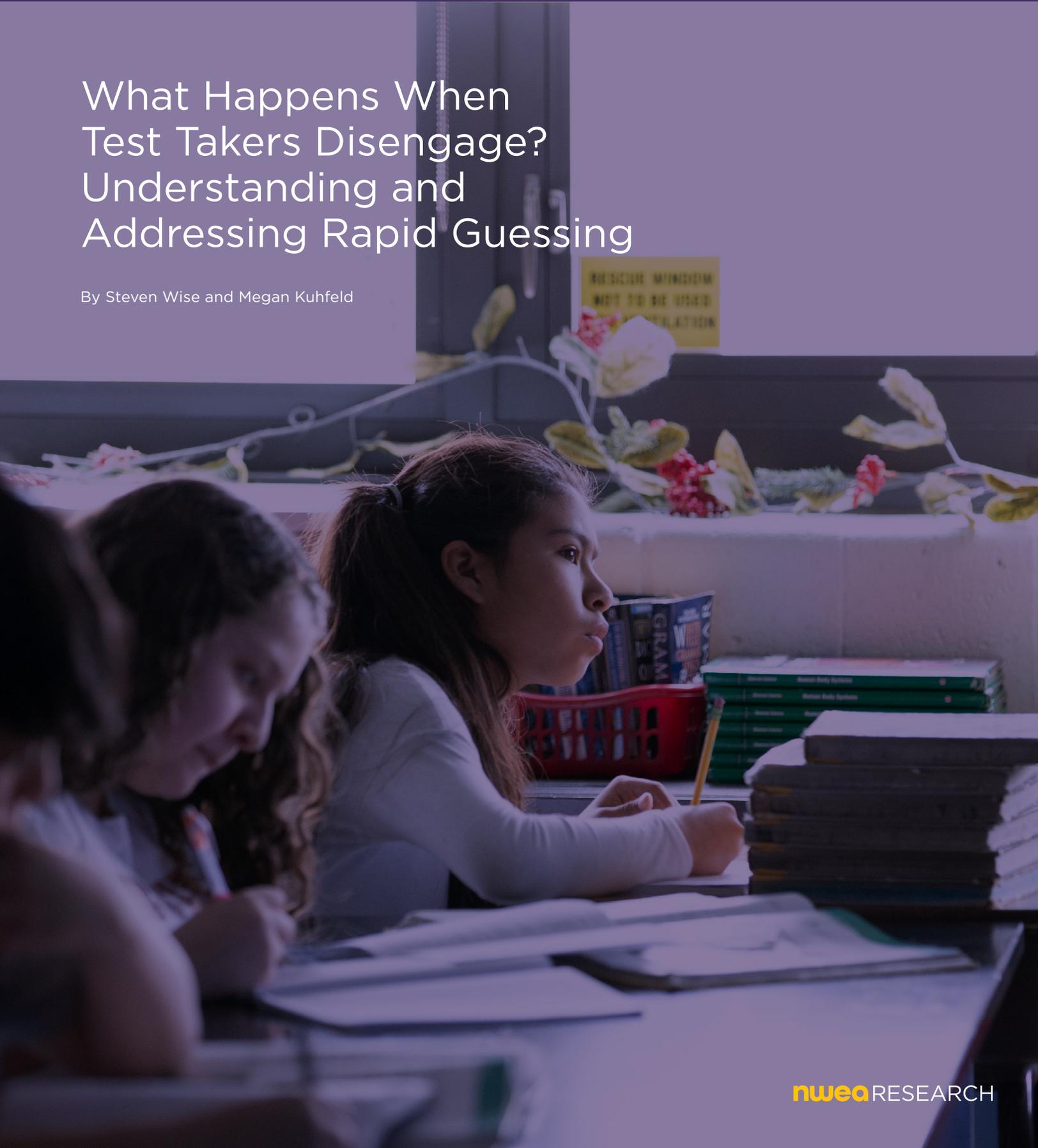


What Happens When Test Takers Disengage? Understanding and Addressing Rapid Guessing

By Steven Wise and Megan Kuhfeld



KEY FINDINGS

- Disengaged responses reflect a momentary pause of measurement. Rapid-guessed item responses, whether correct or incorrect, provide little, if any, information about the test taker’s achievement level. They reflect a choice by the test taker not to apply their skills and knowledge in answering the test item.
- Test takers generally disengage only during a portion of test items. Rapid guessing is higher for items that contain more reading or that are perceived as more taxing, in male test takers, and, at the K-12 level, increases with grade level.
- The choice to include or exclude rapid guesses in scoring depends on whether testing is intended to indicate as accurately as possible what the test taker knows and can do, or if it represents an achievement “hurdle” considered the test taker’s responsibility to clear.

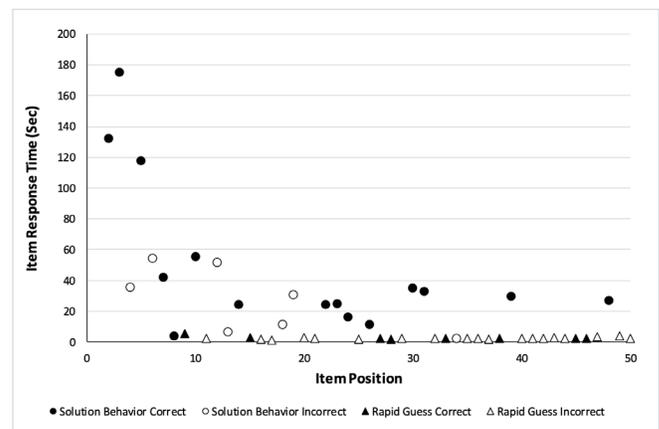
Academic assessments are used to measure educational attainment, assess proficiency, evaluate schools and programs, for certification/licensure, and to inform other important decisions. For test scores to validly indicate what students know and can do, students must give good effort on the assessment. Measurement practitioners have long known that test takers are not always engaged, however, and that disengagement can threaten score validity and negatively bias test scores^{i,ii}. Disengagement has been seen both in high-stakes timed tests, where some test takers rapidly-guess as time runs out, and in untimed, low-stakes tests, where low test-taker motivation is a more likely cause.

Before computer-based tests (CBTs) were introduced, inferences about test taker engagement had to be made at the test event level, most frequently by asking a test taker immediately after testing to report their own level of engagement on the test. With CBTs, new insights are available: item response time permits an item-by-item assessment of engagement through the identification of rapid-guessing behavior. Using data from MAP® Growth™ assessments, an adaptive assessment system for K-12 students, this research illustrates the nature of rapid-guessing behavior, explores how it differs from solution

behavior, provides a model of what happens when test takers disengage, and addresses how disengagement should be managed during scoring.

Test takers may switch between disengaged and solution behavior throughout a test

Research has shown that test takers rarely exhibit rapid-guessing behavior throughout a test, but rather may move multiple times between solution behavior and rapid guessing. Rapid guessing is affected by characteristics of the item, the test taker, and the context in which the item is administered.



Responses from a test event in which there was frequent switching between the solution behavior and rapid-guessing behavior

Response process for rapid guessing is fundamentally different from solution behavior

Understanding how test takers engage with and respond to items is critical for understanding what test scores mean and for test validity. This research

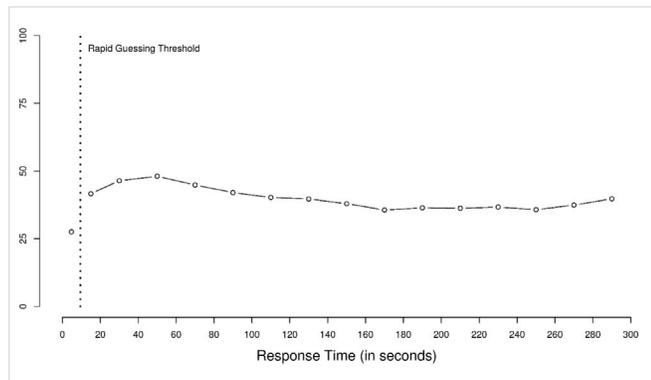
Solution behavior: engaged test taking behavior; student makes an effort to use what she/he/they know and can do to correctly answer a test item.

Rapid guessing: disengaged test taking behavior; student responds to a test item so quickly that she/he/they could not have understood its content.

asserts that the response process underlying rapid-guessing behavior is fundamentally different from that used during solution behavior. It provides three lines of evidence to support this:

1. *Rapid guesses showed accuracy discontinuity.*

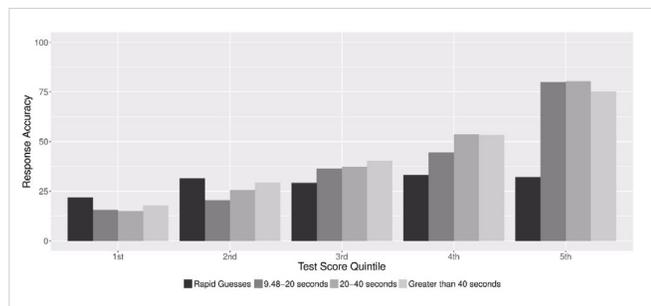
Responses made more quickly than the minimum time required to understand and respond to that item had a markedly lower accuracy rate than responses made over a longer time using solution behaviors. The magnitude of discontinuity was larger in some items than in others.



Discontinuity between rapid-guessing accuracy (left of the dotted line) and solution behavior accuracy for various response time segments associated with responses to a MAP Growth Reading item.

2. *Rapid guesses provided little or no information about achievement.*

Test items are designed to provide information that, across all items on the test, can produce an estimate of the test taker's achievement: higher-achieving students are expected to answer an item correctly more often than lower-achieving students. However, with rapid guesses, this relationship generally did not hold. The psychometric information inherent in the correctness of rapid guesses was very limited.



The relationship between response accuracy and overall test performance for rapid guesses and solution behavior response time segments for a MAP Growth Math item. In solution behaviors, frequency of correct answers increased with overall test achievement. This was not seen for rapid guesses (black bars).

3. *In rapid guesses, test takers are biased toward responses based on position, not correctness or content.*

Under solution behavior, test takers are more likely to select the correct option than any other option. With rapid guessing, however, test takers were more likely to choose options based on their response position rather than correctness. Rapid guess responses were not random; rather, test takers were more likely to choose options in the middle of those listed rather than the first or last option. This contributed to the variation seen in magnitudes of accuracy discontinuity: when option B was the correct answer, rapid guesses were correct about a third of the time, so accuracy discontinuity appeared smaller. When option D was correct, in contrast, rapid guesses were correct only about a seventh of the time, and accuracy discontinuity appeared larger.

A model of rapid-guessing behavior: two choices

To explain what happens when a test taker disengages, whether under time constraints on a high-stakes timed test, or from low motivation where stakes are lower, a model was proposed that involves a test taker making two choices.

The first choice is whether to engage in solution behavior or rapid-guessing behavior. Solution behavior requires slow, effortful, controlled cognition. This requires an intentional choice; by default, people prefer engaging in fast, non-effortful, autonomous cognition. When a test item is given, the test taker quickly weighs their *effort capacity*, the amount of effort they are willing and able to give, against the item's *resource demands*, or how much work the item appears to require to fully answer effortfully. If the test taker's effort capacity is lower than the resource demands of the item administered, they default to non-effortful cognition, and to rapid guessing. Both effort capacity and resource demands may change throughout the test as the test taker becomes bored, fatigued, more interested, or feels increasing time pressure.

Once a test taker makes the decision to rapid-guess, the second choice is selection of a particular response option. While this is often considered a random choice, as other research has shown, and this study illustrates, this is often influenced by biases commonly seen when people are asked to choose randomly. When test takers guess the answers to multiple-choice items, they tend to show *edge aversion*, a common choice bias that appears as test takers showing greater preference for the middle of the set of options, and lesser preference for the first and last option.

What to do about rapid guesses?

Using item response times to identify rapid-guessing behavior provides new insights into test-taking engagement. It shows that a rapid guess represents an item-person encounter during which the test taker chose not to be measured. The resulting item response, whether correct or incorrect, provides little, if any, psychometric information about the test taker's achievement level.

With this new insight comes new questions: should we account for rapid guessing when scoring tests by de-emphasizing or filtering out rapid guesses? If so, when? The choice hinges primarily on context, and whether measurement represents an achievement "hurdle" considered the test taker's responsibility to clear; or if it is intended to indicate with the greatest possible accuracy what test takers know and can do. If the inference to be made focuses on whether the test taker has demonstrated proficiency, and stakes are high for the test taker, including rapid guesses in

scoring may be appropriate, even though these rapid guesses likely negatively distort scores. In this case, if the test score is negatively distorted because of low motivation or running out of time, this could reasonably be viewed as the test taker's responsibility, and the test giver should not feel compelled to correct for the distortion. However, if the inference to be made is not about a particular student's proficiency, but about proficiency rates across a group of test takers, it would generally be advisable to exclude rapid guesses during scoring. Including rapid guesses would likely lead to an underestimation of the true number of proficient test takers—an undesirable result for most test givers. Similarly, in low-stakes tests, where the test-giver seeks the most accurate information on the instructional needs of the test taker, rapid guesses should be excluded from scoring to obtain the most accurate measurement of test taker achievement.

ⁱ Haladyna, T. M., & Downing, S. M. (2004). Construct-irrelevant variance in high-stakes testing. *Educational Measurement: Issues and Practice*, 23(1), 17-27.

ⁱⁱ Wise, S. L., & DeMars, C. E. (2005). Low examinee effort in low-stakes assessment: Problems and potential solutions. *Educational Assessment*, 10, 1-17.

RECOMMENDATIONS

Keep test-taking engagement in mind when interpreting test scores and seek practices that increase test engagement.

While test disengagement is not a new problem, using item response time to identify rapid guessing provides new insights into when and how disengagement occurs. When substantial rapid guessing takes place, the trustworthiness of a student's test score is diminished. These findings underscore the essential role of test-taking engagement in pursuit of valid scores. Research also shows that rapid guessing frequency varies across item types and test contexts. What can we do to make assessments more engaging for students? How can we increase student motivation to make an effort on assessments? Educators and researchers should work together to identify and develop practices that increase test engagement.

Engage in a broader discussion on when and how to account for disengaged test taking in test scores.

We have new ways to measure and account for disengaged item responses. In many cases, however, rapid-guessed item responses are still included in scores even though they provide little or no psychometric information about achievement. The pursuit of accurate scores in most cases implies that rapid guesses should be excluded; achievement hurdles represent an exception for which rapid guesses may be included.

This brief describes research documented in:

Wise, S. L. & Kuhfeld, M. (2019). A cessation of measurement: Identifying test taker disengagement using response time. (The Collaborative for Student Growth at NWEA Working Paper).

Suggested citation:

Wise, S. L. & Kuhfeld, M. (2019). What happens when test takers disengage? Understanding and addressing rapid guessing. (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. 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. 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 masters' degree in statistics from the University of California, Los Angeles (UCLA).



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.

JUN19 | KAP3851