often be a sign of disengagement with school more broadlyⁱ, and engagement varies by gender, age, and across racial/ethnic groupsⁱⁱ. Future qualitative research should explore why students disengage in the first place.

ⁱ Soland, J., Jensen, N., Keys, T. D., Bi, S. Z., & Wolk, E. (2019). Are Test and Academic Disengagement Related? Implications for Measurement and Practice. *Educational Assessment*, 1-16.

ⁱⁱ Soland, J. (2018). The Achievement Gap or the Engagement Gap? Investigating the Sensitivity of Gaps Estimates to Test Motivation. *Applied Measurement in Education*. 31(4), 312-323.

RECOMMENDATIONS

Promote computer-based testing practices that enhance student test-taking engagement

Obtaining valid achievement test scores requires students who are engaged and trying to give their best effort. Rapid guessing, which indicates disengaged test taking, can be identified in real time during a test event. This information can be used to send messages—either to the proctor or directly to the student—that can curtail their disengagement and improve test score validity.

Keep test-taking engagement in mind when interpreting a student’s test score

When substantial rapid guessing occurs, the trustworthiness of a student’s test score is diminished. It becomes unclear to what extent a low score is due to poor effort as opposed to low true achievement. This issue is further complicated by the fact that lower achievers are more likely to exhibit rapid guessing. Thus, when material amounts of rapid guessing are present, we should view scores skeptically and re-test students if possible.

Engage in a broader discussion about how to consider disengaged test taking when aggregating scores

The idea that students don’t always give good effort on tests is not new. The ability to identify disengagement by measuring rapid guessing, though, is a recent development. This capability to measure disengaged test taking complicates our interpretations of test results when we aggregate scores to make evaluative judgments about teachers, schools, education programs, etc. If some scores are less trustworthy due to disengagement, how should they be combined (if at all) with the more trustworthy scores? The reality that we can now distinguish scores substantially differing in validity invites a discussion among policy makers about the most valid and fair ways to aggregate scores.

This brief describes research documented in:

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.

Suggested citation:

Wise, S. L., Kuhfeld, M., & Soland, J. (2019). The Impact of Proctor Notification When Students Disengage. (The Collaborative for Student Growth at NWEA 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. His current research focus is primarily on practical methods for effectively dealing with the measurement problems posed by low examinee engagement on achievement tests. Wise holds a PhD in Educational Psychology, Measurement and Statistics from the University of Illinois at Urbana-Champaign.



Dr. Megan Kuhfeld is a Research Scientist II for the Collaborative for Student Growth at NWEA. Her research seeks to understand students' trajectories of academic and social-emotional learning (SEL) and the school and neighborhood influences that promote optimal growth. Kuhfeld completed a doctorate in Quantitative Methods in Education and a master's degree in statistics from the University of California, Los Angeles (UCLA).



Dr. Jim Soland is a Senior Research Scientist at the Collaborative for Student Growth at NWEA, and is an Adjunct Professor of Quantitative Methodology in the College of Education at Oregon State University. 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. Soland completed a PhD in Educational Psychology at Stanford University with a concentration in measurement and policy.



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 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](https://www.nwea.org) to find out how NWEA can partner with you to help all kids learn.

© 2019 NWEA. NWEA and MAP are registered trademarks, and MAP Growth is a trademark of NWEA in the US and in other countries.

MAY19 | KAP3755