ND State Assessment Scores and Learning Recovery

7th Annual School Board Member Bootcamp

NDDPI and NDSBA

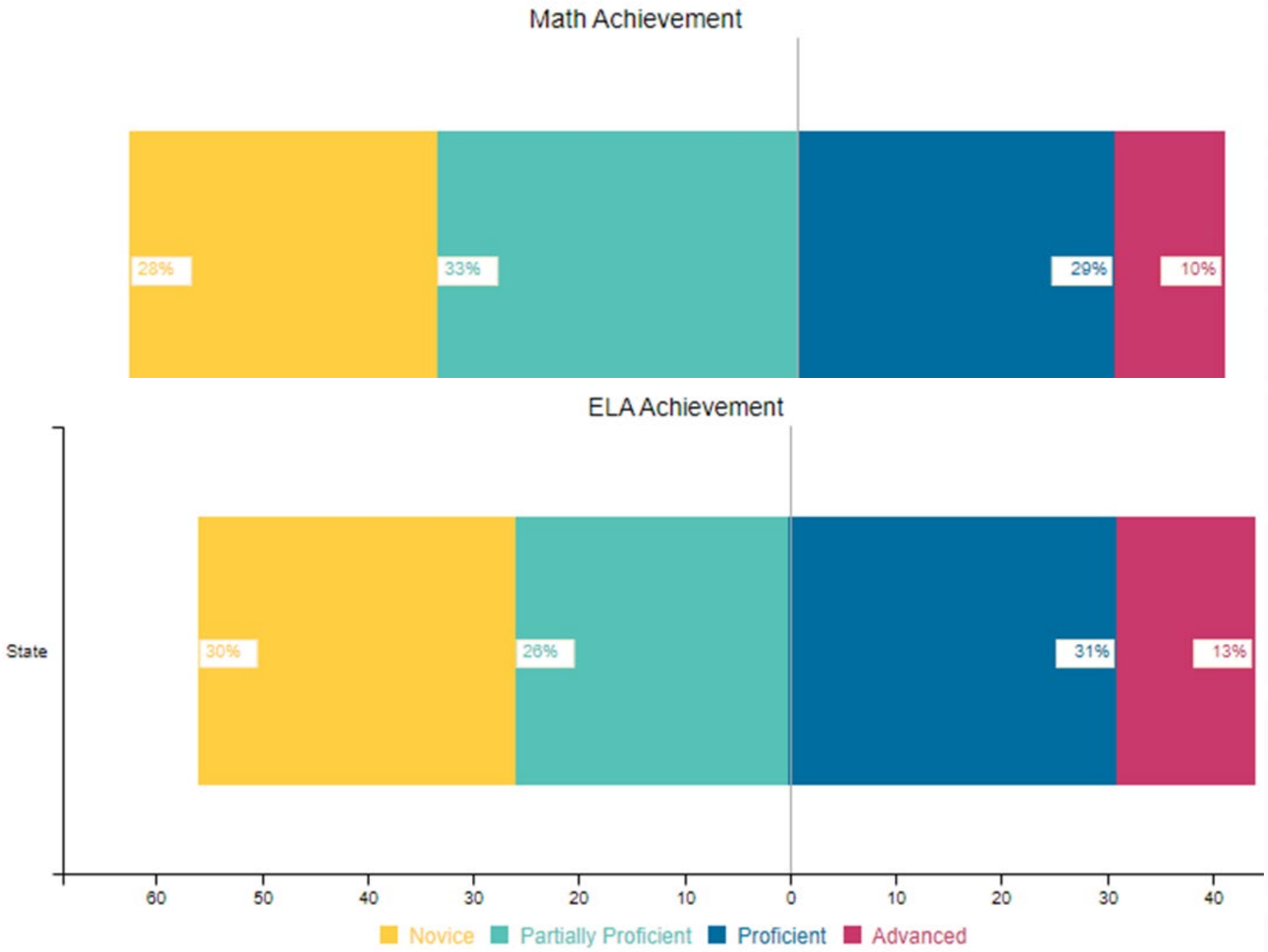
August 1, 2024

Stanley Schauer Jr.

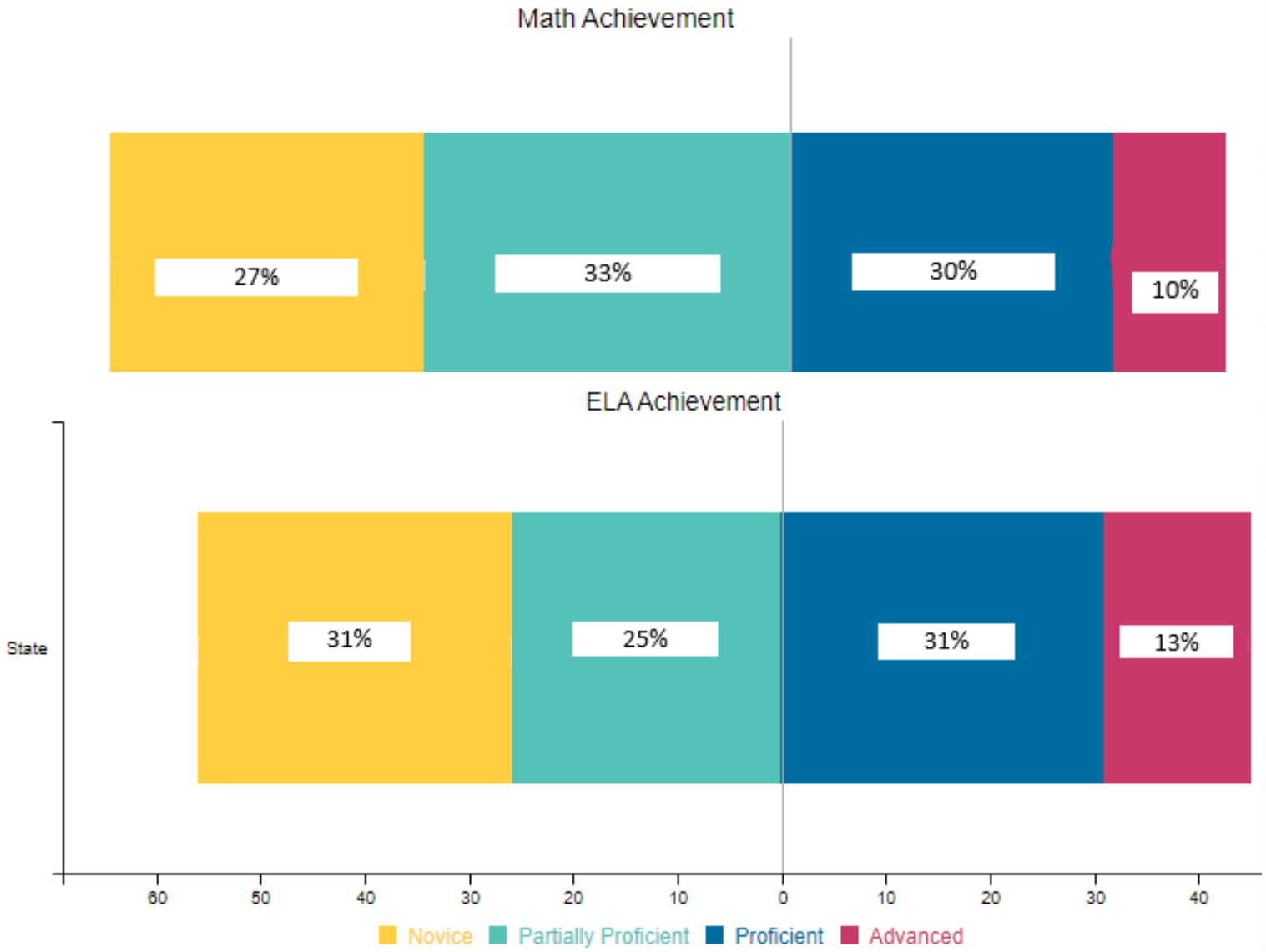
Plan

- 1. Discuss current NDSA scores at state aggregate
- 2. Inform attendees about the data
- 3. Inform attendees about the data sources
- 4. Discuss learning recovery since the pandemic
- 5. Brief - future of assessment in ND

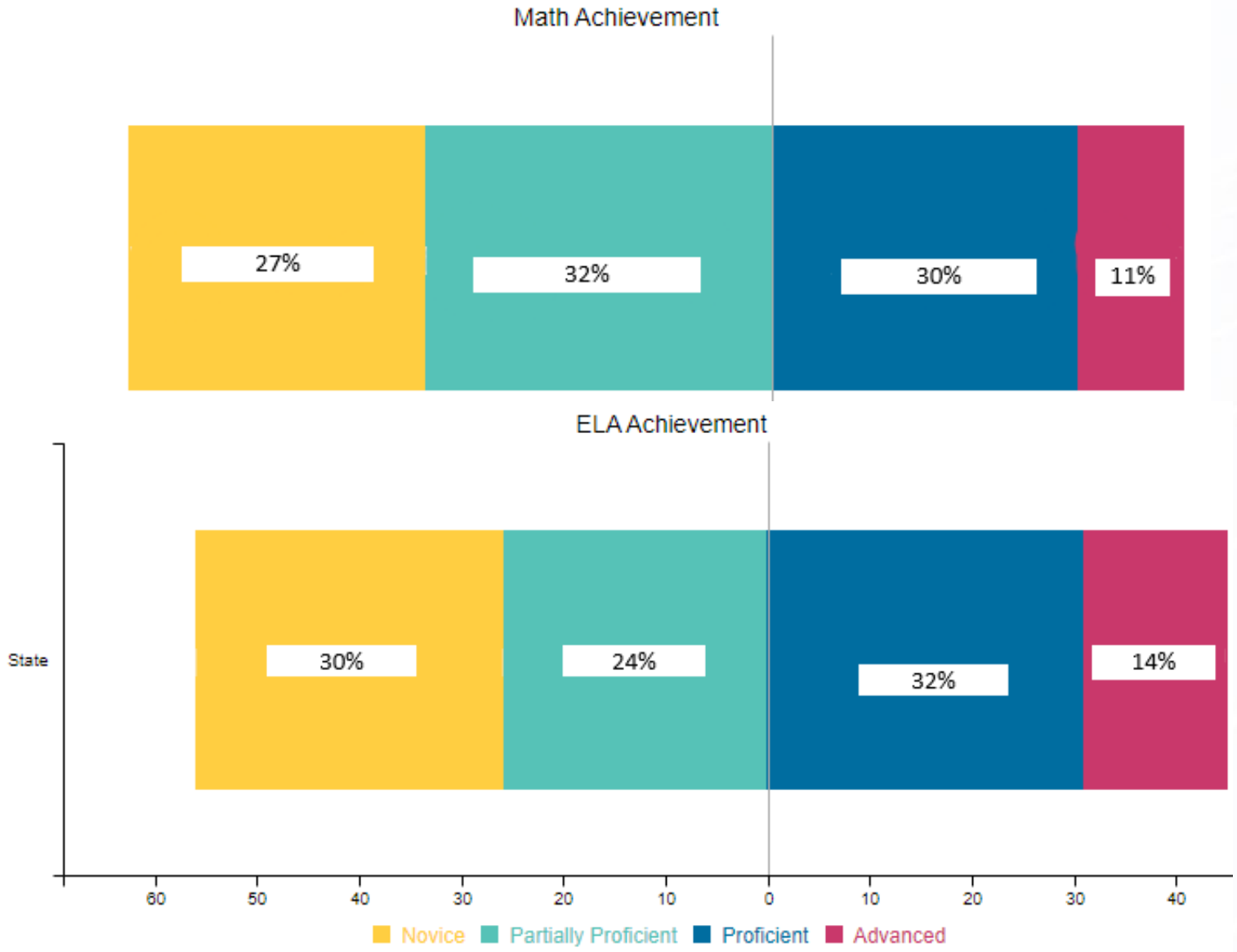
22-23 ND Insights Statewide NDSA Performance (accountability rules applied)



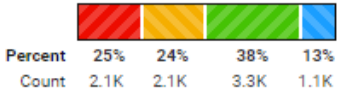
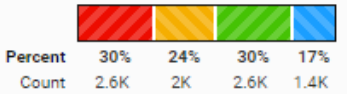
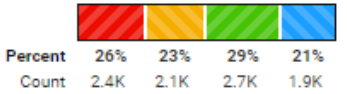
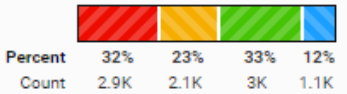
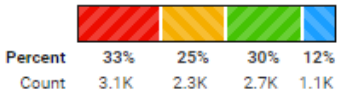
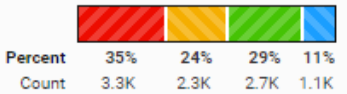
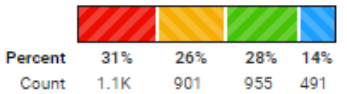
22-23 TIDE-Centralized Reporting System (CRS)*



23-24 TIDE-Centralized Reporting System (CRS)*



ELA Centralized Reporting System (CRS)

Assessment Name	Test Reason	Student Count	Average Score	Performance Distribution
Grade 8 ELA	Spring 2024	8622	648 ± 1	 Percent 25% 24% 38% 13% Count 2.1K 2.1K 3.3K 1.1K
Grade 7 ELA	Spring 2024	8613	634 ± 1	 Percent 30% 24% 30% 17% Count 2.6K 2K 2.6K 1.4K
Grade 6 ELA	Spring 2024	9107	635	 Percent 26% 23% 29% 21% Count 2.4K 2.1K 2.7K 1.9K
Grade 5 ELA	Spring 2024	9053	613	 Percent 32% 23% 33% 12% Count 2.9K 2.1K 3K 1.1K
Grade 4 ELA	Spring 2024	9182	590	 Percent 33% 25% 30% 12% Count 3.1K 2.3K 2.7K 1.1K
Grade 3 ELA	Spring 2024	9223	574	 Percent 35% 24% 29% 11% Count 3.3K 2.3K 2.7K 1.1K
Grade 10 ELA	Spring 2024	3423	652 ± 1	 Percent 31% 26% 28% 14% Count 1.1K 901 955 491

Math Centralized Reporting System (CRS)

Assessment Name	Test Reason	Student Count	Average Score	Performance Distribution
Grade 8 Math	Spring 2024	8706	556 ± 1	Percent 25% 37% 29% 8% Count 2.2K 3.3K 2.5K 734
Grade 7 Math	Spring 2024	8673	529 ± 1	Percent 30% 32% 28% 11% Count 2.6K 2.8K 2.4K 944
Grade 6 Math	Spring 2024	9168	499 ± 1	Percent 25% 34% 30% 11% Count 2.3K 3.1K 2.7K 1K
Grade 5 Math	Spring 2024	9100	477	Percent 22% 32% 32% 14% Count 2K 3K 2.9K 1.3K
Grade 4 Math	Spring 2024	9257	452	Percent 31% 30% 30% 10% Count 2.9K 2.8K 2.7K 896
Grade 3 Math	Spring 2024	9306	426	Percent 27% 23% 39% 11% Count 2.5K 2.1K 3.6K 1K
Grade 10 Math	Spring 2024	3462	600 ± 1	Percent 42% 32% 17% 9% Count 1.4K 1.1K 580 323

Historical NDSA performance levels

	Novice	Partially Proficient	Proficient	Advanced		P+A	Notes
Math							
18-19	24%	32%	34%	11%		45%	Pre-pandemic
19-20	NA	NA	NA	NA		NA	No statewide assessment - pandemic
20-21	28%	34%	29%	9%		38%	
21-22	28%	33%	29%	10%		39%	
22-23	28%	33%	29%	10%		39%	
23-24*	27%	32%	30%	11%		41%	*CRS raw data, no accountaiblity filters like other data
ELA							
18-19	26%	26%	33%	14%		47%	Pre-pandemic
19-20	NA	NA	NA	NA		NA	No statewide assessment - pandemic
20-21	31%	26%	30%	12%		42%	
21-22	30%	26%	31%	14%		45%	
22-23	30%	26%	31%	13%		44%	
23-24*	30%	24%	32%	14%		46%	*CRS raw data, no accountaiblity filters like other data

Understanding NDSA Data

- NDSA or North Dakota State Assessment is a broad-scope, computer adaptive, criterion-referenced (test scores generate a statement about the behavior that can be expected of a person with that score) test that measures student performance against the ND education content standards.
- Standards are used to build Test Blueprints. Test blueprints drive which items will be included in an item pool. Achievement Level Descriptors (ALDs) are part of the statement/expectation(grade, subject, and standard level) mentioned above. ALDs include the four Performance Levels.
- Each student falls into one of the four performance levels (novice, partially proficient, proficient, and advanced) by receiving a single scale score (per subject). The meaning behind these levels is created throughout the test construction process, which heavily involves educator input.

Understanding NDSA Data

- **The length and rigor of the NDSA is determined by the creation process (example – more standards equal a longer assessment). ALDs set the tone for what a proficient or advanced student, at the standard/subject/grade level, can do. Lastly, the standard setting process that occurs after first administration really sets the watermark for years to come.**
- **It is vital to understand that this test is valid/reliable – but no assessment or test should be used in isolation to determine a students/schools/districts performance. Always use as much evidence as possible!**
- **Test is originally designed to allow ranking and comparisons at the district and school levels. As one drops aggregated levels, the standard error of measurement increases. So, the measurement error is highest at the individual student level.**
- **Lastly – coverage or grain size mattes when talking about assessment data of any kind. It is possible that a single test item is addressing an entire content standard. That's considered large scale or large grain size.**
    - Example – with one 100 3rd graders answering 1 or 2 questions related to a standard you can get solid information on the understanding of that standard. What you do not get is detailed information about a single students understanding or weakness with a standard.

RL.3.1 Ask and answer questions to demonstrate understanding of a text, referring explicitly to the text (textual evidence) as the basis for the answers.

English Language Arts Grade 3

ALD	Standard	Novice	Partially Proficient	Proficient	Advanced
Policy		The Level 1 student is below proficient in applying the English language arts/literacy knowledge/skills as specified in the North Dakota English Language Arts and Literacy (ELA) Content Standards. The student generally performs significantly below the standard for the grade level, is likely able to partially access grade-level content and engages with higher-order thinking skills with extensive support.	The Level 2 student is approaching proficient in applying the English language arts/literacy knowledge/skills as specified in the North Dakota English Language Arts and Literacy (ELA) Content Standards. The student generally performs slightly below the standard for the grade level, is likely able to access grade-level content and engages in higher-order thinking skills with some independence and support.	The Level 3 student is proficient in applying the English language arts/literacy knowledge/skills as specified in the North Dakota English Language Arts and Literacy (ELA) Content Standards. The student generally performs at the standard for the grade level, is able to access grade-level content and engages in higher-order thinking skills with some independence and minimal support.	The Level 4 student is highly proficient in applying the English language arts/literacy knowledge/skills as specified in the North Dakota English Language Arts and Literacy (ELA) Content Standards. The student generally performs significantly above the standard for the grade level, is able to access above grade-level content and engages in higher-order thinking skills independently.
		For a range of grade-appropriate texts, the Level 1 student	For a range of grade-appropriate texts, the Level 2 student	For a range of grade-appropriate texts, the Level 3 student	For a range of grade-appropriate texts, the Level 4 student
Reading: Literature/Fiction					
Range	RL.3.1	asks and answers explicit questions to demonstrate understanding of a text.	asks and answers explicit questions to demonstrate understanding text, referring to the text as the basis for answers.	asks and answers questions to demonstrate understanding of a text, referring explicitly to textual evidence as the basis for answers.	asks and answers complex questions to demonstrate understanding of a text, referring explicitly to the textual evidence as the basis for answers.

English Language Arts NDSA Cut Scores

Achievement Level	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 10
Advanced	621	639	661	671	680	702	713
Proficient	585	600	622	638	641	650	667
Partially Proficient	560	572	595	610	611	616	627

North Dakota Grades 3-5 ELA Computer Adaptive Blueprint

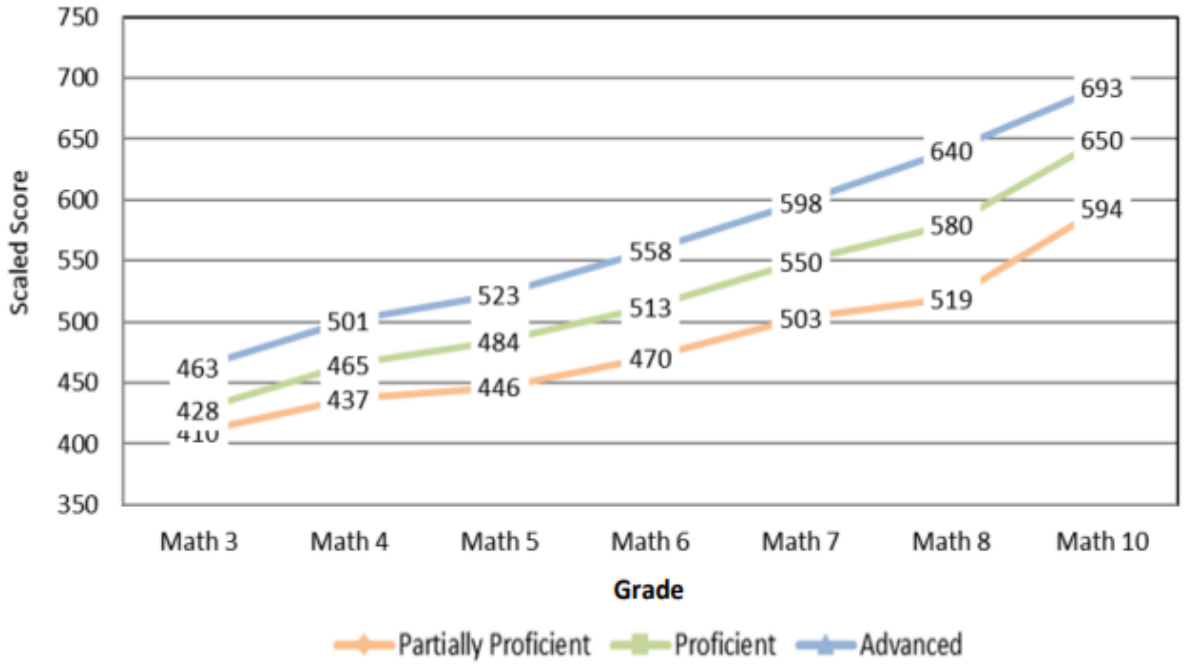
Reporting Category	Standards Assessed	DOK	Approximate Portion of Test 38 – 42 items
Reading Literature Text 15 – 17 items		**1,2,3**	**39-40%**
Key Ideas and Details	RL.1 RL.2 RL.3	1,2,3	
Craft and Structure	RL. 4 RL. 5 RL.6	2,3	
Integration of Knowledge and Ideas	RL. 7 RL.9	2,3	
Listening	SL.2 SL.3	1,2	

Key Ideas and Details

3.RL.1		3.RL.2		3.RL.3	
Proficient?	Weak or Strong?	Proficient?	Weak or Strong?	Proficient?	Weak or Strong?
✖	+	✖	+	✖	+
✖	+	✖	+	✖	+
⊖	=	⊖	=	✓	=
✖	=	✖	+	✖	+
✓	=	✓	+	⊖	=

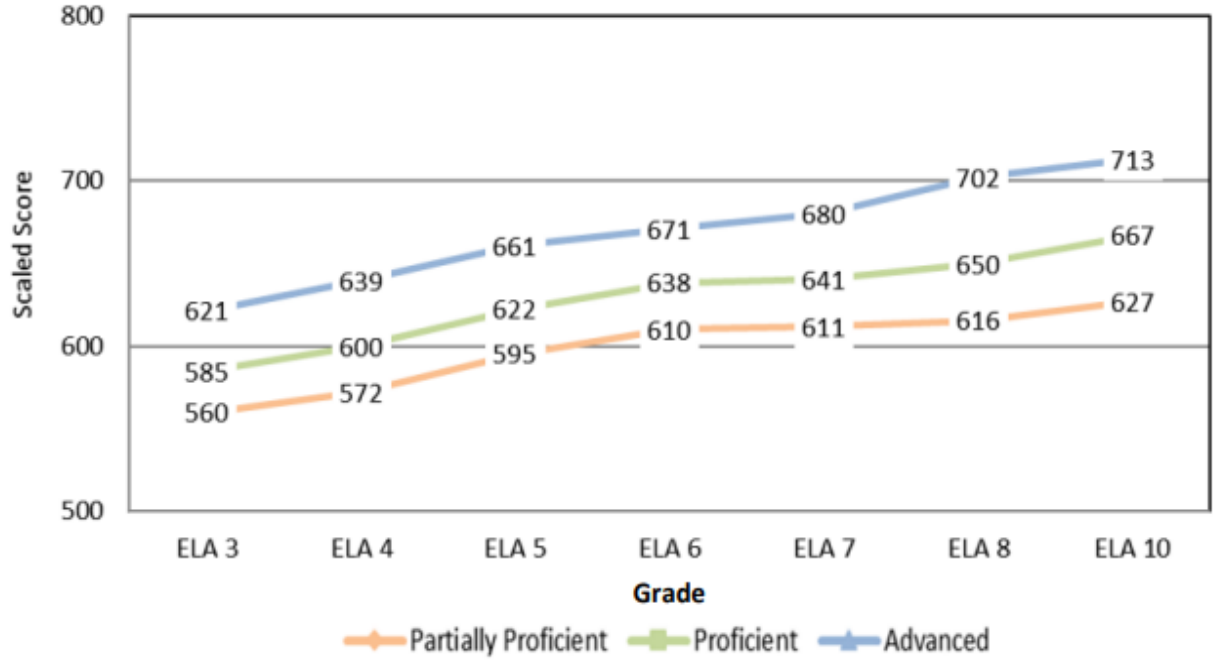
Mathematics
NDSA Cut Scores

Achievement Level	Grade						
	3	4	5	6	7	8	10
Advanced	463	501	523	558	598	640	693
Proficient	428	465	484	513	550	580	650
Partially Proficient	410	437	446	470	503	519	594



English Language Arts
NDSA Cut Scores

Achievement Level	Grade						
	3	4	5	6	7	8	10
Advanced	621	639	661	671	680	702	713
Proficient	585	600	622	638	641	650	667
Partially Proficient	560	572	595	610	611	616	627



Average testing time

Math	mean	Reading	mean	Writing	mean
G3	0:57	G3	1:06	G3	0:58
G4	0:59	G4	0:57	G4	0:59
G5	1:04	G5	1:03	G5	1:00
G6	1:07	G6	0:59	G6	0:54
G7	0:56	G7	0:53	G7	0:47
G8	0:51	G8	0:47	G8	0:45
G10	0:52	G10	0:43	G10	0:42

NDSA Data Sources

- **There are three main data locations that contain NDSA data:**
  - 1. ND Insights
    - This is a public website and data that relates to performance on the NDSA can be found (for state, district, school) under Academic Progress and then either Student Achievement or Student Growth. Contains no individual student data and suppression rules are applied. Accountability rules applied.
  - 2. STARS Reporting and Analytics
    - This is not a public website and is typically only accessed by educators. This is essentially the data that can be found on ND Insights, but one can access individual student data and see the data behind the scenes of ND Insights. No suppression rules and accountability rules are applied or not applied, can filter data sets.
  - 3. Centralized Reporting System (CRS)
    - This is not a public website and is typically only accessed by educators. This is the direct reporting system for the NDSA and is updated in real-time when students finish testing. Slides 6 and 7 are from the CRS. This is more “raw” data, but still formatted in a way that can be used. Contains individual student data, no suppression rules, no accountability rules.

Learning Recovery

- Measuring learning recovery is very complex and should involve more than just a statewide summative like NDSA. Below are three examples of how this can be done:
- 1. A simplified look at a cohort:

		Grade 3	Grade 8	
		18-19	23-24	Difference
Math	N	28%	25%	-3%
	PP	23%	37%	14%
	P	40%	29%	-11%
	A	9%	9%	0%
		18-19	23-24	Difference
ELA	N	27%	25%	-2%
	PP	26%	24%	-2%
	P	35%	38%	3%
	A	12%	13%	1%

- **2. Highest aggregate comparison of performance levels:**

	Novice	Partially P	Proficient	Advanced		P+A	Notes
Math							
18-19	24%	32%	34%	11%		45%	Pre-pandemic
23-24*	27%	32%	30%	11%		41%	*CRS raw data, no accountaiblity filters like other data
ELA							
18-19	26%	26%	33%	14%		47%	Pre-pandemic
23-24*	30%	24%	32%	14%		46%	*CRS raw data, no accountaiblity filters like other data

- **3. SAS-Education Visualization and Analytics Solution (EVASS) study of learning disruption and recovery.**
  - This is far more "scientific" than the previous two slides. We partnered with SAS for two different data runs (20-21 showed disruption and 21-22 showed recovery).
  - The report uses effect sizes. In very simple terms, at the student level, the data uses previous test results and builds a projection. Then, it measures actual performance against that projection and measures the difference, or the effect. Positive is good and negative is bad.
  - Key observations from the 20-21 data run:
    - Overall: "…there is evidence of a decline in learning, but the observed impacts are not especially large relative to the variation observed in pre-pandemic results."
    - ELA: "The negative effect sizes provide evidence of learning decline, but the magnitude of this impact states and national studies." tended to be less than what has been reported in research from other
    - Math: "In grades 5–8 and 10 NDSA Math, there was an observable amount of learning decline represented by effect sizes of -0.18 to -0.10. The magnitude of this impact was less than what has been reported in research in other states and national studies."

- What we concluded from the analysis of our own matched key findings:
  - We saw a larger decline in Math than in ELA and our drop in students who were proficient and advanced were evident, but nothing like we were hearing in other states.
- Key observations from the 21-22 data run:
  - "The recovery effect considers whether the 2021-22 results indicate a flat, positive, or negative recovery. Statistically significant positive results were observed for NDSA ELA in grades 4, 6, and 7, demonstrating recovery toward the pre-pandemic trend. Statistically significant negative results were observed for NDSA ELA in grade 8 and NDSA Math in grades 7 and 8. This suggests that students taking these assessments are tending to show slower recovery."
  - What we concluded from the analysis also matched these findings:
    - Decline leveling off and started to see hold steady/recovery in certain grade levels.

Future of Assessment in ND

- **Summative assessment: ND A+ Summative**
  - NDSA and ACT will be replaced by ND A+ summative. Same grade levels (no 11th grade w/ ACT), same use and purpose, same reporting measures. New math and ELA standards, new vendor, new platform.
  - Spring 25 will be the first administration
- **Interim assessment: ND A+ interim**
  - State provided is ND A+ interims.
  - 24-25 is second year
- **Formative assessment: ND A+ classroom formative - Navvy**
  - New in 24-25 – the start of SY

- **The above three make up the new ND A+ system.**

Thank you!

seschauer@nd.gov for follow up questions or conversations