



# NEBRASKA

## DEPARTMENT OF EDUCATION

### Interpreting Student NSCAS Growth Scores

How the computer adaptive design impacts understanding student scores

As a District Assessment Coordinator (DAC), you play an important role in supporting district and school staff as they administer NSCAS Growth and use the data from the assessment. As school leaders and educators dig into their student's results from this spring's NSCAS Growth administration, the Nebraska Department of Education (NDE) wants to equip DACs with information on interpreting student NSCAS Growth scores.

NSCAS Growth is a computer adaptive assessment designed to both provide students a chance to answer items above or below grade level to allow them to demonstrate where they are in their learning while also ensuring there are enough on-grade level items for summative purposes. Computer adaptive tests select items that are tailored to a student's performance throughout the assessment. Items get easier if a student answers an item incorrectly and harder if a student answers an item correctly. As a student responds to items, the computer adaptive assessment better understands where a student is in their learning.

**More personalized and more precise than linear or fixed form assessments**

**Administers each item based on student response and item difficulty**

**Items can be difficult or easy for students even within grade**



NSCAS Growth generates two different scores that aim to provide comprehensive information on a student's results – the NSCAS scale score (summative) and RIT score (diagnostic).

- The NSCAS scale score and associated achievement level is based on a subset of items (about 27 items) that are on-grade and aligned to the state summative blueprint.
- The RIT score is based on all items that a student took during that test (about 27 on-grade-level items plus about 13 items that adapt on- and off-grade-level as needed).



### 3-8 Tests Adapt

- Only on-grade for summative scores
- Across all grades, though no more than two grades above student's enrolled grade for diagnostic scores

## Adaptivity and Item Selection: Diagnostic

On the diagnostic portion of NSCAS, the test selects the item based on the difficulty level of the items – not the grade-level of the item. But a student may see an easy item that is aligned to standards two grade levels above or they may see a difficult item that is any number of grade levels below. High school standard-aligned items are treated as one grade-level for purposes of choosing items for a student. A student in grades 7 and 8 could see items aligned to 9-12th grade-level standards based on the difficulty of that item.

**The resulting scores a student receives are calculated based on a mixture of student responses and item difficulty.** Because students take different items through the assessment, comparing just the total number of items correct from student to student is like comparing apples to oranges.



# What does this mean when looking at a student's scores?

With students receiving different items because of the computer adaptivity of the test, there are nuances to each student's results that will explain their specific achievement level and RIT scores.

## What contributes to a student's scores?

Specific set of items the student sees and their item difficulty

Student's results on specific categories (e.g., numbers and algebra in mathematics)

Total number of points the student received on an item

Number of diagnostic items

Students engage with EASY, MEDIUM, and HARD items through their test. A classical way to describe item difficulty is in terms of the proportion of kids who get a correct answer, particularly if you give the same test (same set of items) to the same group of students, then you have a meaningful basis for comparing easier items from more difficult items.

In our adaptive tests, different students see different items on the test, so we can't make the above kind of description of item difficulty – however, the way we define the difficulty categories is analogous to the above description. We use a psychometric model for item difficulty that expresses the item difficulty in terms of how likely a student is to correctly answer the question based on their ability level. For the EASY, MEDIUM, and HARD categories we set those thresholds in terms of the barely proficient student – a hypothetical student whose final test score is right at the cut score for proficiency. The difficulty criteria for the test items are then defined as:

★  
EASY

The barely proficient student has a **75%** probability or better of correctly answering.

★★  
MEDIUM

The barely proficient student has between a **40% and 75%** probability of correctly answering.

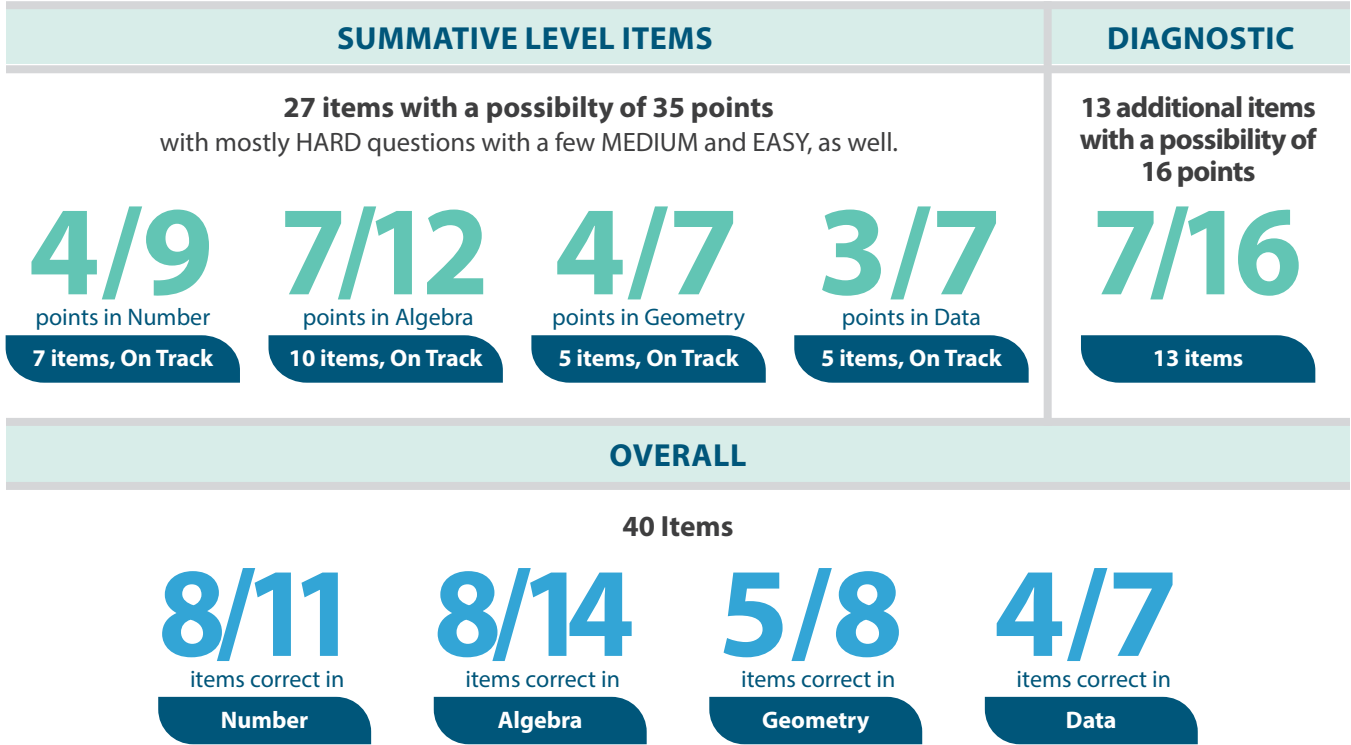
★★★  
HARD

The barely proficient student has a **40%** probability or less of correctly answering.

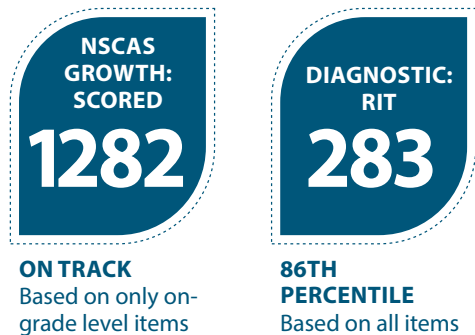
## Student A:

Student A is in 6th grade and took the NSCAS Growth mathematics assessment.

### Student A's Responses



### Student A's Results



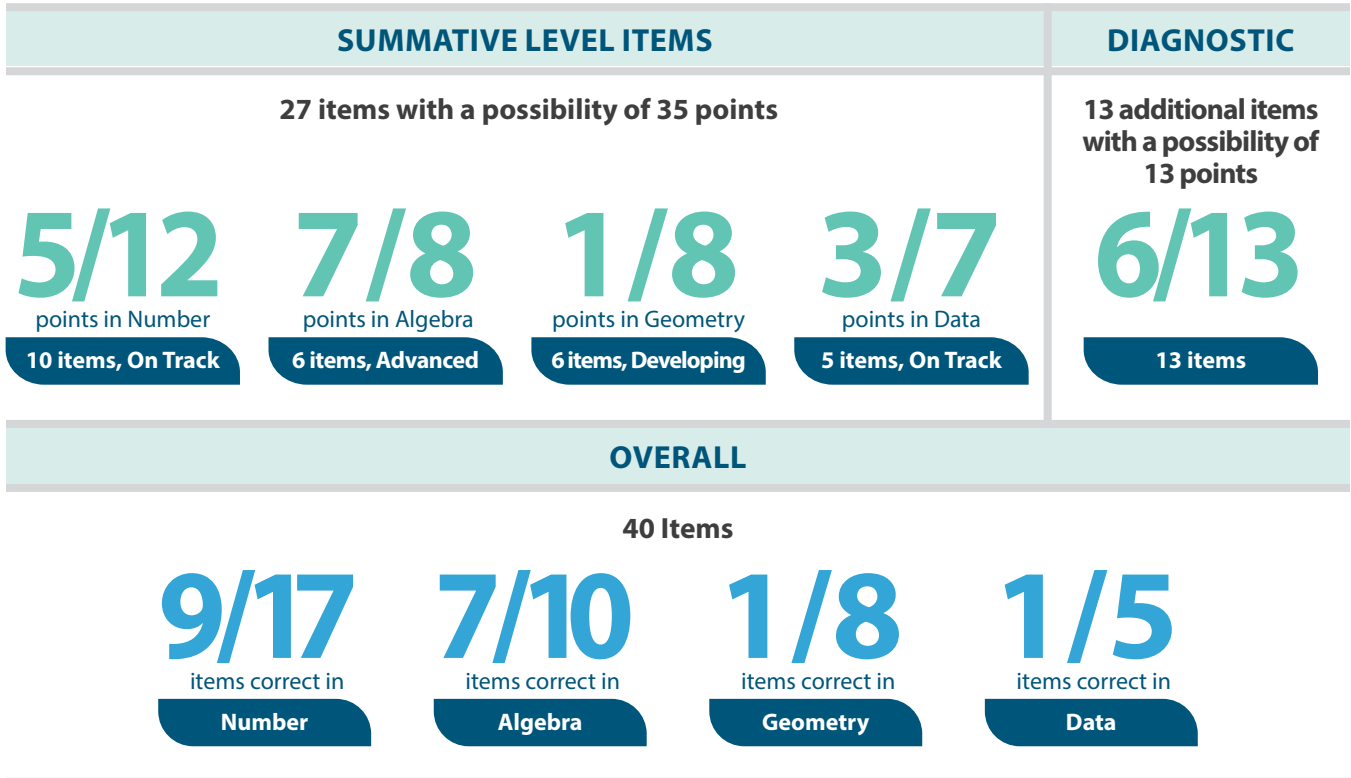
### Understanding Student A's Scores

On this test, Student A received 27 items that contributed to their NSCAS score and received 18 out of 35 points from answering the 27 items correctly or partially correct. Even though this student answered 23 items correctly overall, the student only received 18/35 points from the NSCAS portion of the exam, and the item difficulty might have been lower than other students.

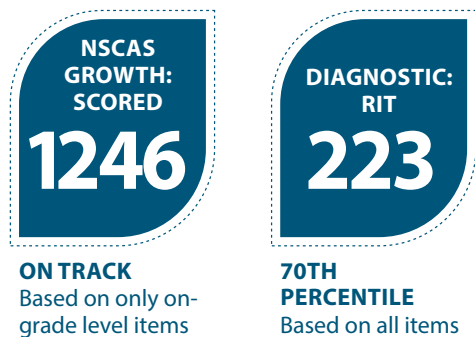
## Student B:

Student B is in 5th grade and took the NSCAS Growth mathematics assessment.

### Student B's Responses



### Student B's Results



Note: NWEA understands the concern and is targeting to implement increasing the Standard Error of Measurement (SEM) reporting to the nearest hundredth for State RIT results starting in fall 2023 as part of ongoing development for MAP Growth Information from NSCAS Growth.

### Understanding Student B's Scores

On this test, Student B received 27 items that contributed to their NSCAS score and achieved 16 out of 35 points for that portion of the assessment. Student B was able to achieve "Advance" in Algebra, "On Track" in Number and Data and "Developing" in Geometry. To understand Student B's placement for Data from seemingly only answering one item, it's important to understand this student achieved full credit from one 2-point item and partial credit (i.e., 1 point) from another 2-point item, achieving 3 points out of 7 points. Also, although this student's performance on Geometry items is not proficient, whose percentage is 22%, overall performance can be proficient, as this student is proficient on the other three areas. In addition, Student B received an additional 13 diagnostic items that contributed to their RIT score.