



Pearson

Transcend™  
Cumulative  
Blueprint  
and Growth  
Overview

TRANSCEND

## Designs for Growth

Many interim assessments claim to support inferences related to learning gains, often referred to as *growth*. For an assessment to facilitate valid inferences related to growth, examinees must be assessed by the instrument over multiple occasions and the *construct model*, or what is being measured, must be stable over the specified time frame. There are two types of learning gains to consider when the construct model is of a large grain size, such as a subject-level construct (e.g., Mathematics) or a domain-level construct (e.g., Operations & Algebraic Thinking): *within-year growth* and *between-year growth*. Within-year growth describes how learning progresses across an academic year, while between-year growth aims to describe the trajectory of student achievement over multiple academic years. Although designing for between-year growth inferences is an attractive attribute of an assessment, it often requires students to encounter items aligned to off-grade standards, impacting the instructional validity of the test scores. Because of this, Transcend has adopted a design that focuses on within-year growth.

To compute within-year growth, Transcend adopts a model-based gain score calculations. Simple gain scores are the difference between two scores at two different points in time. When there are three or more scores for an individual student, the gain can be generalized over multiple time points by averaging and expressing progress as an average change per unit of time. The Transcend gain score model facilitates descriptions of growth based on the available scores as well as growth predictions to future time points for individual students and class, school, and district aggregates.

Transcend is the first assessment to adopt a *cumulative blueprint* for facilitating within-year growth inferences. The cumulative blueprint improves the validity of growth-related insights by distinguishing between the reasons why a score might have changed: individual learning vs. curriculum progression (opportunity to learn). Prior to describing the cumulative blueprint and how it improves the validity of inferences related to learning gains, this paper describes two popular interim blueprints: the *differing blueprint* and the *common blueprint* design.

To describe the relationship between a specific test blueprint and growth, this paper uses three Grade 3 Mathematics interim assessments aligned to the Common Core State Standards as an illustrated example. For each blueprint, this paper walks through a series of diagrams that represent the fall, winter, and spring administrations. While the blueprint differs across examples, the test assembly and construct model remain consistent. This paper adopts the Transcend test assembly and construct model, which are described in Figure 1.

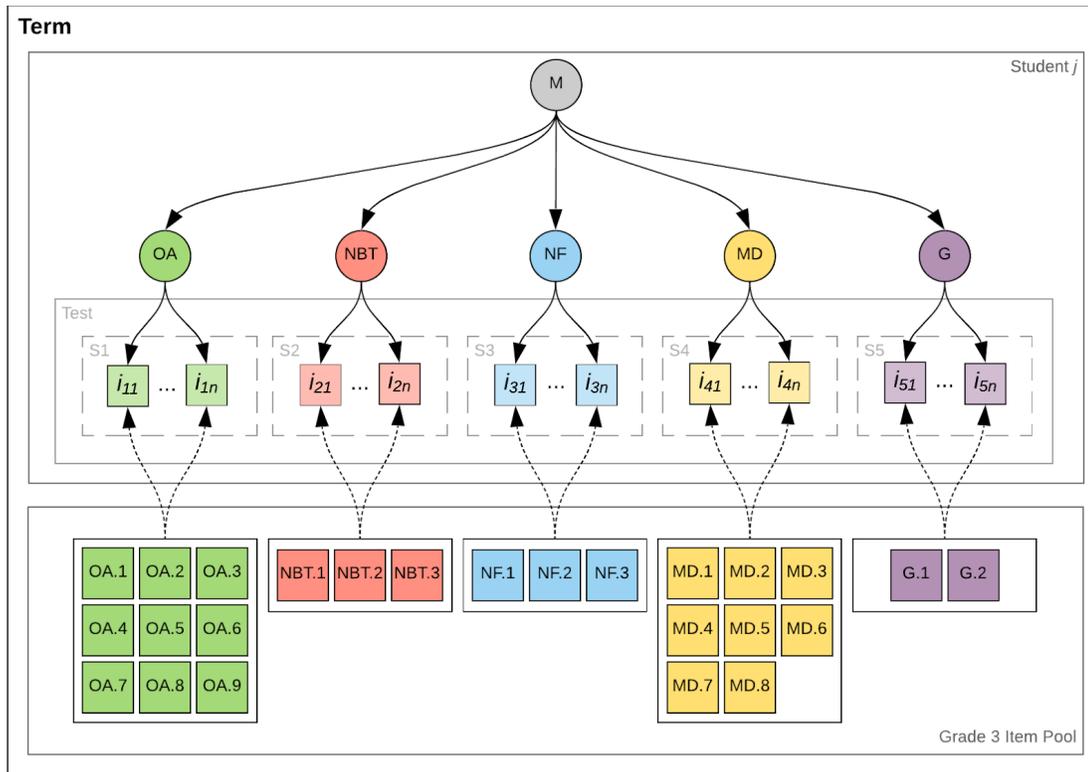
The bottom of Figure 1 represents the Grade 3 item bank. The representation includes one box for each mathematics standard<sup>1</sup>. Notice, the standards are partitioned into *domain-specific* item banks where the nine green “OA” boxes represent *Operations & Algebraic Thinking* standards; the three red “NBT” boxes represent the three *Numbers & Operations in Base Ten* Standards; the three blue “NF” boxes represent the three *Numbers & Operations—Fractions* standards; the eight yellow “MD” boxes represent the eight *Measures & Data* standards, and the two purple “G” boxes represent the two *Geometry* standards.

Next, the bottom component of the “Student  $j$ ” box labeled “Test” represents the personalized assessment. The Transcend assessment design groups items by domain-specific sections to lower the cognitive complexity caused by random shifting between domains that can confound measures of domain understanding. That is, Transcend reduces the cognitive load associated with *taking a test* to focus on the measurement of the target constructs. These sections are represented by the dashed boxes labeled “S1” through “S5.” Within each test section, items are represented by a color-filled box labeled “ $i$ ” with a subscript to indicate the section number and item number. To represent the variable length, we use an  $n$  subscript to indicate the last item within the test section.

Dotted arrows connect each domain-specific item bank partition to a domain-specific test section of Student  $j$ 's test. These arrows represent the adaptive item selection algorithm that builds a personalized assessment based on a student's domain understanding.

The final aspect of the diagram is the Transcend *construct model*. The directed graph of round nodes represents each construct being measured by the Transcend assessment. The grey node labeled “M” represents the subject-level construct, mathematics. Solid arrows connect the subject-level construct to the domain-level constructs. That is, we say Grade 3 mathematics is a composite construct made up of the five Grade 3 mathematics domains. Next, solid arrows connect the domain-level constructs to the specific items on the domain-specific section of the Student  $j$ 's test. That is, we say Student  $j$ 's domain understanding elicits their responses to the items on the domain subtest.

<sup>1</sup> The diagram adopts a *domain.standard* format (e.g., NF.2) which differs from the official *grade.domain.cluster.standard.sub-standard* format (e.g., 3.NF.A.2.A)



**Figure 1.** Transcend assessment design using Grade 3 Common Core mathematics standards.

## Differing Blueprint

Interim assessments that adopt a differing blueprint will include only items aligned to standards that were taught within a specific unit of instruction. The fall assessment depicted at the bottom of Figure 2 was designed to assess those standards addressed by the district’s curriculum during the fall instructional unit. For this particular district, items associated with the first eight Operations & Algebraic Thinking (OA) standards were eligible for selection, as were the first two standards from the Numbers & Operations, Fractions (NF) domain, and standards 4 through 8 from the Measures & Data (MD) domain. As a result of the available items, the fall measurement model included domains OA, NF, and MD which contribute to a composite construct defined generally as *Grade 3 Fall Mathematics*.

During the winter months, the curriculum progressed to cover standards in the Number & Operations, Base Ten (NBT); Number & Operations, Fractions (NF); and Measures & Data (MD) domains. Note, however, that while NF and MD were recurring domains, the actual standards that made up the domains differed from fall to winter. In the fall curricular unit, NF standards 1 and 2 were addressed while in the winter curricular unit, standard 3 was addressed. Likewise, MD standards 5 - 8 were addressed during the fall instructional unit while MD standards 1-4 were addressed during winter instructional unit. Thus, the test blueprint for the winter interim differed from the fall interim blueprint. The items available from

the winter blueprint resulted in a winter construct model including domains NBT, NF, and MD which contributed to a composite construct defined as *Grade 3 Winter Mathematics*.

Finally, the spring curricular unit addressed the remaining Operations & Algebraic Thinking standard (OA.9); all three Number & Operations, Base Ten (NBT) standards; and both Geometry (G) standards. As a result, the spring interim blueprint differed from both the fall and winter. As a result, the construct model included OA, NBT, and G domains, and a composite construct defined as *Grade 3 Spring Mathematics*.

Each assessment derived from the differing blueprint design in Figure 2 is of high instructional validity. Students were only assessed on content standards they have had the opportunity to learn and the resulting scores aligned to what had been taught in the classroom during that instructional unit. However, there are two aspects of a differing blueprint that limit inferences. First, none of the assessments provided an understanding of how students perform on the broader, *Grade 3 Mathematics* construct, the target construct of state summative assessments used for federal accountability. Second, the design prevents valid growth inferences as the construct model is unstable over the academic year.

The shift in available standards within and between domains from occasion to occasion resulted in an unstable construct model. Figure 3, a detail of Figure 2, illustrates an unstable construct at the domain-level. Measures & Data was assessed in fall and winter, one cannot make any inferences as to whether or not a student has made learning gains in the MD domain since the items associated with the domain at each occasion differed. That is, a student may struggle to demonstrate understanding of MD.5 through MD.8 in the fall, and thus, earn a low score on the fall MD subtest. However, that same student may understand MD.1 through MD.4 when it is taught in the winter, for which they earn a high score on the winter MD subtest. Because the two subtests are assessing different content related to the MD domain, one cannot make a valid inference about how that student's understanding of the MD domain has changed between fall and winter as he or she may still struggle to understand the second half of the MD standards (MD.5 - MD.8) at the time of the winter assessment.

This irregularity of construct definition across occasions extends to the composite construct over time. Recall that each occasion has a newly defined composite construct, as illustrated in Figure 4. None of these composite constructs carry a Grade 3 Mathematics definition, but a temporally specific subset of it. Because of this, any trend in the series of composite construct scores over time cannot be interpreted as growth.

## Common Blueprint

To facilitate inferences pertaining to learning gains, interim assessments employ a *common blueprint* design. Figure 5 illustrates the same Transcend assessment design as Figure 2, but now uses a common blueprint. That is, items aligned to every Grade 3 standard are available

for selection on the fall, winter and spring test. As a result, the construct model is stable over the academic year.

The common blueprint has long been considered necessary to make inferences related to growth because the construct model is stable over time. However, such a blueprint is of limited instructional validity as students are tested on the same content at each occasion, regardless of whether or not the curriculum has covered that specific material—the student is assessed on content he or she has not yet had the opportunity to learn.

Consider the student who encountered an item that asked them to “[u]se place value understanding to round whole numbers to the nearest 10 or 100” (3.NBT.A.1) during the fall assessment, before the third grade mathematics curriculum was designed to cover the topic. The student answers the item incorrectly. Was the incorrect answer because the student had trouble understanding the concept, or because they have not yet had an opportunity to learn the concept? Thus, the scores from an assessment that adopts a common blueprint may not provide instructionally relevant information.

Opportunity to learn jeopardizes the validity of the growth inferences made from such an assessment as well. Again, consider the same student, but now they take the winter test and another item aligned to 3.NBT.A.1 appears. Now, they answer it correctly. Is this evidence that the student has made learning gains, or is it evidence of what the curriculum had implemented? That is, the student never had any problem with using place value understanding to round whole numbers to the nearest 10 or 100, rather, they never had the opportunity to learn the concept until the winter term. Thus, while the common blueprint is seen as a requirement for valid inferences related to learning gains, scores from common blueprints conflate learning gains with opportunity to learn.

It is important to realize that this illustration used a grade-specific item bank resulting in students only seeing items aligned to grade-appropriate standards. Many interim assessments, however, combine a common blueprint with a grade-band item bank. This is typically done to support claims of between-year growth. Such a combination exacerbates the conflation of learning gains with opportunity to learn as students are potentially confronted with items aligned to off-grade standards.

## Cumulative Blueprint

The differing blueprint is of high instructional validity because it assesses only what has been covered in the curriculum, but is unable to facilitate inferences of learning gains because the construct model is unstable. Common blueprints are of low instructional validity because students are assessed on items that have not yet been covered in the curriculum, but claim to measure growth. However, such growth measures conflate curriculum progression and learning gains.

To optimize instructional validity and valid inferences pertaining to learning gains, Transcend developed the cumulative blueprint. The cumulative blueprint design was developed by considering how a construct model expands as a student progresses through the curriculum, and how this expanding construct model retains its stability through the expansion. That is, a construct model is defined within the context of what has been made available to learn and the expanding definition, while non-constant, is not unstable.

In Figure 6, the fall assessment was defined by a subset of standards, and only the domains associated with those standards are part of the construct model. In the winter, the item bank expanded to include newly addressed standards in addition to those standards addressed in the fall. Likewise, the construct model expanded to include those new domains. Finally, in the spring, after the curriculum has addressed all Grade 3 standards, all items in the item bank are made available and the construct model represents the complete Grade 3 construct.

Because curriculums cover more material over time, the construct to be measured is then expanding over the course of an academic year. Thus, it is the cumulative blueprint that expands as the construct expands, as defined by the curriculum over the course of the academic year. Such a design separates learning gains from opportunity to learn, providing a purer measure of learning growth.

Transcend™ Cumulative Blueprint and Growth Overview

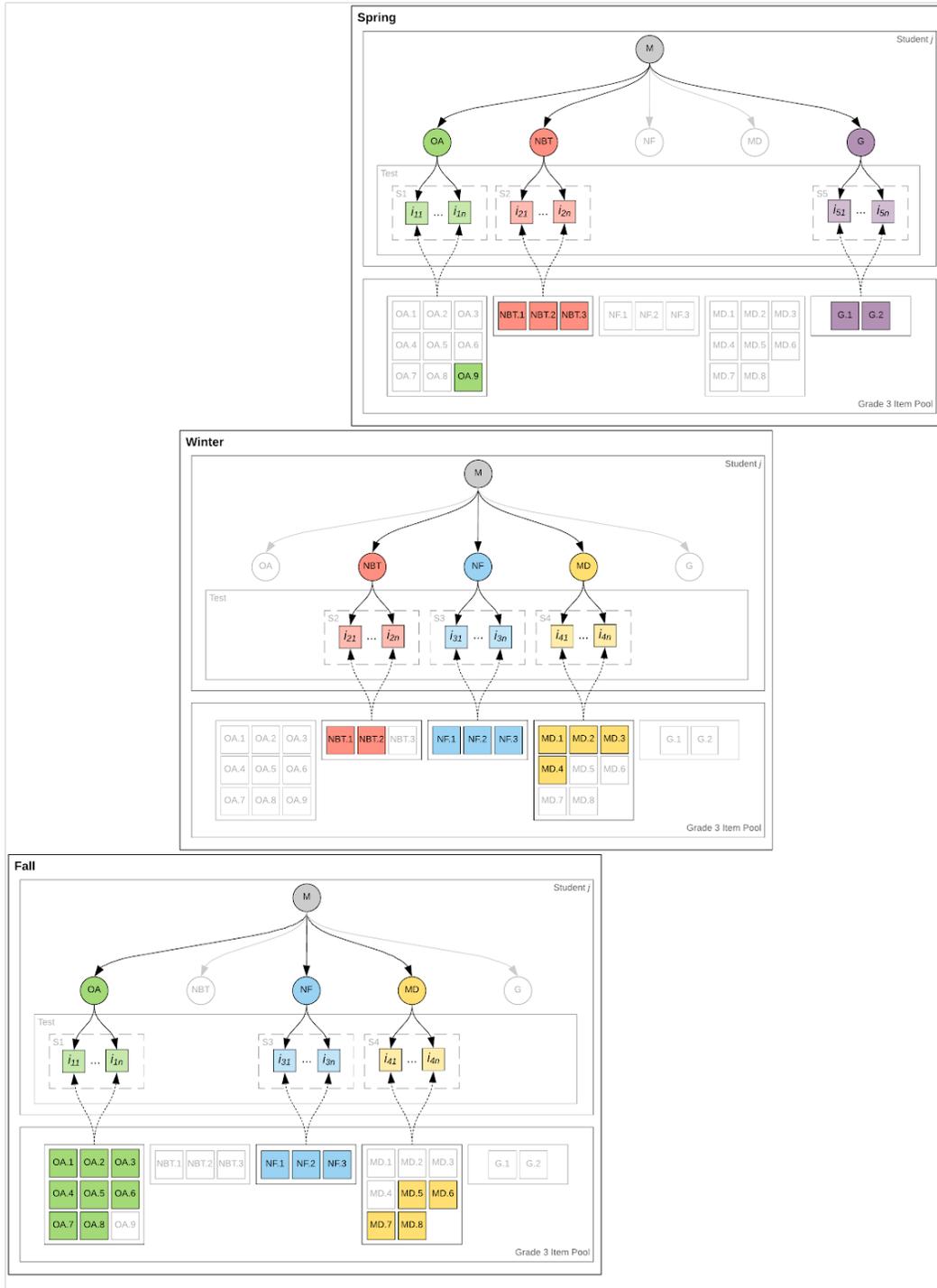


Figure 2. Transcend assessment design with a differing blueprint and Grade 3 Common Core mathematics standards.

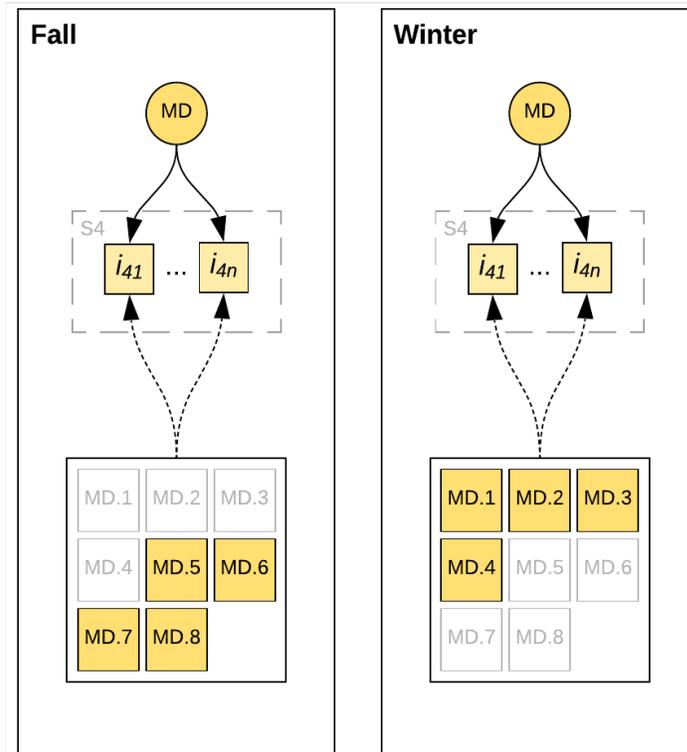


Figure 3. Detail of the fall and winter Measures & Data domain from the differing blueprint

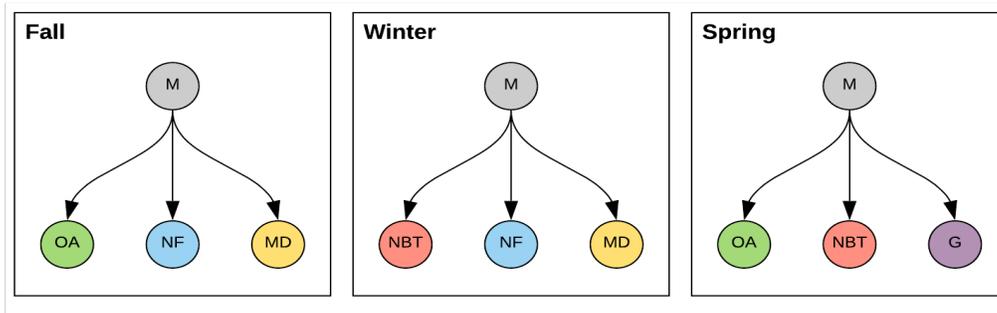
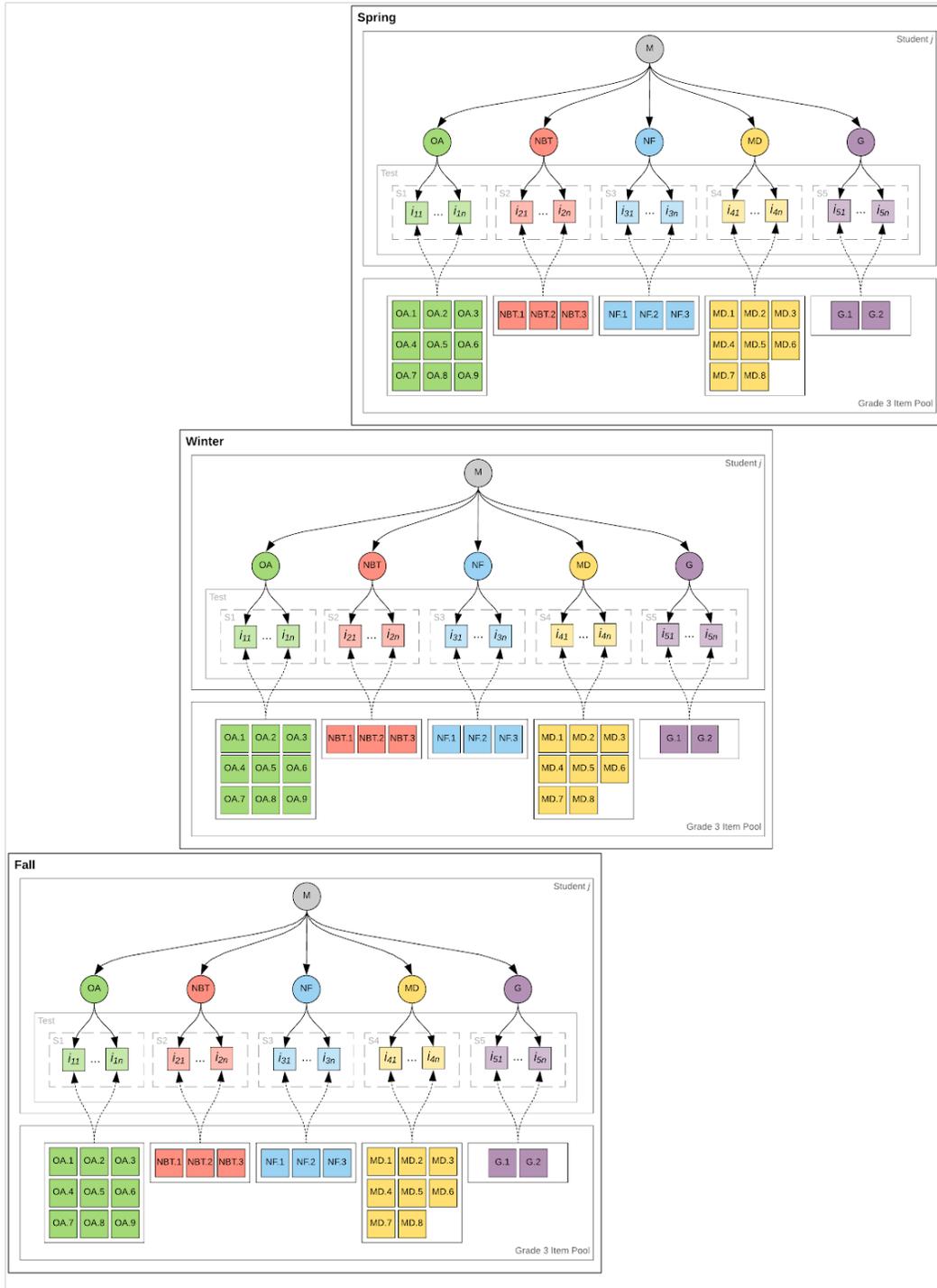


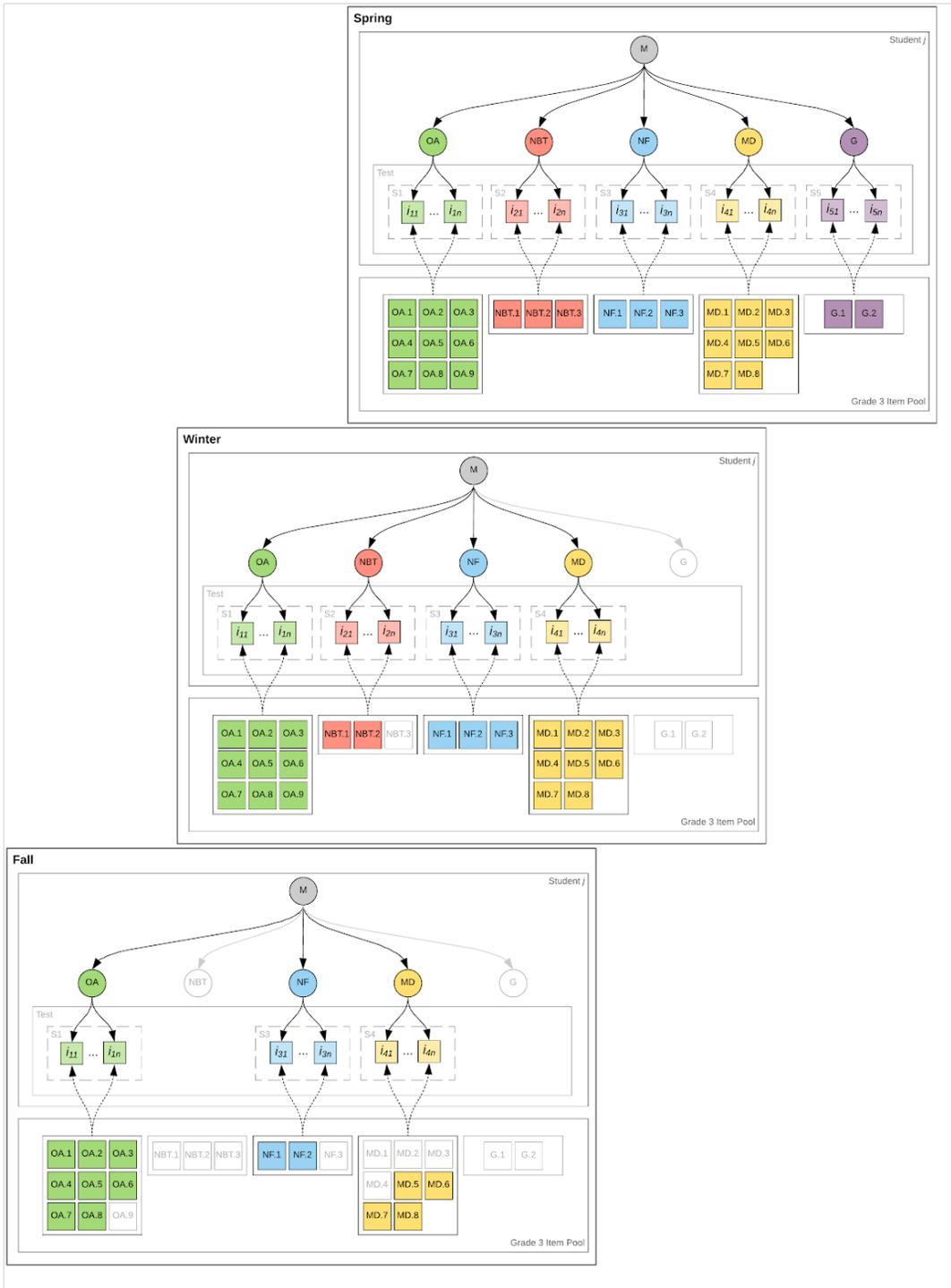
Figure 4. Detail of the construct models for fall winter and spring differing blueprint

Transcend™ Cumulative Blueprint and Growth Overview



**Figure 5.** Transcend assessment design with a common blueprint and Grade 3 Common Core mathematics standards.

Transcend™ Cumulative Blueprint and Growth Overview



**Figure 6.** Transcend assessment design with a cumulative blueprint and Grade 3 Common Core mathematics standards.