ELLs and Their Performance on State-Mandated Assessments

Presented By:

Christina Kitson

Doctoral Candidate at the University of Kansas

Instructor at Kansas State University

Growing Population

- There is a general consensus in the research that this is the fastest growing group of students in the K-12 system in the United States
  - Wolf et al., 2008a; Tsang, Katz, & Stack, 2008; Abedi, 2004; Kim & Herman, 2009

- In Kansas- ELLs made up between 7 – 14 % of student population in the 2009-2010 school year.

- Kansas experienced a five percent population growth in ELLs in public schools from the school year 2000-2001 to 2009-2010.

Equal Opportunity for Education

- Education Acts to ensure Equal Opportunity:
  - Equal Educational Opportunities Act (1974) (PL 93-380)
  - No Child Left Behind Act (2001) (PL 107-110)
    - Title I – All students have to perform in terms of accountability and adequate yearly progress
    - Title III- All ELLs have to show they are making progress in acquiring English Language proficiency
- Influential court case-
  - Lau v. Nichols (1974) (414 U.S. 563)
  - **Courts found that ELLs should be treated equally with other students and given the same opportunities to learn, including additional English instruction to facilitate English language acquisition**

Current Research

- Often groups all ELLs into a single sub-group when looking at their content performance

- Abedi, & Gandara, 2006; Pitoniak, Young, Martiniello, King, Buteux, & Ginsburgh, 2009; Young, Holtzman, & Steinberg, 2011

Requirements of Language Assessment

- All language assessments have to have:
  - Reading, Writing, Listening, and Speaking sections
  - Assess the ‘academic language’ of the student
  - Align with the states language proficiency and state content standards.

- Scores are used for:
  - What services to offer ELLs
  - Report for school and district accountability

Validity of Content Assessments

- Wolf, Herman, and Dietel (2010) say, validity is, “the degree to which an assessment system produces accurate information about ELL students’ performance and provides a sound basis for policy decision-making”.
- If language proficiency influences students’ scores, then the validity of the content assessments is in question for use with ELLs
- Quoting from The National Research Council directly Abedi and Gandara say, “if a student is not proficient in the language of the test, her performance is likely to be affected by construct-irrelevant variance.”

More on Validity

- “Performance on these tests may reflect the English language abilities of ELL students rather than their knowledge of the content material the tests are designed to measure (e.g., mathematics skills, scientific knowledge, etc.)” (Abedi, J., Bailey, A., Butler, F., Castellon-Wellington, M., Leon, S., & Mirocha, J., 2005)

- According to Wolf, Herman, and Dietel (2010) even math tests are in a way English language tests for the ELL students, “the language demands of any test may get in the way of ELL students showing what they know and inappropriately constrain their performance”

Assumptions of Study

- Year to year students will perform similarly on their assessments
  - Both their content and proficiency assessments which allows for generalizations

- That a single state may have implications for other states
  - The study is looking only at one year in Kansas
  - Each state has unique tests (or is part of a consortium)
  - Each state has unique interpretations of test scores
    - Unique standards, cut scores, and proficiency levels

Required Tests

- During the first year
  - All students identified as ELLs take proficiency assessment (within 30 days)
  - All students take the Math content assessment (3rd grade and up)
  - Have the option to take the Reading 3rd grade and up as well
- After the first year all grade required testing and the proficiency assessment
  - English Proficiency Assessment
  - Math from 3rd grade to 11th
  - Reading from 3rd grade to 11th
  - Science in grades 4, 7, with two opportunities in 9-11th

Data used in current study

- KELPA score (proficiency group)
  - Fluent
  - Advanced
  - Intermediate
  - Beginning
- Content test scores (Math, Reading, Science)
  - Meeting (grouped Exemplary, Exceeds Standard, & Meets Standard)
  - Not Meeting Standards (grouped Approaches Standard & Academic Warning)
- Demographic Variables
  - Gender; Eligibility for National School Lunch Program; Exceptionality Code; First Language; and Number of Years in the US

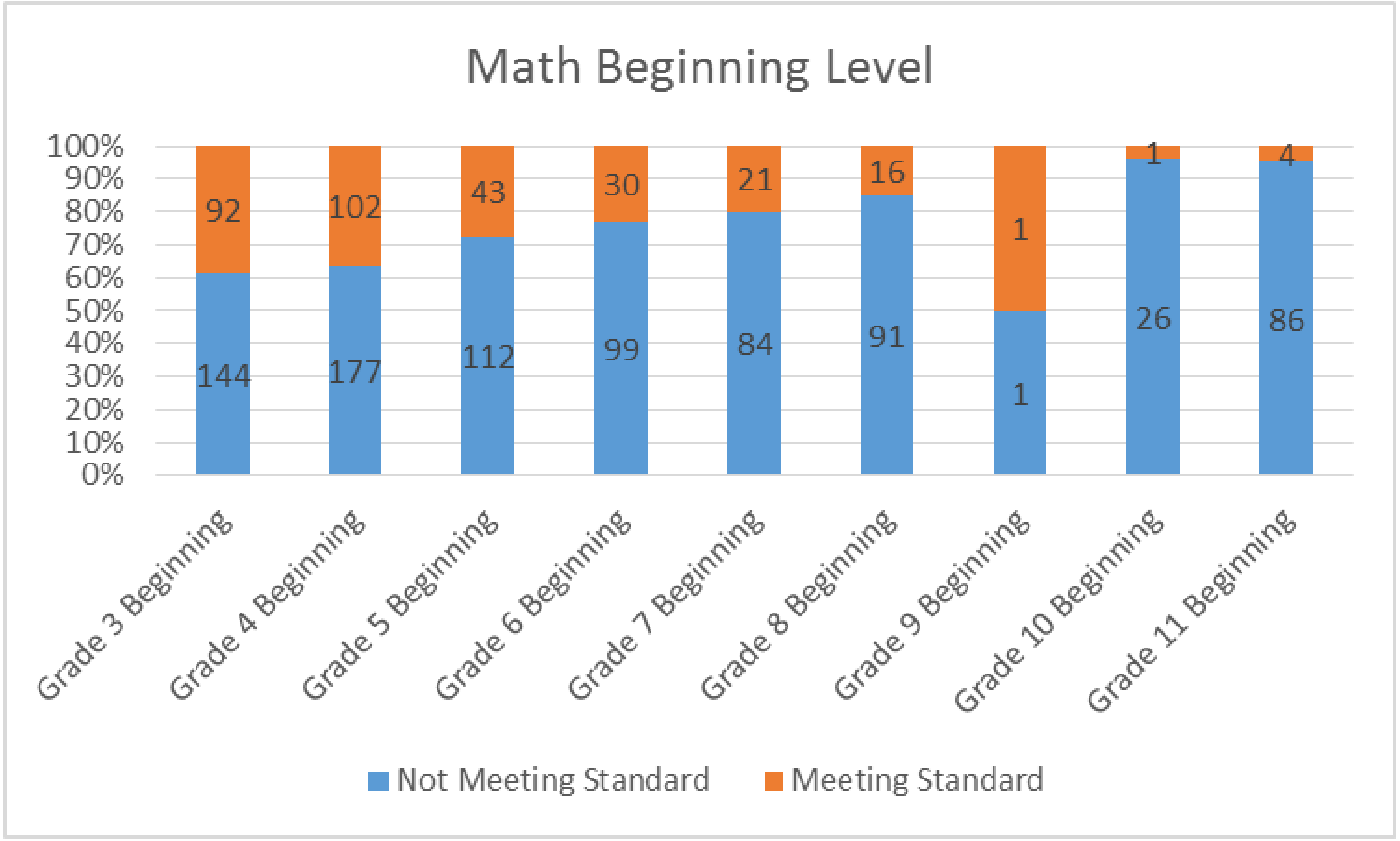
RQ#1

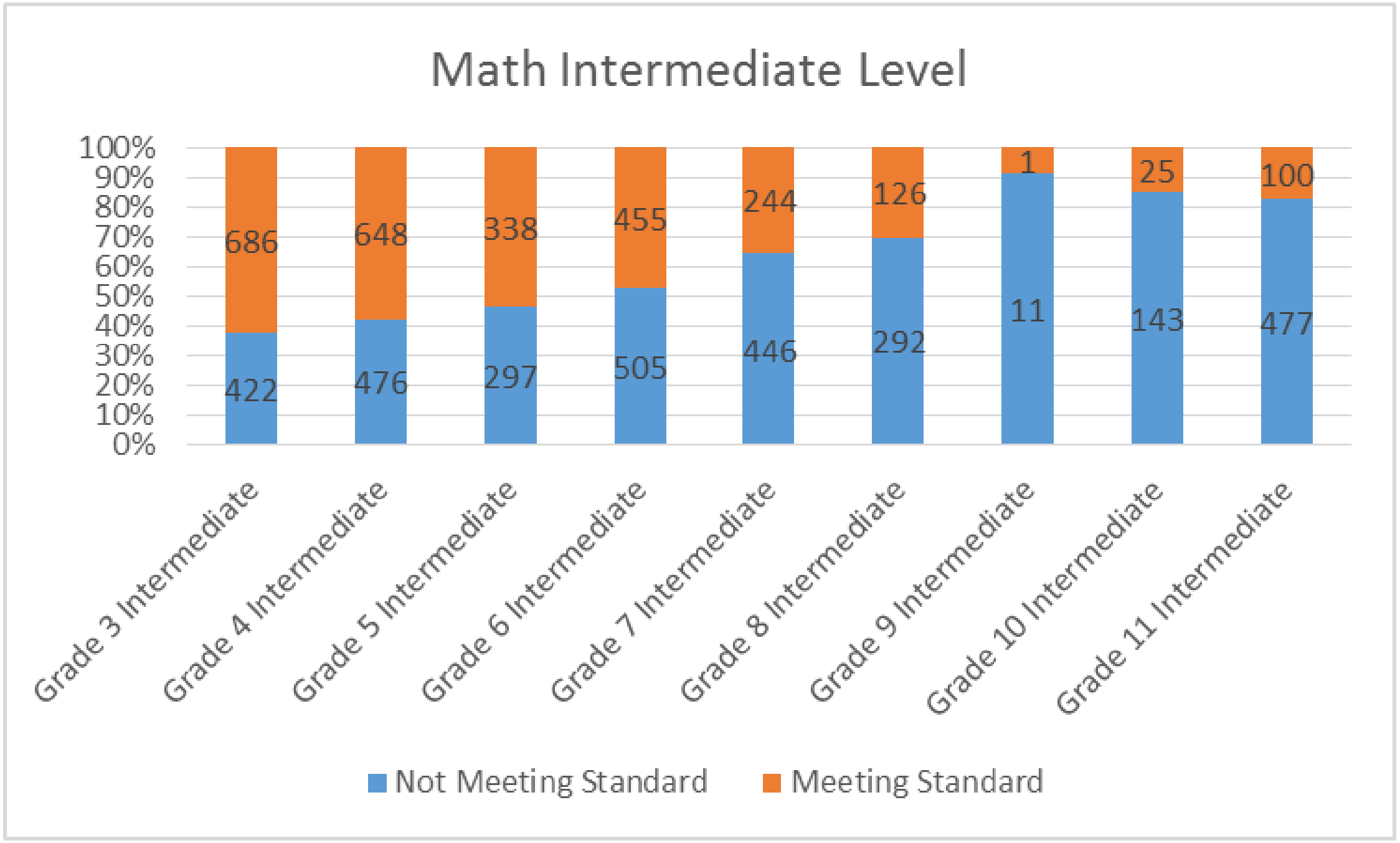
- What are the relative effects of proficiency level on assessment scores across grade levels?

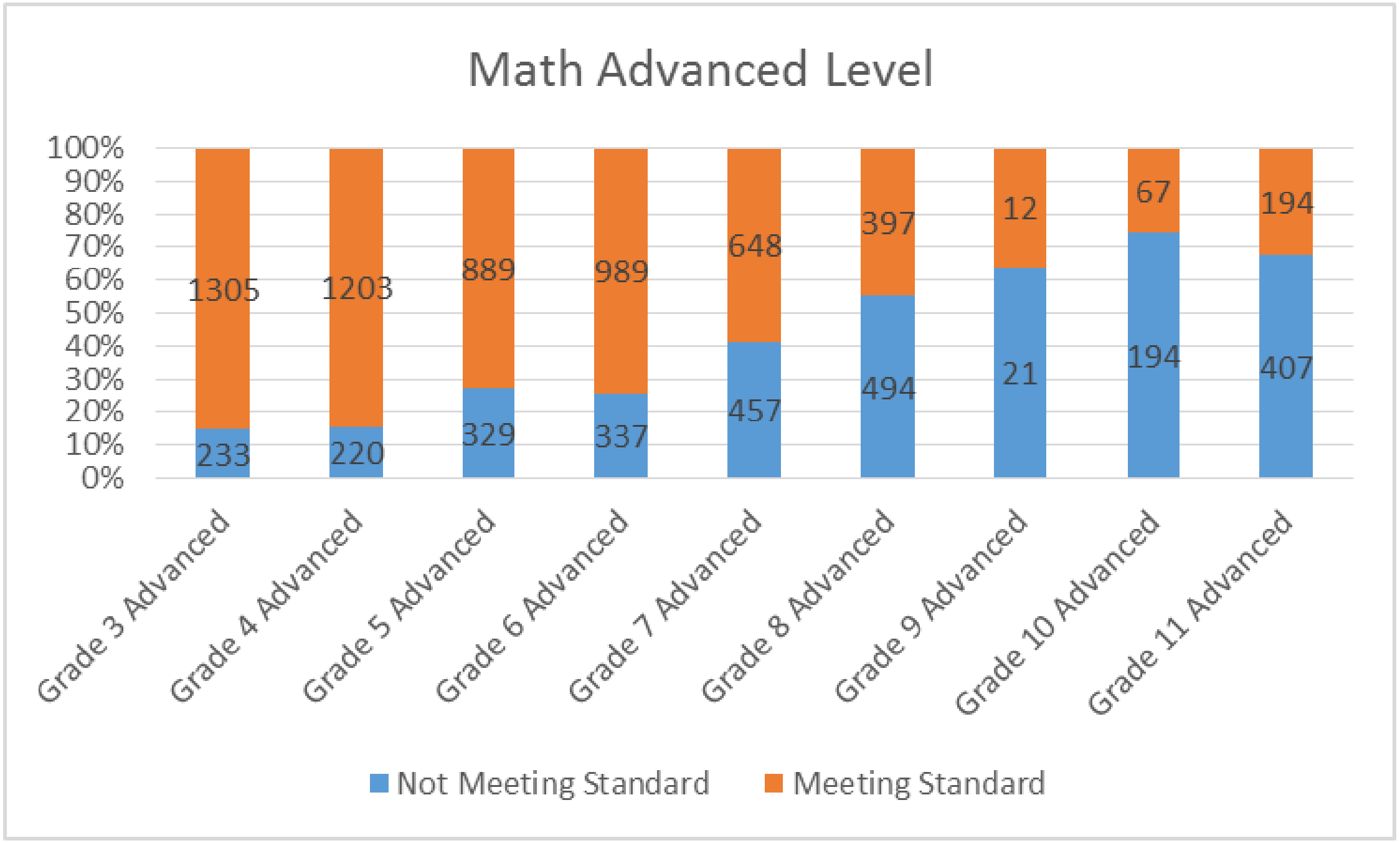
- Methods-
- Proportions will be displayed showing how students performed on each content assessment based on grade level.

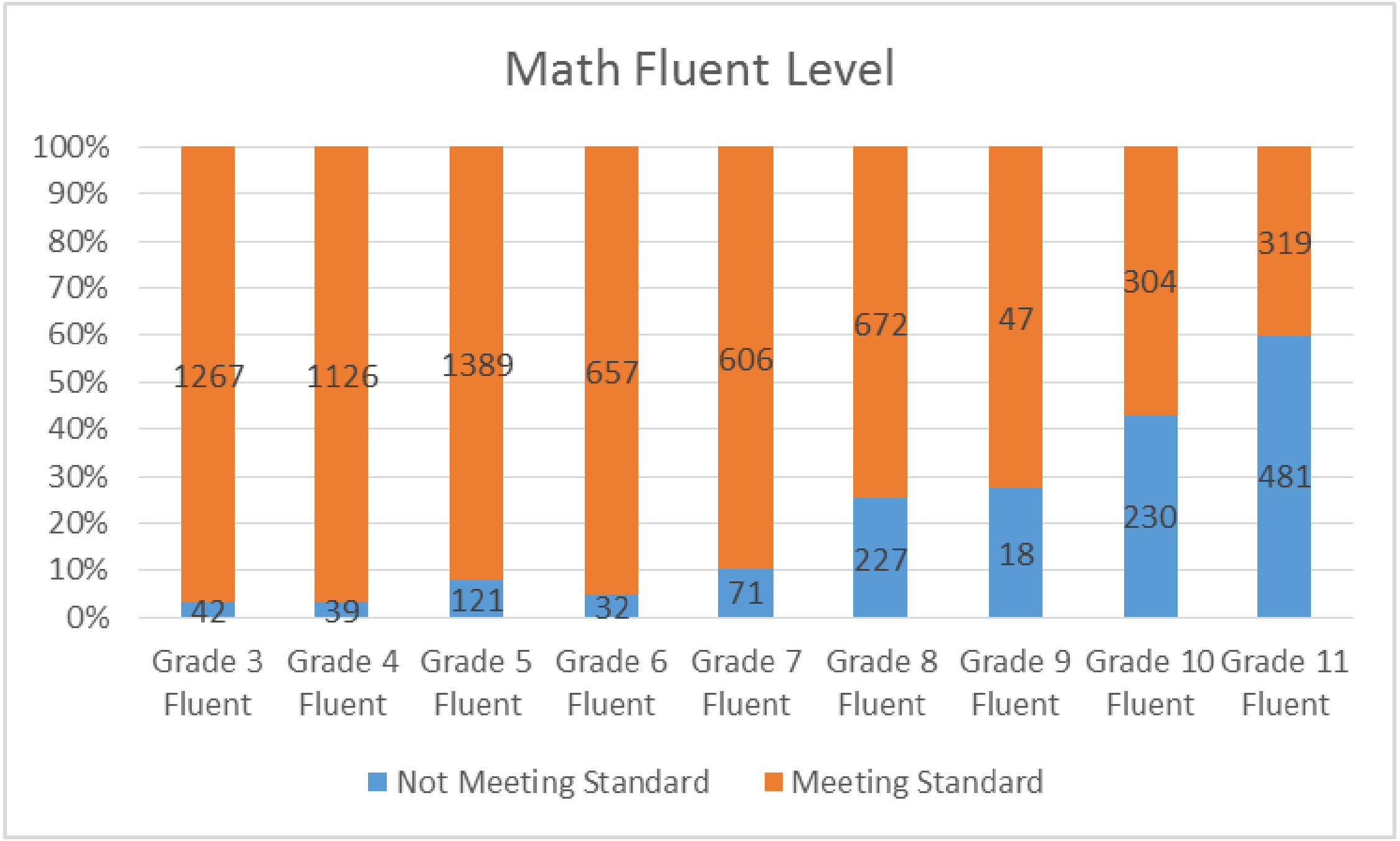
Math

- Look for patterns in the next few charts…


Math Beginning Level
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
92
102
43
30
21
16
1
1
144
177
112
99
84
91
1
26
86
Grade 3 Beginning
Grade 4 Beginning
Grade 5 Beginning
Grade 6 Beginning
Grade 7 Beginning
Grade 8 Beginning
Grade 9 Beginning
Grade 10 Beginning
Grade 11 Beginning
Not Meeting Standard
Meeting Standard


Math Intermediate Level
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
686
648
338
455
244
126
1
25
100
422
476
297
505
446
292
11
143
477
Grade 3 Intermediate
Grade 4 Intermediate
Grade 5 Intermediate
Grade 6 Intermediate
Grade 7 Intermediate
Grade 8 Intermediate
Grade 9 Intermediate
Grade 10 Intermediate
Grade 11 Intermediate
Not Meeting Standard
Meeting Standard


Math Advanced Level
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
1305
1203
889
989
648
397
12
67
194
233
220
329
337
457
494
21
194
407
Grade 3 Advanced
Grade 4 Advanced
Grade 5 Advanced
Grade 6 Advanced
Grade 7 Advanced
Grade 8 Advanced
Grade 9 Advanced
Grade 10 Advanced
Grade 11 Advanced
Not Meeting Standard
Meeting Standard


Math Fluent Level
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
1267
42
1126
39
1389
121
657
32
606
71
672
227
47
18
304
230
319
481
Grade 3 Fluent
Grade 4 Fluent
Grade 5 Fluent
Grade 6 Fluent
Grade 7 Fluent
Grade 8 Fluent
Grade 9 Fluent
Grade 10 Fluent
Grade 11 Fluent
Not Meeting Standard
Meeting Standard

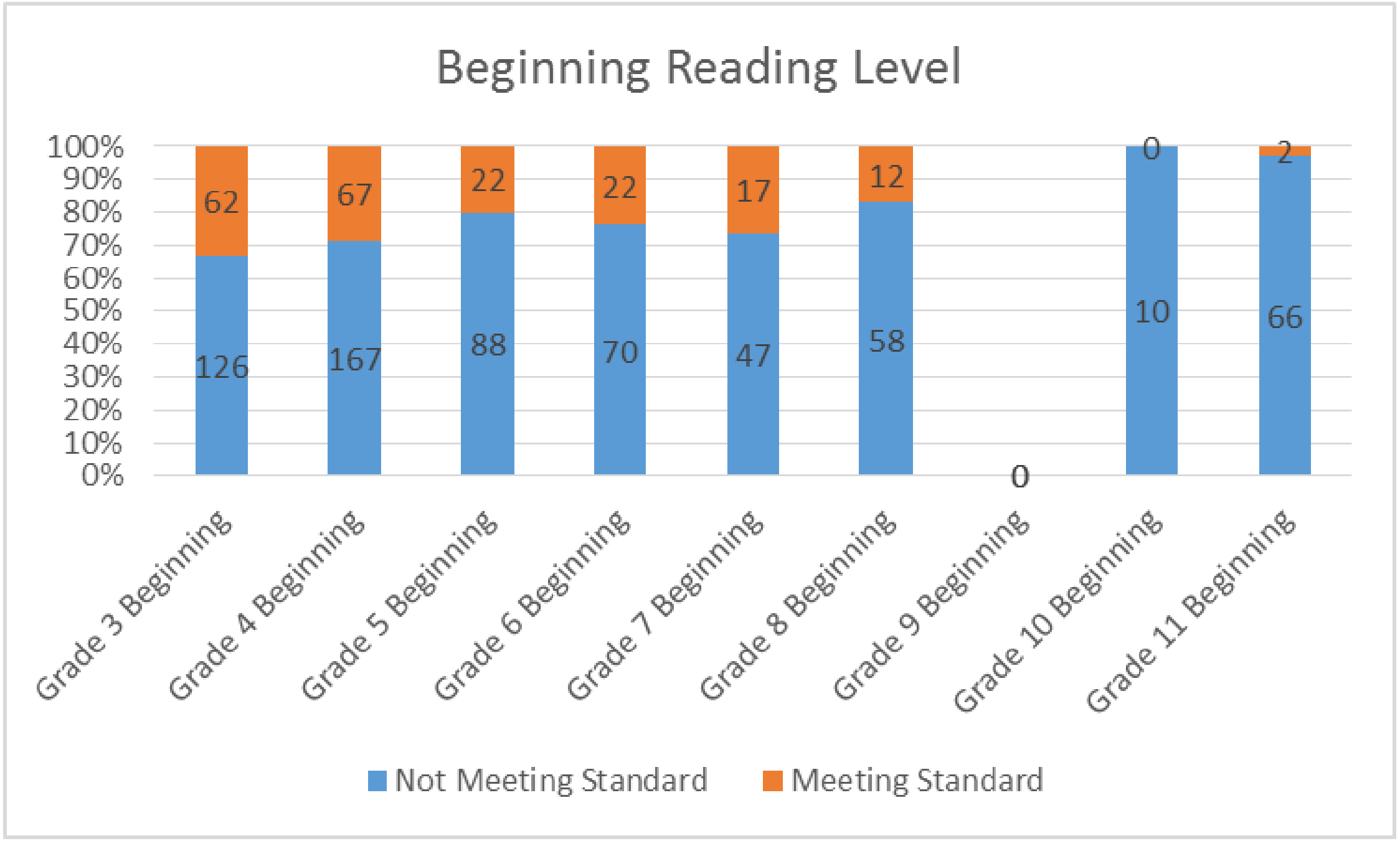
Patterns?

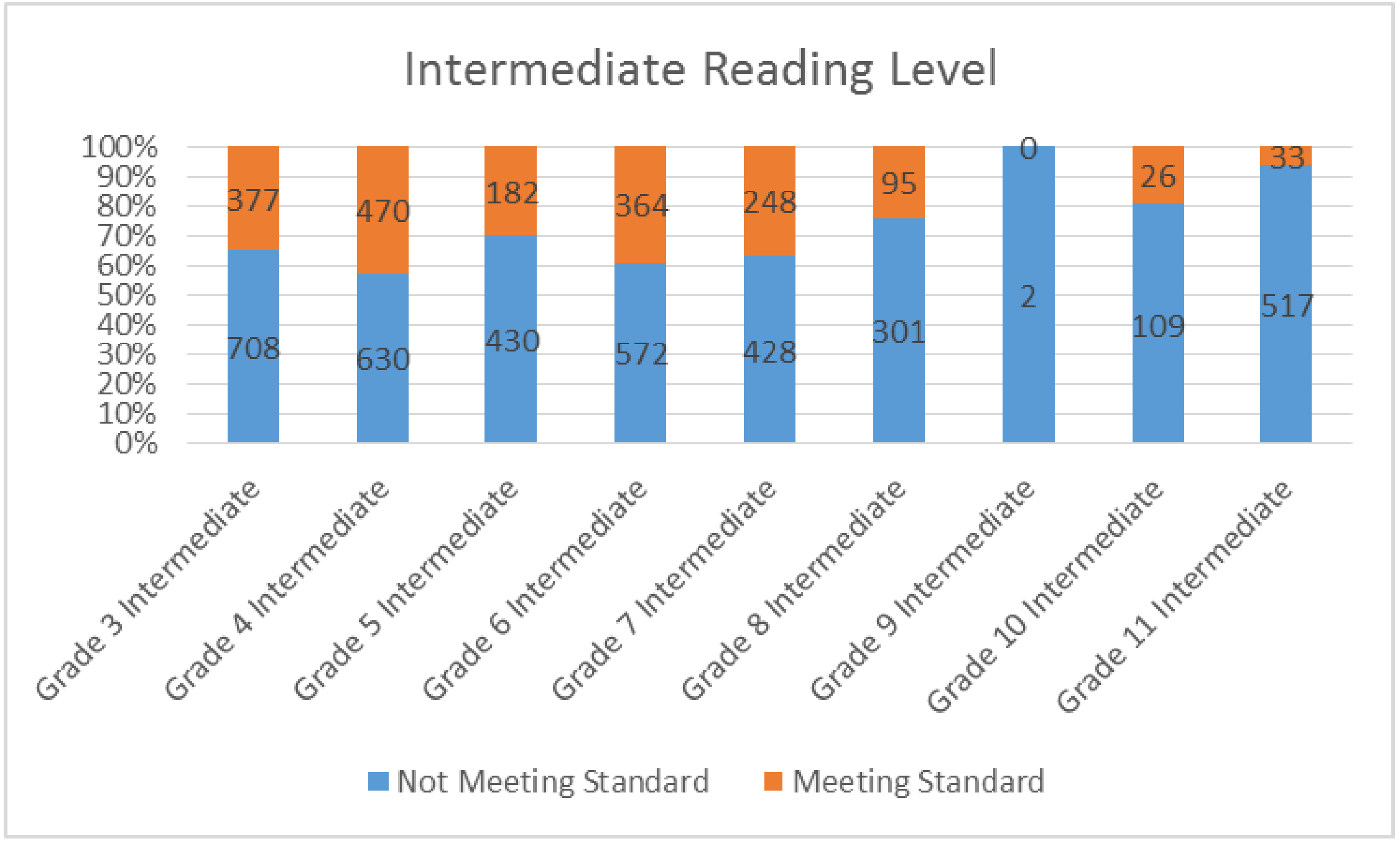
- What did you notice?

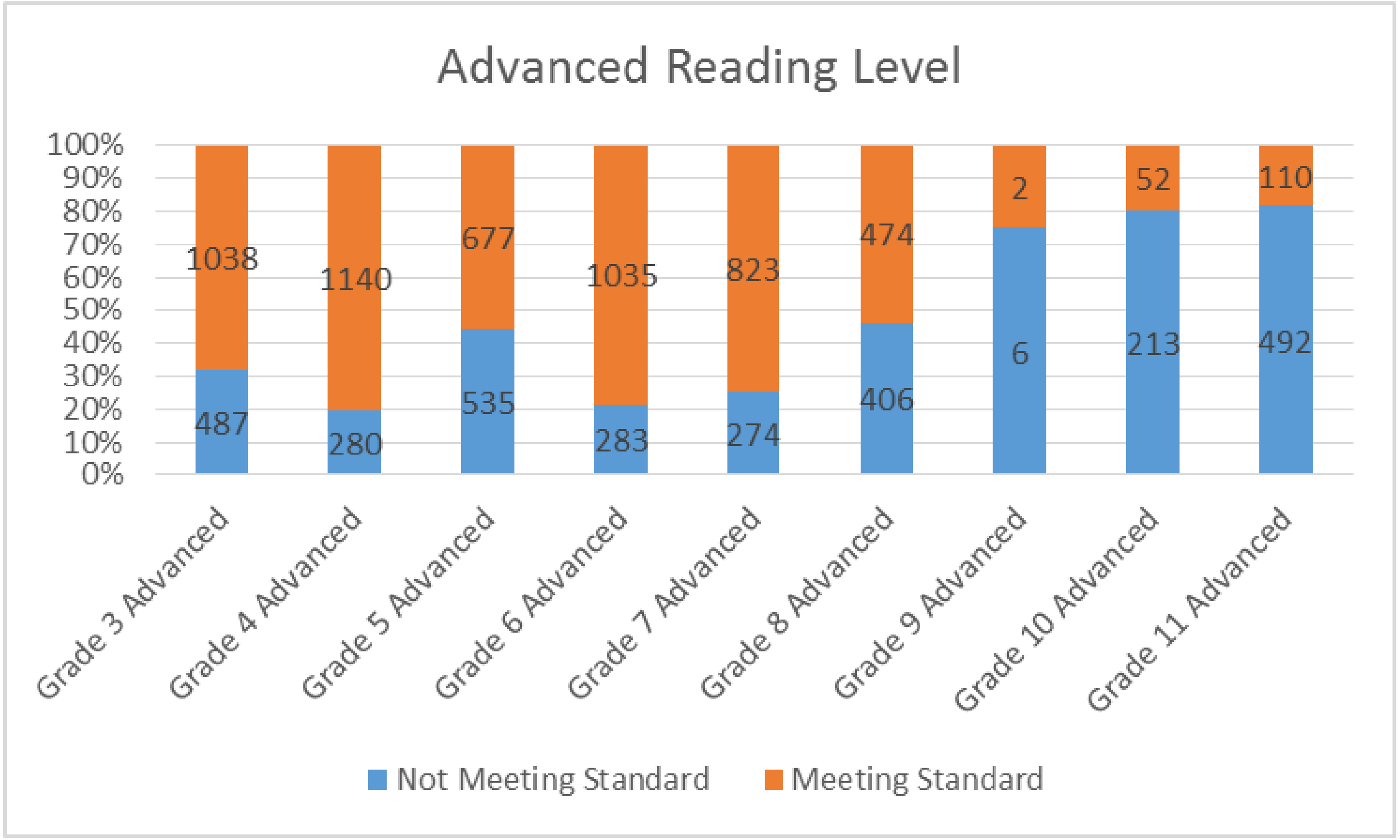
- I noticed…
  - Beginning students did the worst as a group through all grades
  - As the grade went up the scores tended to go down across all proficiencies
  - There is a big difference between Fluent and Beginning performance
  - Fluent had the best performance overall, but…
    - It got worse as the grades went up
    - Why?
      - Content gets harder as well…

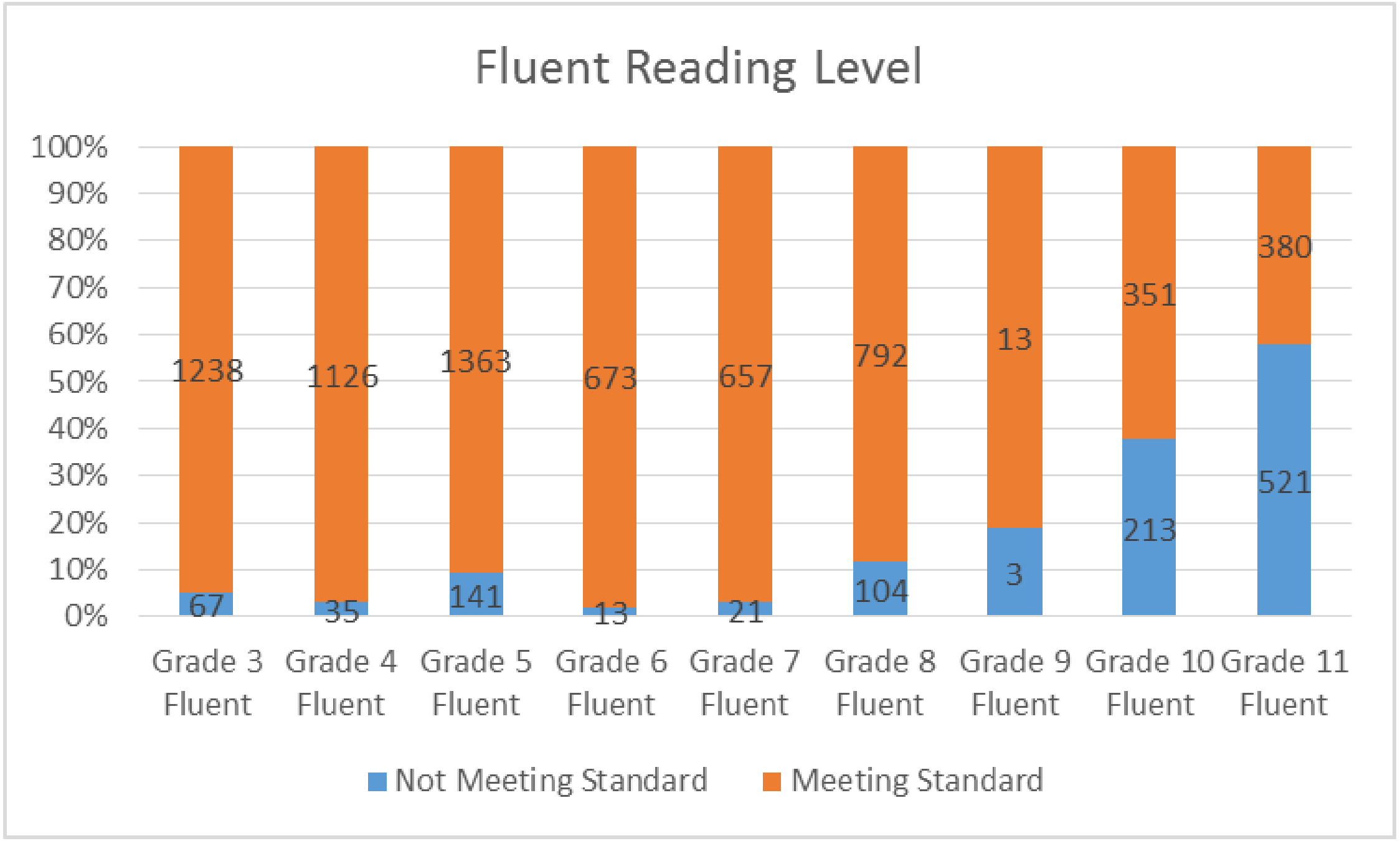
Reading

- Again look for patterns in the next few charts…


Beginning Reading Level
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
62
67
22
22
17
12
0
2
126
167
88
70
47
58
0
10
66
Grade 3 Beginning
Grade 4 Beginning
Grade 5 Beginning
Grade 6 Beginning
Grade 7 Beginning
Grade 8 Beginning
Grade 9 Beginning
Grade 10 Beginning
Grade 11 Beginning
Not Meeting Standard
Meeting Standard


Intermediate Reading Level
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
377
470
182
364
248
95
0
26
33
708
630
430
572
428
301
2
109
517
Grade 3 Intermediate
Grade 4 Intermediate
Grade 5 Intermediate
Grade 6 Intermediate
Grade 7 Intermediate
Grade 8 Intermediate
Grade 9 Intermediate
Grade 10 Intermediate
Grade 11 Intermediate
Not Meeting Standard
Meeting Standard


Advanced Reading Level
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
1038
487
1140
280
677
535
1035
283
823
274
474
406
2
6
52
213
110
492
Grade 3 Advanced
Grade 4 Advanced
Grade 5 Advanced
Grade 6 Advanced
Grade 7 Advanced
Grade 8 Advanced
Grade 9 Advanced
Grade 10 Advanced
Grade 11 Advanced
Not Meeting Standard
Meeting Standard


Fluent Reading Level
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
1238
1126
1363
673
657
792
13
351
380
67
35
141
13
21
104
3
213
521
Grade 3 Fluent
Grade 4 Fluent
Grade 5 Fluent
Grade 6 Fluent
Grade 7 Fluent
Grade 8 Fluent
Grade 9 Fluent
Grade 10 Fluent
Grade 11 Fluent
Not Meeting Standard
Meeting Standard

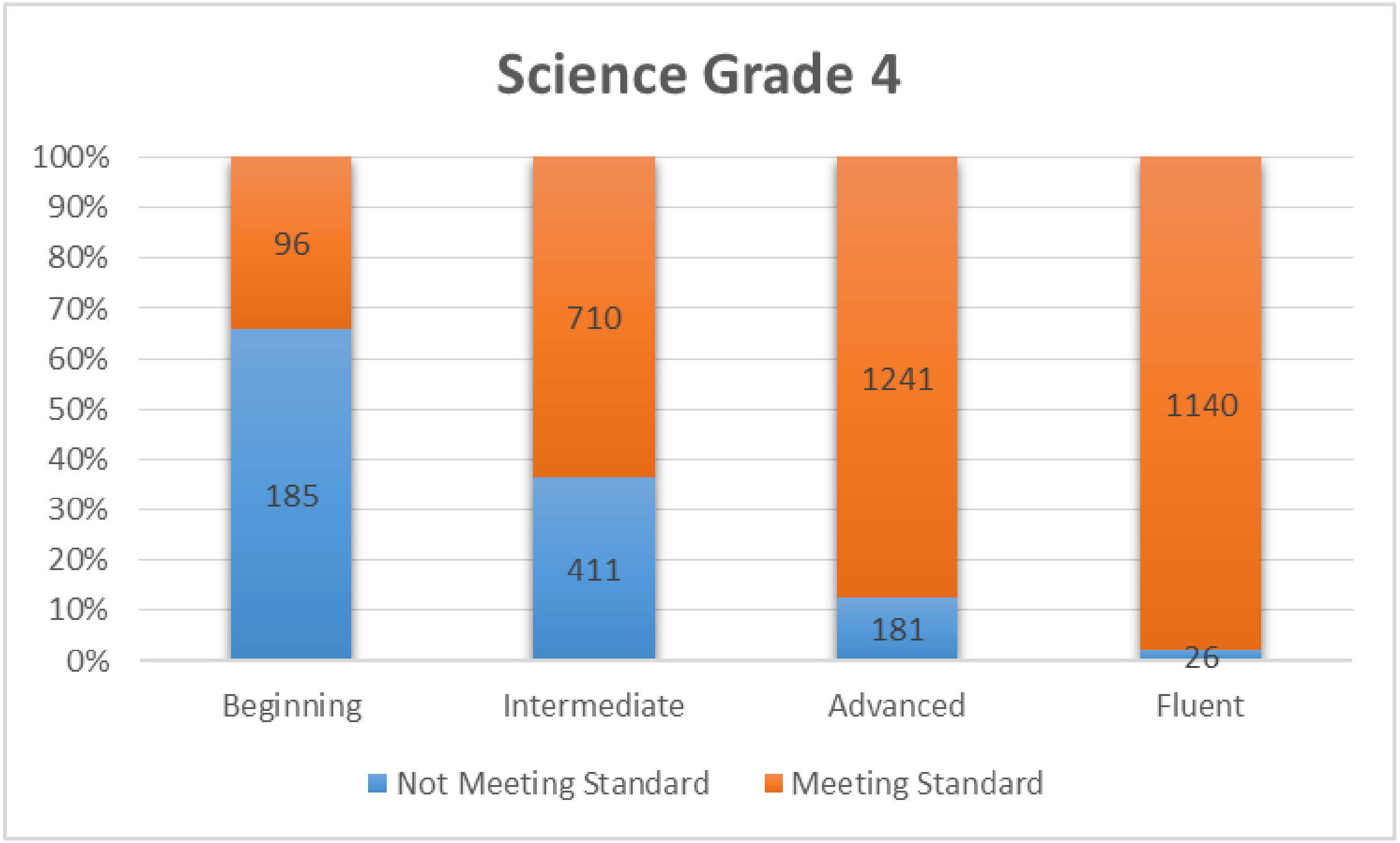
Patterns?

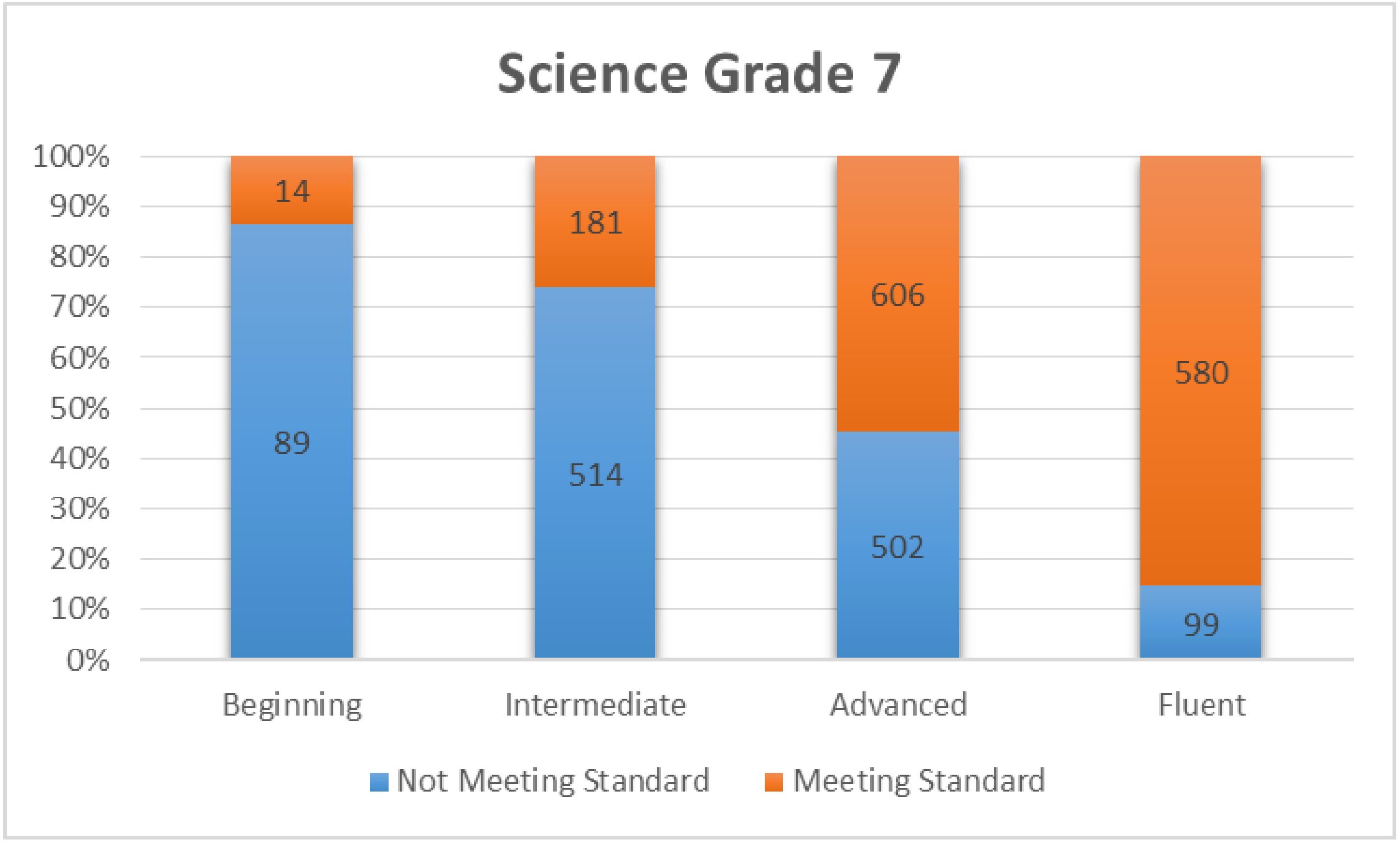
- What did you notice?

- I noticed…
  - Beginning students did the worst as a group through all grades
  - As the grade went up the scores tended to go down across all proficiencies
  - There is a big difference between Fluent and Beginning performance
  - Fluent had the best performance overall, but…
    - It got worse as the grades went up even more so than Math
    - Why?
      - Content gets harder and is directly linked to English ability…

Science

- Here we only have two grades to look at..


Science Grade 4
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
96
185
710
411
1241
181
1140
26
Beginning
Intermediate
Advanced
Fluent
Not Meeting Standard
Meeting Standard

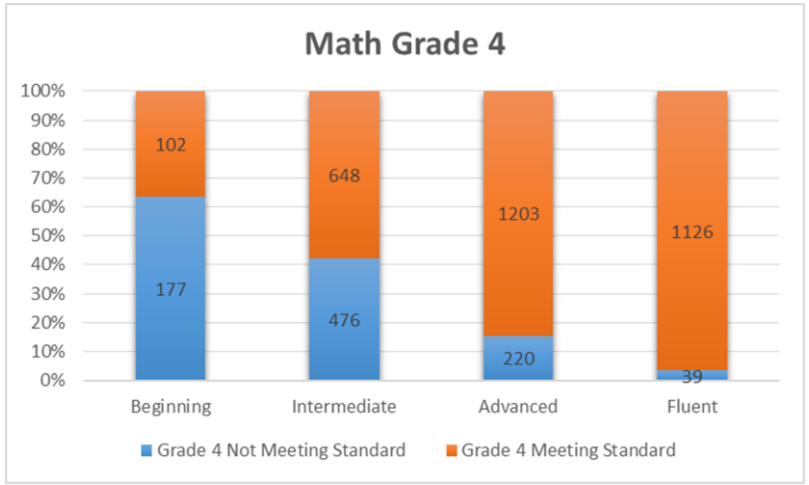

Science Grade 7
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
14
89
181
514
606
502
580
99
Beginning
Intermediate
Advanced
Fluent
Not Meeting Standard
Meeting Standard

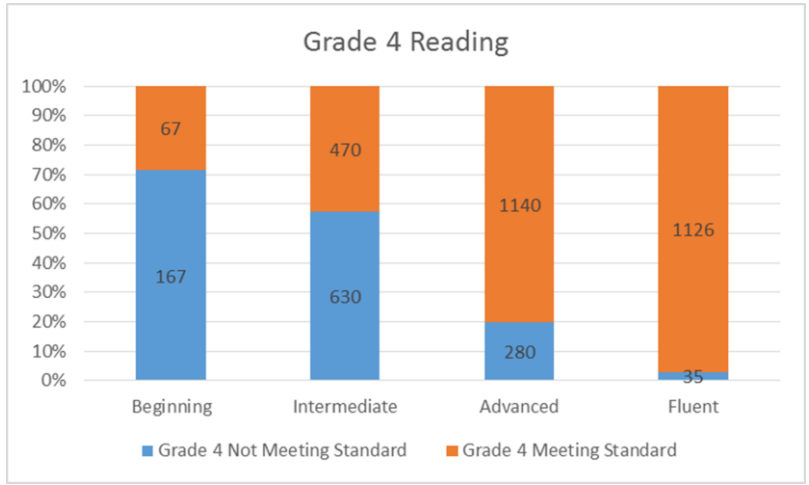
Patterns?

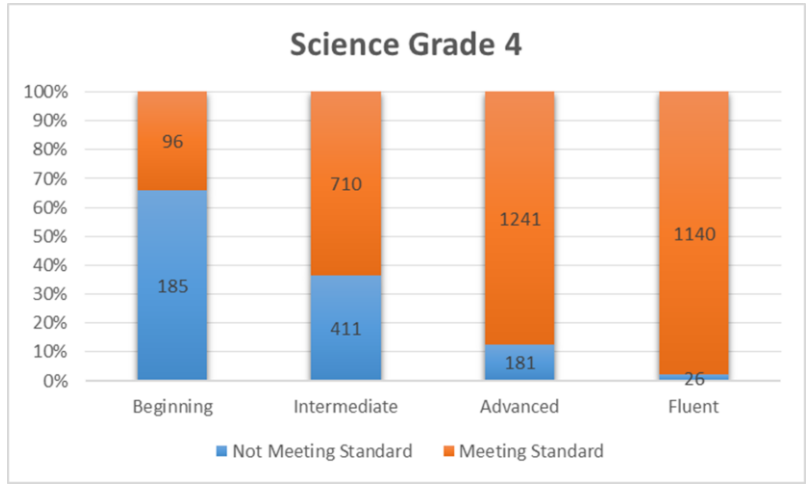
- What did you notice?

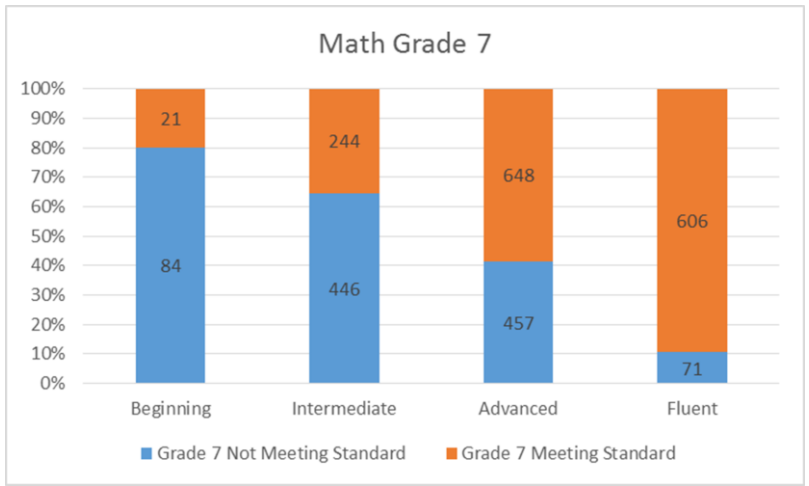
- I noticed…
  - Again as the proficiency went up the students performed better
  - From 4th to 7th grade less Fluent students meet standards
  - There is a big difference between Fluent and Beginning performance
  - Fluent had the best performance overall, but…
    - It got worse as the grades went up
    - Why?
      - Content gets harder

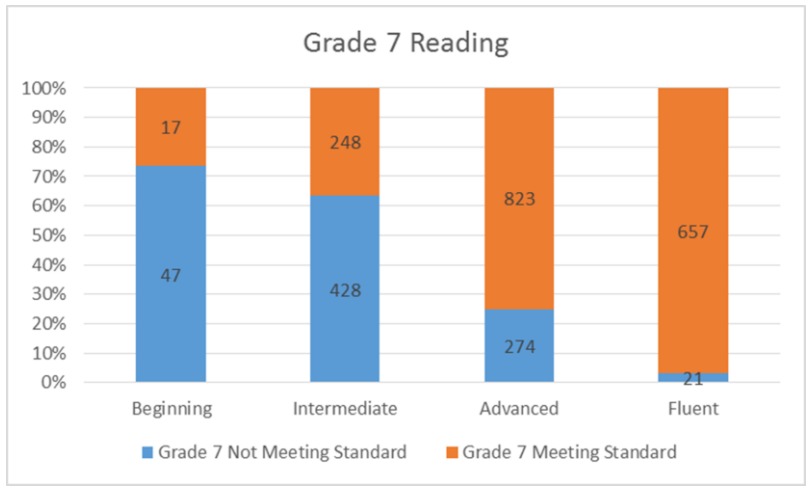
What patterns do you see in the scores for 4th and 7th grades?

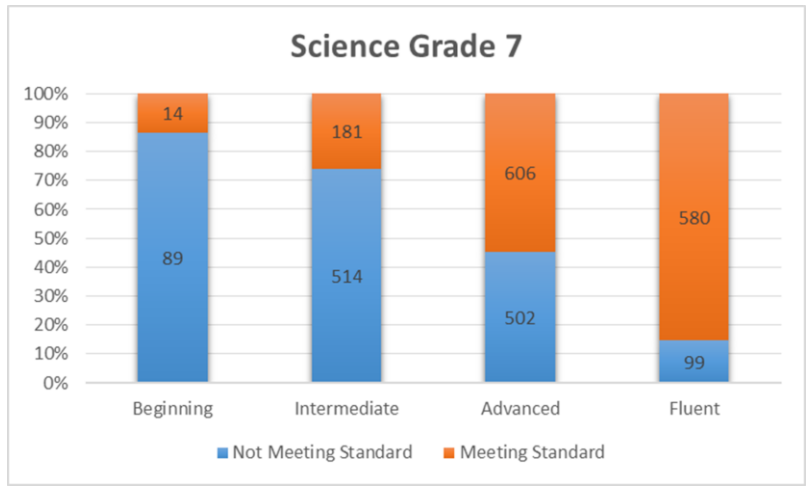












Patterns

- The pattern seems pretty obvious
  - As the grade goes up the number of students Not Meeting Standards also goes up
  - As proficiency goes up the number of students Meeting Standards also goes up
  - There is the largest difference between Fluent and Beginning

RQ#2

- What role do other demographic variables (such as free and reduced lunch, native language, gender, length of time in the US, or learning disability) play in student achievement on content assessments for English Language Learners?

- Methods-
- Set of multiple linear regression analysis
  - Dependent Variable- content assessment score
  - Independent Variable- student proficiency group **and** then each of the demographic variables

First we will look at overall performance

- Each instance when a demographic was identified as being significant (Pr > F) of 5% or less was compiled
- This was done for each Content Area (Math, Reading, & Science)
- All grades were grouped together by Content Area
- The Over-all Predictor Quality was described using this information

Predictor Quality	
Blue = High (100-70)	Grey = Low (29-1)
Orange = Medium (69-30)	Pink = None

Type III SS with (Pr > F) of less than 5%					
Demographic Names	Math % all grades	Reading % all grades	Science % all grades	Total % of all	Over-all Predictor Quality
One Way Relationships					
except	78%	67%	100%	81%	High
totalcat	78%	67%	80%	75%	High
years	89%	22%	100%	70%	High
lunch	78%	67%	60%	68%	Medium
lang	78%	67%	60%	68%	Medium
gender	44%	44%	100%	63%	Medium

Predictor Quality	
Blue = High (100-70)	Grey = Low (29-1)
Orange = Medium (69-30)	Pink = None

Type III SS with (Pr > F) of less than 5%					
Demographic Names	Math % all grades	Reading % all grades	Science % all grades	Total % of all	Over-all Predictor Quality
Two Way Relationships					
totalcat*except	56%	67%	80%	67%	Medium
lang*years	56%	0%	40%	32%	Medium
gender*except	11%	22%	60%	31%	Medium
totalcat*lang	11%	33%	20%	21%	Low
lang*gender	22%	22%	20%	21%	Low
totalcat*lunch	22%	11%	20%	18%	Low
years*except	11%	22%	20%	18%	Low
totalcat*gender	0%	0%	40%	13%	Low
lunch*gender	0%	11%	20%	10%	Low
lunch*lang	22%	0%	0%	7%	Low
lunch*years	22%	0%	0%	7%	Low
lunch*except	0%	0%	20%	7%	Low
gender*years	0%	0%	20%	7%	Low
totalcat*years	0%	0%	0%	0%	None
lang*except	0%	0%	0%	0%	None

Predictor Quality
Blue = High (100-70)
Orange = Medium (69-30)
Grey = Low (29-1)
Pink = None

Predictability

- The best single demographics for predictability are (above 70%):
  - Exceptionality (Learning Disabled, Gifted, None, or Other)
  - Total Proficiency Category
  - Number of Years in the U.S.
- There are no best categories for the two-way the highest is:
  - Total Proficiency Category * Exceptionality (67% over-all)

- Let’s look at the breakdown of the demographics

Exceptionality									
	Math			**Reading**			**Science**		
	Estimate	Pr > F	R-Score	Estimate	Pr > F	R-Score	Estimate	Pr > F	R-Score
Grade 3			31%	-5.8765	<.0001	41%			N/A
Grade 4	-2.8574	0.0053	33%	-0.2374	<.0001	33%	-3.5882	0.0036	44%
Grade 5	-4.1481	0.0003	30%	-4.8852	<0.0001	44%			N/A
Grade 6	-6.2981	<0.0001	34%	-4.8830	0.0002	45%			N/A
Grade 7	-5.1562	0.0014	30%			49%	-4.8295	0.0010	37%
Grade 8	-4.7721	0.0141	26%	-8.6270	<.0001	46%			N/A
Grade 9			19%	12.0184	0.0336	63%	-13.4391	0.0016	36%
Grade 10	-3.2436	0.0268	29%	-16.9082	<.0001	39%	-4.6795	0.0075	16%
Grade 11	2.0296	0.0002	19%			39%	-3.4417	0.0379	21%

- Universally applicable
- Mostly Negative relationships
- Very significant (Pr > F less than 1%)

Total Proficency Category									
	Math			Reading			Science		
	Estimate	Pr > F	R-Score	Estimate	Pr > F	R-Score	Estimate	Pr > F	R-Score
Grade 3	9.2702	<0.0001	31%	6.6749	<0.0001	41%			N/A
Grade 4	5.9462	<0.0001	33%	0.6238	<.0001	33%	5.1944	0.0003	44%
Grade 5	4.4226	0.0011	30%	4.8709	0.0215	44%			N/A
Grade 6			34%			45%			N/A
Grade 7	5.1981	0.0083	30%	13.5756	<0.0001	49%	4.7578	0.0039	37%
Grade 8	4.5991	0.0228	26%	5.5644	0.0007	46%			N/A
Grade 9	13.9182	<.0001	19%			63%			36%
Grade 10	10.6517	<.0001	29%			39%	7.1134	<.0001	16%
Grade 11			19%	10.9264	<.0001	39%	8.7680	0.0008	21%

- Universally applicable
- Entirely positive relationships
- Very significant (Pr > F less than 1%)

Number of Years in the U.S.									
	Math			**Reading**			**Science**		
	Estimate	Pr > F	R-Score	Estimate	Pr > F	R-Score	Estimate	Pr > F	R-Score
Grade 3			31%			41%			N/A
Grade 4	0.1143	0.0111	33%			33%	0.4122	0.0004	44%
Grade 5	-0.6321	0.0298	30%			44%			N/A
Grade 6	-1.2020	0.0049	34%			45%			N/A
Grade 7	-2.5474	<0.0001	30%			49%	-1.5095	<0.0001	37%
Grade 8	-2.0645	<.0001	26%	0.4743	0.0026	46%			N/A
Grade 9	-0.3835	<.0001	19%			63%	0.2670	<.0001	36%
Grade 10	-3.1872	<.0001	29%	0.0778	0.0093	39%	0.0790	0.0036	16%
Grade 11	0.1311	<.0001	19%			39%	-1.3635	<.0001	21%

- More applicable for Math and Science (not much for Reading)
- Mostly Negative relationships
- Very significant (Pr > F less than 1%) Math & Science

Free and Reduced Lunch Program Eligibility									
	Math			**Reading**			**Science**		
	Estimate	Pr > F	R-Score	Estimate	Pr > F	R-Score	Estimate	Pr > F	R-Score
Grade 3			31%	-2.9557	<0.0001	41%			N/A
Grade 4			33%			33%	-3.0859	<.0001	44%
Grade 5	-1.5662	0.0209	30%			44%			N/A
Grade 6	-10.9333	0.0010	34%	-9.4489	0.0028	45%			N/A
Grade 7	-2.6347	0.0029	30%	-2.3144	0.0010	49%	-3.8119	<0.0001	37%
Grade 8	-2.6553	0.0129	26%	-2.8237	0.0014	46%			N/A
Grade 9	-66.1633	0.0356	19%			63%			36%
Grade 10	-5.6149	<.0001	29%	-4.0528	0.0004	39%	-14.3022	0.0158	16%
Grade 11	-14.1405	0.0084	19%	-6.2318	<.0001	39%			21%

- Most applicable to Reading then Math; Not very for Science
- Entirely Negative relationships
- Very significant (Pr > F less than 1%) Reading & somewhat Math

First Language									
	Math			**Reading**			**Science**		
	Estimate	Pr > F	R-Score	Estimate	Pr > F	R-Score	Estimate	Pr > F	R-Score
Grade 3	-0.2738	0.0052	31%	-0.5036	0.0002	41%			N/A
Grade 4	-0.2785	0.0080	33%	6.3095	0.0407	33%	-0.5619	<.0001	44%
Grade 5	-0.5339	<0.0001	30%	-1.1727	0.0104	44%			N/A
Grade 6	-0.7167	0.0002	34%	-1.3341	0.0034	45%			N/A
Grade 7	-1.1397	<0.0001	30%			49%	-0.8398	<0.0001	37%
Grade 8	-1.1974	<.0001	26%	-0.3755	0.0095	46%			N/A
Grade 9			19%			63%			36%
Grade 10			29%			39%			16%
Grade 11	-1.4506	0.0148	19%	0.4002	0.0240	39%	1.0395	0.0167	21%

- Most significant in Math then Reading; Not very in Science
- Mostly Negative relationships
- Very significant (Pr > F less than 1%) Math

Gender									
	Math			Reading			Science		
	Estimate	Pr > F	R-Score	Estimate	Pr > F	R-Score	Estimate	Pr > F	R-Score
Grade 3	2.0035	<0.0001	31%	-5.9785	0.0132	41%			N/A
Grade 4	2.5902	<0.0001	33%	-4.0051	0.0020	33%	9.1613	0.0002	44%
Grade 5			30%			44%			N/A
Grade 6	1.2272	0.0137	34%			45%			N/A
Grade 7			30%	1.2015	0.0068	49%	-6.6878	0.0068	37%
Grade 8			26%			46%			N/A
Grade 9			19%			63%			36%
Grade 10			29%	-15.5454	0.0008	39%	4.4694	<.0001	16%
Grade 11	11.1023	0.0052	19%			39%	16.9798	0.0002	21%

- Negative and positive relationships
- Very significant (Pr > F less than 1%) when it appears usually significant (or just above 1%)

Exceptionality * Total Category									
	Math			Reading			Science		
	Estimate	Pr > F	R-Score	Estimate	Pr > F	R-Score	Estimate	Pr > F	R-Score
Grade 3			31%	1.3224	0.0041	41%			N/A
Grade 4	1.2357	0.0101	33%	1.3367	0.0011	33%	1.4638	0.0026	44%
Grade 5	1.6877	0.0003	30%	1.3121	0.0017	44%			N/A
Grade 6	2.5020	0.0001	34%	1.5207	0.0078	45%			N/A
Grade 7	2.0268	0.0026	30%			49%	1.4481	0.0092	37%
Grade 8	2.1122	0.0025	26%	2.9604	<.0001	46%			N/A
Grade 9			19%			63%	4.2674	0.0013	36%
Grade 10			29%	3.8265	0.0001	39%			16%
Grade 11			19%			39%	2.0398	0.0004	21%

- Universally applicable
- Entirely Positive relationships
- Very significant (Pr > F less than 1%)

Exceptionality * Gender									
	Math			Reading			Science		
	Estimate	Pr > F	R-Score	Estimate	Pr > F	R-Score	Estimate	Pr > F	R-Score
Grade 3			31%			41%			N/A
Grade 4			33%	-1.9988	0.0038	33%	-2.0014	0.0149	44%
Grade 5			30%			44%			N/A
Grade 6			34%			45%			N/A
Grade 7			30%			49%	1.6292	0.0348	37%
Grade 8			26%			46%			N/A
Grade 9			19%			63%			36%
Grade 10	4.0937	0.0258	29%	5.0127	0.0017	39%			16%
Grade 11			19%			39%	-2.1726	0.0339	21%

- Not very Universally applicable
- Positive and Negative relationships
- Overall not a very good predictor

Exceptionality * Number of Years in the U.S.									
	Math			**Reading**			**Science**		
	Estimate	Pr > F	R-Score	Estimate	Pr > F	R-Score	Estimate	Pr > F	R-Score
Grade 3			31%			41%			N/A
Grade 4			33%			33%	-0.1250	0.0012	44%
Grade 5			30%			44%			N/A
Grade 6			34%			45%			N/A
Grade 7			30%			49%			37%
Grade 8	-0.1381	0.0372	26%	-0.1546	0.0040	46%			N/A
Grade 9			19%			63%			36%
Grade 10			29%			39%			16%
Grade 11			19%	0.1155	0.0294	39%			21%

- Not very Universally applicable
- Mostly Negative relationships
- Overall not a very good predictor

Exceptionality * Free and Reduced Lunch Eligibility									
	Math			**Reading**			**Science**		
	Estimate	Pr > F	R-Score	Estimate	Pr > F	R-Score	Estimate	Pr > F	R-Score
Grade 3			31%			41%			N/A
Grade 4			33%			33%			44%
Grade 5			30%			44%			N/A
Grade 6			34%			45%			N/A
Grade 7			30%			49%			37%
Grade 8			26%			46%			N/A
Grade 9			19%			63%			36%
Grade 10			29%			39%	4.6486	0.0223	16%
Grade 11			19%			39%			21%

- Not a good predictor

Exceptionality * Language									
	Math			**Reading**			**Science**		
	Estimate	Pr > F	R-Score	Estimate	Pr > F	R-Score	Estimate	Pr > F	R-Score
Grade 3			31%			41%			N/A
Grade 4			33%			33%			44%
Grade 5			30%			44%			N/A
Grade 6			34%			45%			N/A
Grade 7			30%			49%			37%
Grade 8			26%			46%			N/A
Grade 9			19%			63%			36%
Grade 10			29%			39%			16%
Grade 11			19%			39%			21%

- No instances of predictability

Findings

- Beginning students uniformly perform poorly on content assessments
- Fluent students perform best on content assessments
- As the grade goes up the percent of students meeting standards goes down
- There are a few good predictors for score
  - Exceptionality code (Learning Disabled, Gifted, N/A, or Other)
  - English Proficiency (Beginning, Intermediate, Advanced, and Fluent)
  - Number of years in the U.S.
  - Exceptionality and English Proficiency

Questions?

Thank you for coming!

Christina Kitson

clkit96@gmail.com