SCHOOL OF INFORMATION
UNIVERSITY OF MICHIGAN

Learning Analytics (680)

Office Hour 3

N. Sheltrown

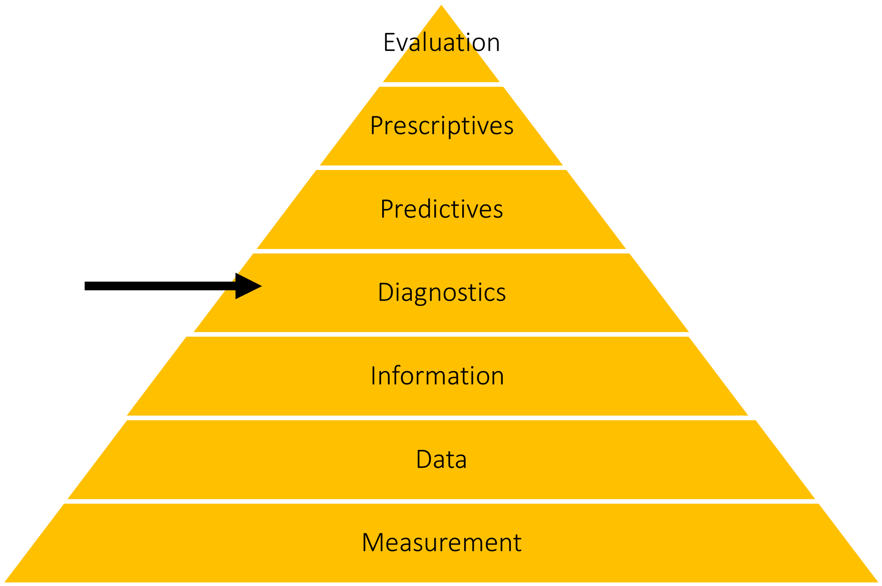
Agenda

- Clustering Example [teachers]
- Learning analytics pyramid: diagnostic analytics

Learning Analytics:

Diagnostic Analysis

The Analytics Pyramid



The Diagnostic Layer

Diagnostics is analyzing the information we have for either positive or negative aberrations that should lead us to adjust our course of action. A blood test may reveal that your body has a problematically low vitamin D level, which may present a problem to your health. Such a diagnosis should lead to an adjusted course of action (take vitamin D pills or spend more time in the sunshine). In learning analytics, we seek to diagnosis challenges (deficiencies in student learning) and opportunities (situations where students are ready for above grade-level content) that should lead to a material adjustment of the student's instructional treatment. Similar to medical diagnosis, the primary focus of the activity is to identify abnormalities that can lead to suboptimal outcomes. Information from the diagnostic layer may lead an educator to adjust a student's learning path to more appropriately reflect his/her particular strengths and weaknesses.

Item Analysis

- Provides information about what students do and do not know
- Highlights different response patterns that will require different instructional strategies (whole class, small group)

CORRECT ANSWER	2	1	4	2	3	4	4
POINTS POSSIBLE	1	1	1	1	1	1	1
%CORRECT	60%	58%	53%	65%	26%	40%	32%
ITEM ANALYSIS							
% Choosing A or (1)	32%	58%	9%	13%	32%	13%	18%
% Choosing B or (2)	60%	26%	23%	65%	10%	23%	8%
% Choosing C or (3)	4%	14%	13%	13%	26%	22%	40%
% Choosing D or (4)	3%	0%	53%	8%	30%	40%	32%

- Item Analyses appear at the bottom of all **roster** reports

Diagnostic analysis defined.

A diagnosis is a means by which we identify the root cause of a problem or illness. Traditionally, it begins with suboptimal conditions such as an illness. The symptoms of the illness are analyzed through diagnostic tests to determine the cause so that some form of treatment can be pursued. Once a patient has a medical diagnosis, the patient will receive a prescription to help remedy the ailment.

In general, diagnosis and prescription work are also closely linked: diagnosis identifies the problem; prescriptive analysis identifies solutions

Diagnostic analysis defined for LA applications.

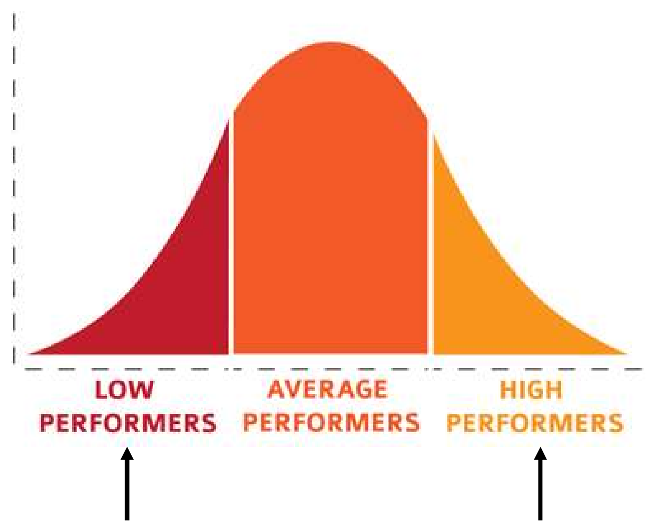
Learning analysts seek to diagnosis both

- *learning deficiencies* (students with learning gaps that prevent them from effectively learning grade-level content) and
- *learning opportunities* (those students who are already proficient at grade-level expectations and would benefit from above grade-level content and resources).

Diagnostic analysis defined for LA applications.

Diagnosis works a little differently in learning applications.

Rather than focusing exclusively on deficit situations (where student performance is below the desired performance) diagnostic analysis seeks to find any meaningful differentiation from the norm – whether students are performing above or below grade-level content expectations. Unlike health care, where an exceptional health would not regularly lead to medical intervention, in learning analytics we always seek to find the unusual because any deviation from the standard educational treatment warrants differentiation.



How Diagnostic Analysis Usually Happens in a K-12 Context

	Gender	Math	Reading
Abagail Adams	F	99	100
Cindy Circumference	F	64	79
Elsa Evengreen	F	68	96
Farro Fontaine	F	85	98
Hemal Hamptom	F	89	93
Ila Ing	F	65	80
Jenny Jasper	F	95	89
Lila Lancer	F	72	82
Nancy Natopole	F	73	84
Ramona Roberts	F	73	88
Syliva Smith	F	99	99
Tona Tanger	F	95	94
Uma Uccello	F	86	90
Valeria Valance	F	91	98
Wendy Weis	F	81	89
Barry Bellowington	M	94	87
D'shawn Deat	M	89	76
Gary Gary	M	79	60
Kyle Kong	M	85	98
Marty Morrin	M	75	100
Oscar Omen	M	87	80
Pedro Pascal	M	92	74
Quincy Quatar	M	95	93
Xavier Xi	M	75	87
Yani Yello	M	99	80
Zach Zelrelli	M	90	66

	Math	Reading
Max	99	100
Min	64	60
Range	35	40
Mean	84.4	86.9
Standard Deviation	10.6	10.3

		Math	Reading
Female			
Max		99	100
Min		64	79
Range		35	21
Mean		82.3	86.9
Standard Deviation		10.9	7.0
Male			
Max		99	100
Min		75	60
Range		24	40
Mean		87.3	81.9
Standard Deviation		8.1	12.6

In these tables, we discover a 5 point difference in math and reading scores between boys and girls (the girls outperform boys in reading, but underperform in math). A half standard deviation difference in the scores may seem like a notable finding, but is it? Two reasons suggest not.

Problem1. Subgroup Analysis Invites Discussion of Meaningless Differences

This problem has been best illustrated in medical research. In his discussion of subgroup analysis, epidemiologist Ben Goldacre highlights an article published by a group of researchers in the medical journal Circulation.

In the study, the researchers selected over 1000 patients with coronary disease from a database and randomly assigned them to one of two treatment groups. After creating two randomized groups, the researchers didn't offer the patients any new treatment, but did collect follow-up data on their progress to see what would happen. In their analysis, they found that the two groups did not differ significantly in their survival rates (as you would expect), but subgroup analysis revealed that a certain subgroup of coronary disease performed significantly better in the first treatment group than the second. Normally, this would be a key finding—were it not for the fact that there was no difference in treatment. The subgroup differences were the result of random chance.

Was this a fluke finding? Not really. This phenomenon is closely connected to the idea of "insensitivity to sample size." Mathematically, we know that the smaller sample, the greater the variation in the data. This means that a small collection of data is much more likely to give an extreme average either much higher or much lower than the rest of our data.

In some cases, this might be an example of *experiment-wise error.*

Clinical Judgment and Statistics

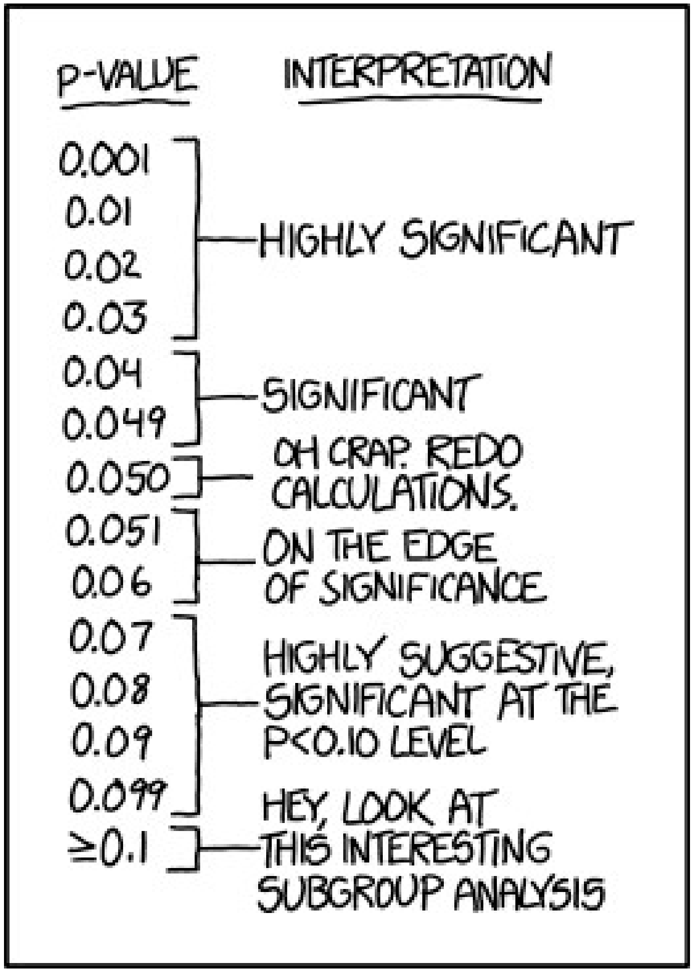
Lessons from a Simulated Randomized Trial in Coronary Artery Disease

KERRY L. LEE, PH.D., J. FREDERICK MCNEER, M.D., C. FRANK STARMER, PH.D., PHILIP J. HARRIS, M.B., D.PHIL., AND ROBERT A. ROSATI, M.D.

SUMMARY A simulated randomized clinical trial in coronary artery disease was conducted to illustrate the need for clinical judgment and modern statistical methods in assessing therapeutic claims in studies of complex diseases. Clinicians should be aware of problems that occur when a patient sample is subdivided and treatment effects are assessed within multiple prognostic categories. In this example, 1073 consecutive, medically treated coronary artery disease patients from the Duke University data bank were randomized into two groups. The groups were reasonably comparable and, as expected, there was no overall difference in survival. In a subgroup of 397 patients characterized by three-vessel disease and an abnormal left ventricular contraction, however, survival of group 1 patients was significantly different from that of group 2 patients. Multivariable adjustment procedures revealed that the difference resulted from the combined effect of small imbalances in the distribution of several prognostic factors. Another subgroup was identified in which a significant survival difference was not explained by multivariable methods.

These are not unlikely examples in trials of a complex disease. Clinicians must exercise careful judgment in attributing such results to an efficacious therapy, as they may be due to chance or to inadequate baseline comparability of the groups.

CLINICAL JUDGMENT in chronic illness involves a knowledge of the natural history of the disease, the ability to assess the validity of therapeutic claims and a means of applying what is known to the individual patient. Much has been written about the natural history of angina pectoris.[1-10] The literature has emphasized the importance of multiple factors in determining outcome and of the heterogeneity of task of interpreting therapeutic claims particularly difficult. In this paper, a simulated randomized clinical trial in patients with coronary artery disease is presented to illustrate the effects that the heterogeneity of patients with angina pectoris may have on the results of clinical experiments. The role of statistics and clinical judgment in solving the problems encountered is explored. The randomized design is



Source http://xkcd.com/1478/

We see this kind of analysis in schools every day: "The test scores are generally down, but our 7th grade girls improved notably in math."

We comb the data looking for patterns that are not really there. As the old adage goes, if you torture the numbers long enough, they'll confess anything.

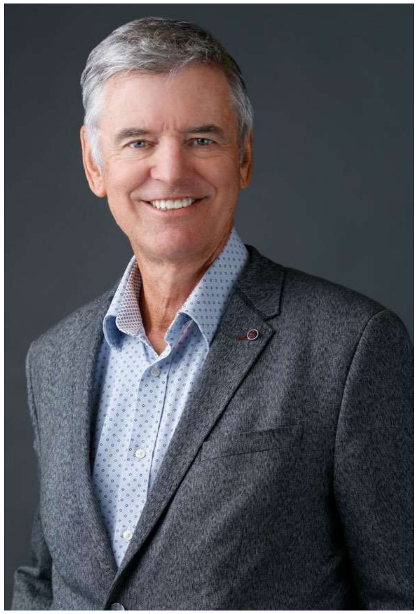
Ethics!

Problem 2. Given the subgroup analytical insight, what do we do next?

Suppose we detect a difference between the girls' and boys' performance, as we did in our example: how will we adjust our actions? Do we divide the class by gender and provide boys with further reading instruction and girls more math support? Perhaps our initial set of instructional activities are gender biased. This may lead to a conversation of whether boys and girls learn differently, but contrary to popular belief, boys and girls are more alike than different when it comes to learning.

In this case, we are lumping students together by gender rather than learning need. There certainly are some boys who need further support in reading, even though they excel at math (e.g. Zach Zelreli | Math: 90, Reading 66). However, there are also boys who do not conform to the aggregate trend (e.g. Marty Morrin | Math: 75, Reading: 100). These data demonstrate the difficulty of leaning on aggregate analyses like subgroup analysis.

The above reveals why John Hattie argues in his work that we should "know the kids and forget the labels." This sentiment is at the heart of learning analytics, and particularly diagnostic analysis. Our default should be to cluster students by learning need not label.



Hattie, J. (2012). *Visible Learning for Teachers: Maximizing Impact on Learning*. Routledge, New York, NY., p. 72

*

Subgroup analysis may make sense if the subgroup is the membership of a special program, like special education or English Language Learners. In these situations, we are really examining how a program is serving the needs of a special population. Students in these programs receive specialized services that should be evaluated for fidelity and efficacy.

However, the statistical limitations of subgroup analysis still apply.

3 Primary Use Cases for Diagnostics in K-12

1. Directing instructional broadcast.
2. Organizing small groups.
3. Understanding the needs of individual learners.

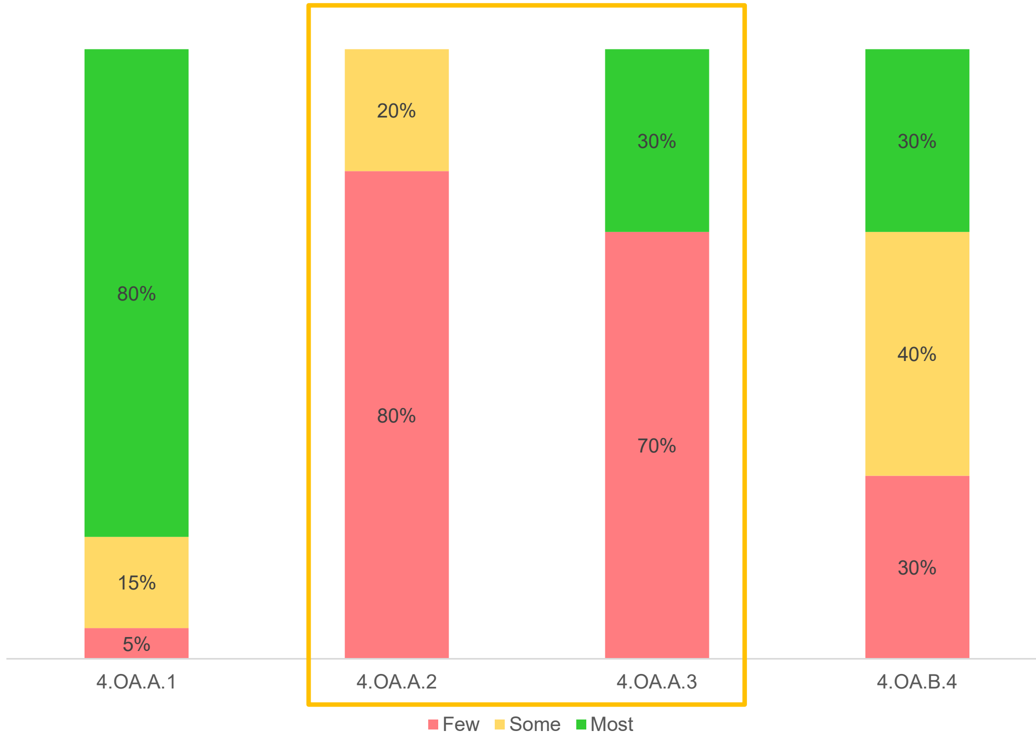
#1: Instructional Broadcast

Whole-group instruction is a regular and necessary component in many classrooms. Instructional broadcast informs the content and rigor that learning. It guides teachers in what to teach the entire class, targeting the instruction at the optimal learning zone for the group.

What do the majority of my learners need next?

#1: Instructional Broadcast

Suppose a teacher gives an assessment covering some of the math standards required for 4th grade students in her state. Each standard had 8-10 questions on the assessment she gave. A key diagnostic question is how she should adjust her whole group instruction now that she has some data on the class's performance. A simple but effective diagnostic technique is to display each standard that was assessed on the assessment and calculate the percentage of students who answered most of the questions correctly, some questions, or only a few (or no) questions correctly. If the assessment has fewer questions per standard, then you may want to simply distinguish between students who answered most of the questions correctly from those who did not.



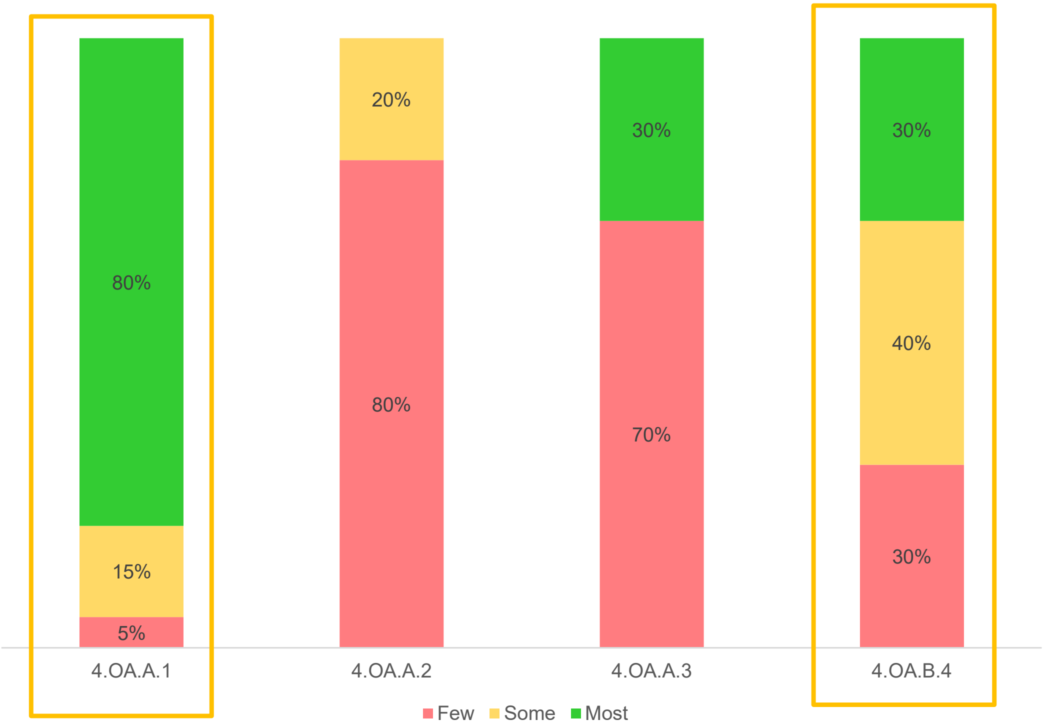
#2: Organizing Small Groups

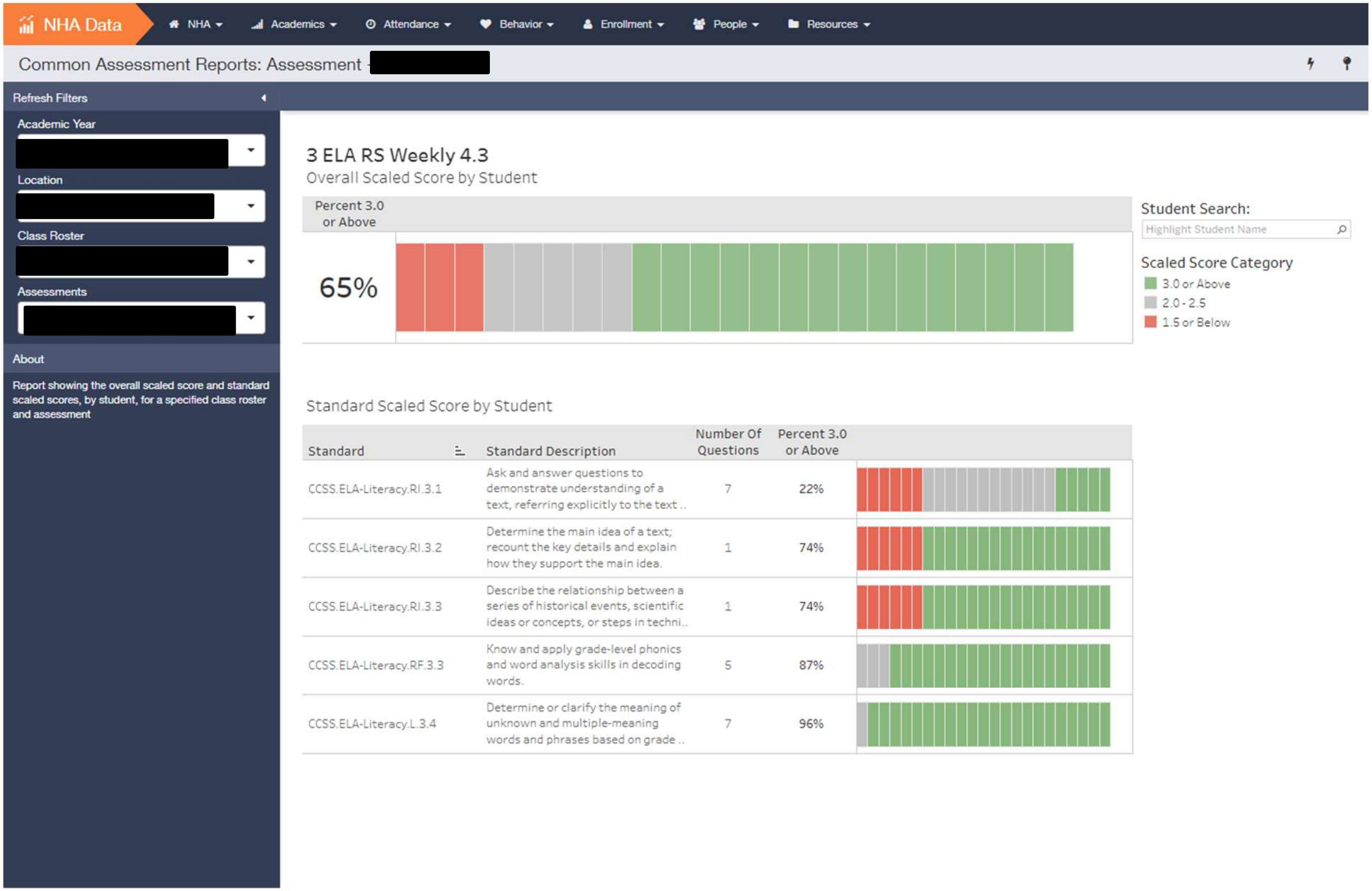
How do my students' academic needs cluster together?

Most classrooms include a diverse group of learners with specialized learning needs. Diagnostic analysis can assist educators in grouping students into similar groups. A key balance in this exercise is balancing group size with group similarity. The smaller the group, the more homogenous their learning needs; however, the smaller the group size, the greater the number of groups and potentially more resources needed to service them.

#2: Organizing Small Groups

4.OA.B.4: This standard is best defined by its variance. A meaningful percentage of students falls into each of the three categories. This standard resists whole group instruction, not because the students know the standard well (70% of the students answered some or few questions correctly), but because the academic diversity in this standard means that whole group instruction will miss most of the class no matter which group it is targeted to. Rather than relying on whole group, the teacher should consider providing a differentiated experience to the three groups of students.




NHA Data
NHA
Academics
Attendance
Behavior
Enrollment
People
Resources
Common Assessment Reports: Assessment [redacted]
Refresh Filters
Academic Year
[redacted]
Location
[redacted]
Class Roster
[redacted]
Assessments
[redacted]
About
Report showing the overall scaled score and standard scaled scores, by student, for a specified class roster and assessment
3 ELA RS Weekly 4.3
Overall Scaled Score by Student
Percent 3.0 or Above
65%
Student Search:
Highlight Student Name
Scaled Score Category
3.0 or Above
2.0 - 2.5
1.5 or Below
Standard Scaled Score by Student
Standard
Standard Description
Number Of Questions
Percent 3.0 or Above
CCSS.ELA-Literacy.RI.3.1
Ask and answer questions to demonstrate understanding of a text, referring explicitly to the text ..
7
22%
CCSS.ELA-Literacy.RI.3.2
Determine the main idea of a text; recount the key details and explain how they support the main idea.
1
74%
CCSS.ELA-Literacy.RI.3.3
Describe the relationship between a series of historical events, scientific ideas or concepts, or steps in techni..
1
74%
CCSS.ELA-Literacy.RF.3.3
Know and apply grade-level phonics and word analysis skills in decoding words.
5
87%
CCSS.ELA-Literacy.L.3.4
Determine or clarify the meaning of unknown and multiple-meaning words and phrases based on grade ..
7
96%

#3: Meeting the needs of individuals

Knowing the needs of individual students can be particularly challenging when teachers serve more than 150 students. Diagnostic data can help foster a familiarity with individual student needs. Where this manifests itself directly is in making programmatic decisions for individuals (intervention, G/T). Individual students often have learning needs that require special program (academic intervention, special education services, gifted/talented). Diagnostic data is useful to inform the nature and intensity of the intervention data.

How do my students' academic needs cluster together?

The link to prescription

Notice at this level, informing instructional broadcast through diagnostic analysis is uninvolved. Our analysis hasn't told the teacher *where* students struggle in standard 4.OA.A.2, only that they did struggle with that standard. To go deeper diagnostically, the teacher would need to examine more evidence in the assessment. If the items are closed-response (multiple choice), item analysis is a helpful next step. Item analysis is not a mathematical or statistical analysis, but rather a content analysis. For multiple choice, test makers can build options to capture common misconceptions. Consider the admittedly low-rigor example to the right.

Find the sum of 12+3.

A. 9
B. 36
C. 42
D. 15
E. None of the above

Obviously, the correct response is (D). However, we consider the other options, we may be able to learn about student misconceptions by analyzing the distractors:

A). Students may have confused subtraction and addition (12 – 3 = 9).
B). Students may have multiplied rather than add (12 x 3 = 36).
C). Students may have confused place value:

$$\begin{array}{r} 12 \\ +\underline{3\ } \\ 42 \end{array}$$

D). Correct Answer: 12 + 3 = 15.
E). Incorrect answer.