

# **NWEA Guidance on the Interpretation and Use of New Engagement Metrics in MAP Growth Reports**

December 2017

© 2017 NWEA. All rights reserved. No part of this document may be modified or further distributed without written permission from NWEA.

MAP® Growth™ is a registered trademark of NWEA.

Disclaimer: This report is the product of research conducted by NWEA.

## Table of Contents

1. Introduction .....	1
2. Engagement Metrics .....	1
2.1. Percent of Disengaged Responses .....	1
2.2. Estimated Impact of Disengagement on RIT Score .....	2
3. Recommendations .....	2

## 1. Introduction

This document describes the following two new student engagement metrics now included on NWEA™ MAP® Growth™ reports, and provides guidance on how to interpret and use these metrics:

- Percent of Disengaged Responses
- Estimated Impact of Disengagement on RIT

These metrics will inform educators about what percentage of items from a student's test event were rapidly guessed, and what the estimated impact of that rapid guessing was on a student's final RIT score. When combined with the notification system that alerts proctors when students are disengaged, these features should provide educators with actionable information about a student's overall engagement during a test. This information is useful when interpreting student test scores, and provides additional insight around which students might be considered for retesting to maintain the validity and integrity of test results.

## 2. Engagement Metrics

### 2.1. Percent of Disengaged Responses

When a student responds to a test item, the amount of time he or she took to respond can be compared to the response times from other students who answered the same item. This comparison allows us to determine if the student responded so quickly that he or she could not have reasonably read the question, considered the content, and attempted to correctly answer the question.

When this occurs, the student's response is characterized as a rapid guess, or a "disengaged response." Too many disengaged responses can impact the validity of a student's test score. In general, the more disengaged responses in a test event, the greater the risk to test score validity.

The "percent of disengaged responses" metric provides information about the overall percentage of operational items from a student's test event identified as disengaged responses. Most students do not have any disengaged responses. For these students, their reports will currently show an "N/A", indicating that they did not rapidly guess on any of their test items. When students provide disengaged responses on less than 10% of items during a test event, the resulting impact on test score validity is generally minimal. Having these students retest is likely not necessary; however, it is recommended to have a conversation with these students about the importance of trying their best on these tests.

However, in situations where students rapidly guess on 10% or more of their items, additional steps may need to be taken, including decisions around whether to retest the student. Retesting decisions based on this metric can be informed by how much of an impact that disengagement had on the student's estimated RIT score (described in the next section). A school's written test policy should provide guidance on how to proceed with retesting based on this metric. A student's test event will automatically be invalidated, starting in July of 2018, if the student provides disengaged responses on 30% or more of test items.

## **2.2. Estimated Impact of Disengagement on RIT**

The greater the number of disengaged responses during a student's test event, the greater the impact on the validity of a student's final RIT score. This is because if a student does not try on a test item, the likelihood is high that the student's response to that item will be incorrect. For example, on an item with four answer options, a student who randomly guesses will have an approximate 75% chance of answering the item incorrectly. This is not particularly problematic when only a few items are rapidly guessed. However, when there are a high number of disengaged responses, the overall impact on a student's RIT score can be pronounced.

One way to assess this risk is to estimate the impact disengaged responses had on a student's test result. This new metric quantifies the estimated impact of disengaged responses on a student's score by re-estimating a student's RIT score with all disengaged item responses removed and calculating how much this estimate differs from the student's demonstrated RIT score based on all responses. This difference represents the "estimated impact of disengagement on RIT."

For example, a value of -3 on a report means that a student's disengaged responses might have reduced his or her actual score by 3 points. As this number gets further from 0, there will be a greater impact on the validity of the student's score. This can result in less accurate student achievement estimates and could affect decisions made about students, teachers, or schools that are based, at least in part, on a score that does not reflect reasonable effort by the student.

## **3. Recommendations**

In July of 2018, our reporting system will automatically invalidate test events in which students provide disengaged responses on 30% or more of test items. This is because when this much rapid guessing occurs, it severely limits our ability to estimate a credible score for a student. If a student shows this much disengagement during a test event, we recommend that the student be retested.

For other situations, our broad guidance would be that when a student provides disengaged responses on less than 10% of the items, retesting is likely not necessary. The decision on whether or not to retest a student when the proportion of disengaged responses is between 10% and 30% is a matter of professional judgment for educators. In circumstances where the assessment is administered for stakes—for example, when it is used for teacher or school evaluations—educators might consider retesting students with this level of disengagement in order to get accurate measures for all students. In these circumstances, the criteria for retesting should be established in a written policy and followed consistently at all testing terms. To offer a second example, in circumstances where a student's eligibility for a program may have been impacted by disengagement—for example, a high-performing student falling just below the cut score for a gifted program because he or she disengaged on 20% of test items—the district might also consider retesting to ensure that students who are qualified for a program have the opportunity to participate. Conversely, in circumstances where the results are not used for high stakes purposes, an educator might choose to not retest a student if the overall impact of disengagement on a student's RIT score was minimal.

Finally, NWEA strongly recommends against combining the overall impact of disengagement on a student's RIT score to the student's reported score, and using that estimate as the student's "official" RIT score moving forward. This is especially true in situations when student scores are used as the basis for high-stakes decisions. For example, extending the example of assessing a student for placement in a gifted and talented program, let's assume eligibility was based on a student scoring at or above a cut score of 200. A student's demonstrated achievement level was a score of 198, but the student provided disengaged responses on 20% of test items, and the overall impact of disengagement on the student's score was -4 points. In this situation, we would not recommend that the school interpret this student's score as 202 (198 + 4). Instead, NWEA's general guidance would be to discuss the need for good test effort with the student, retest the student to improve the overall validity and reliability of the achievement estimate, and then make a placement decision for that student based on this new RIT score.