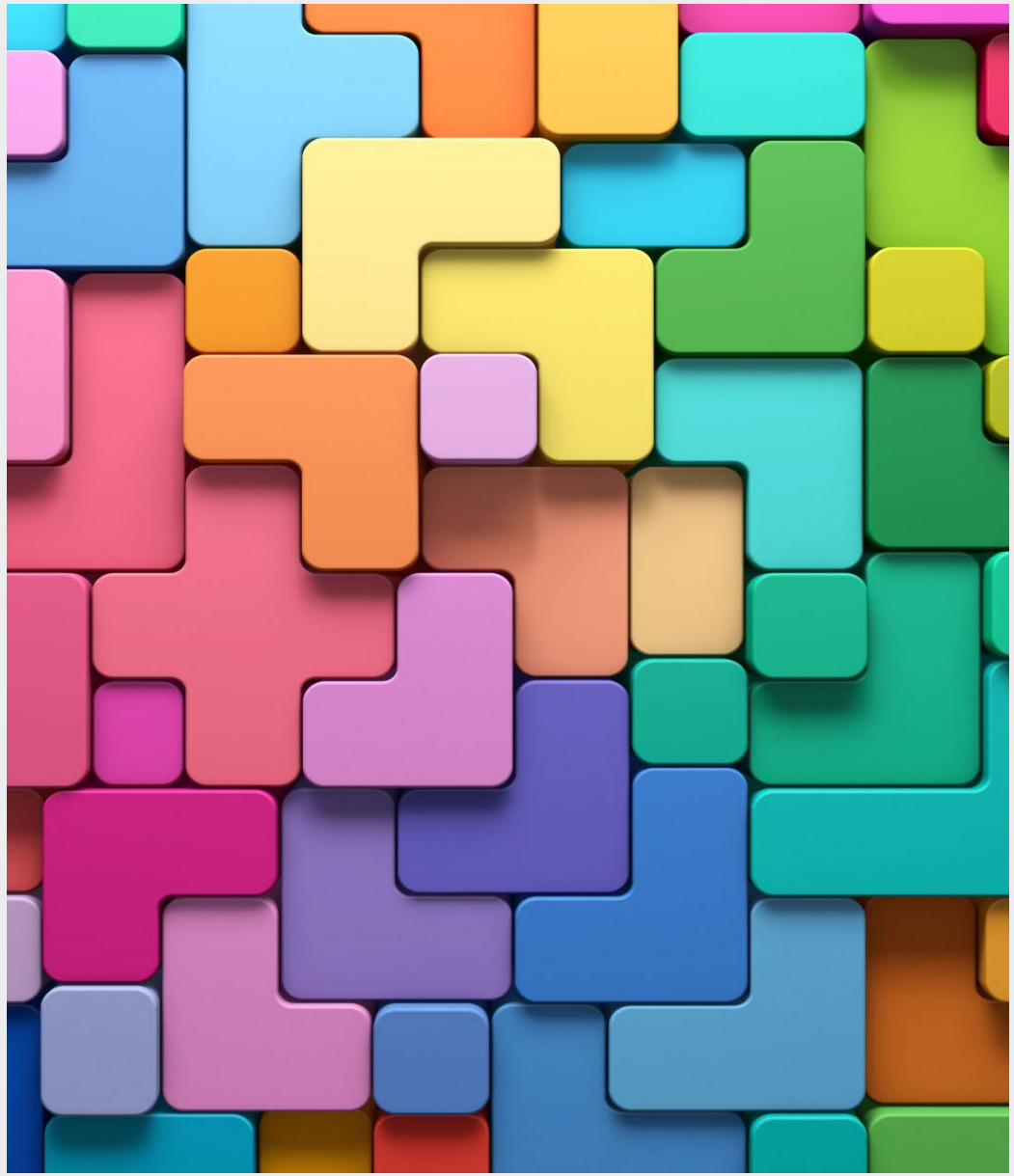
Completing the Data Puzzle

What happens when the puzzle pieces don't quite fit? Do our statistical models always work?



Fred Cohen
Nassau BOCES
fcohen@nasboces.org

Do the data reports we use actually measure what we think they measure?

Remember that the goal must be the improvement of teaching and learning.

And that is why we need “Fat Tony.”

Fat Tony was created by economist Nassim Taleb who suggests that statistical models don’t always work in complex situations (like schools).

Fat Tony is a man who lives by his wits in contrast to “**Dr John,**” a man of pure science and logic.

They are presented with the following problem: If a fair coin is flipped 99 times and comes up heads, what are the odds that the 100th flip will also be heads?

Dr. John: “50%, the previous 99 flips have no effect on the next flip.”

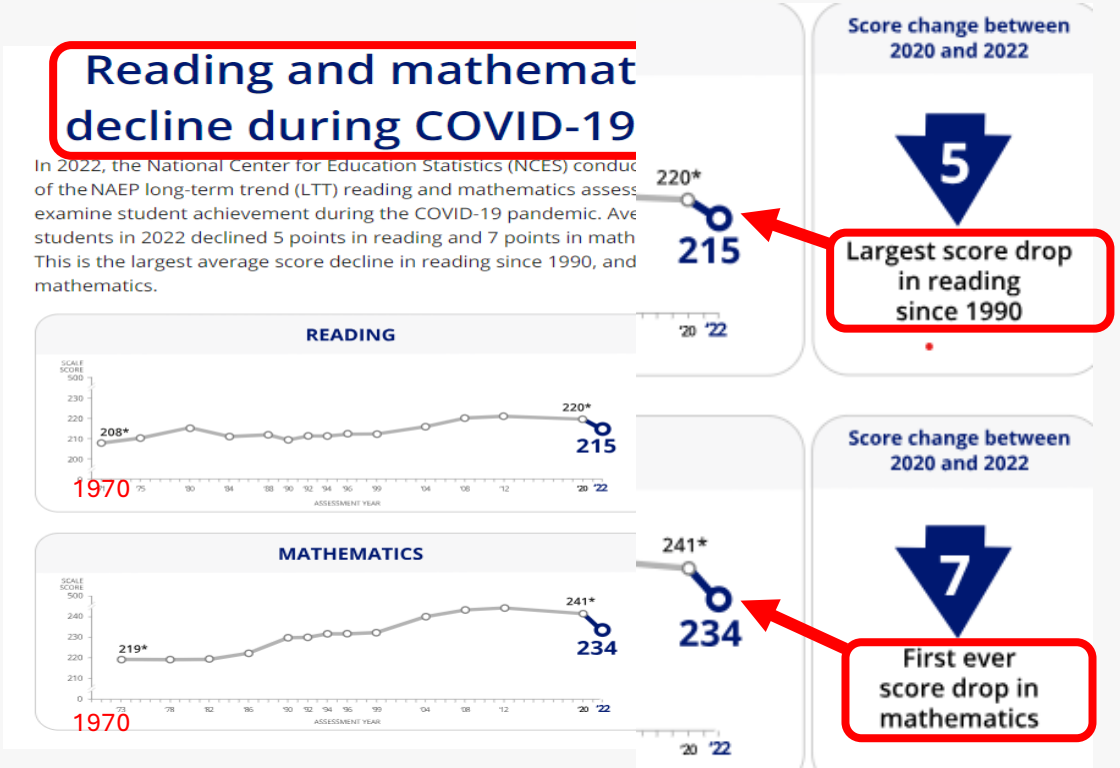
Fat Tony: “100% heads. It’s not a fair coin!”

When seeking answers to some school data problems, Fat Tony's "smarts" are sometimes needed.

Here are four problems to puzzle everyone.

Problem 1: Since the pandemic, what has happened to **graduation rates** and **Advanced Regents Diploma rates**? What would Dr. John say? What data might he reference?

The National Assessment on Educational Progress (NAEP) is the gold standard. He might cite the long-term trends for Reading and Math scores .

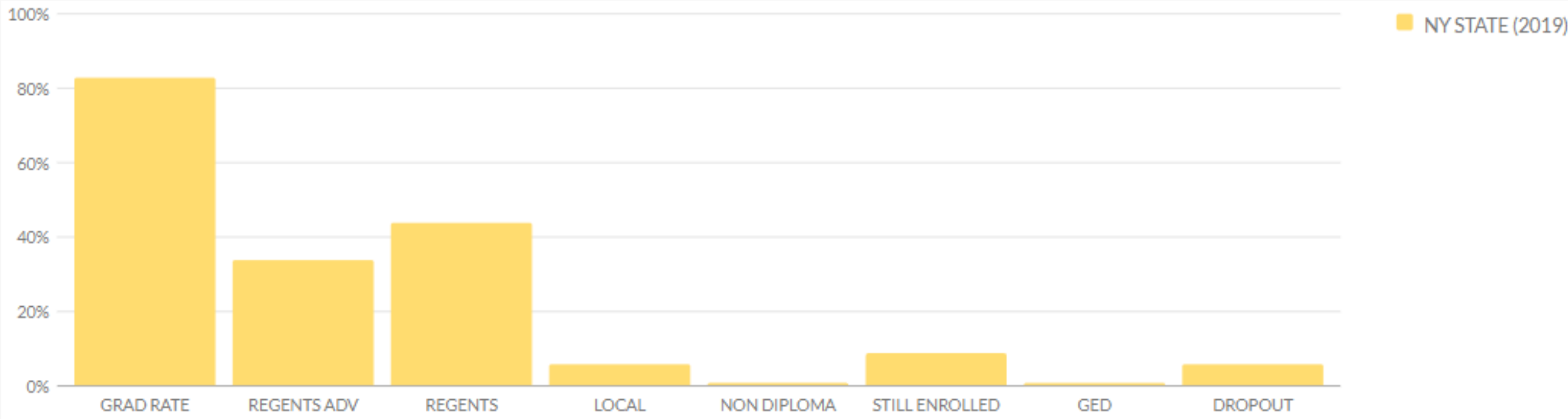


Based on the NAEP's overwhelming data, we might predict Dr. John's response.

But what does **Sam Smart** say? Like Fat Tony, Sam Smart uses data but has also been a teacher for many years.

Sam Smart says, "Let's go to the School Report Card."

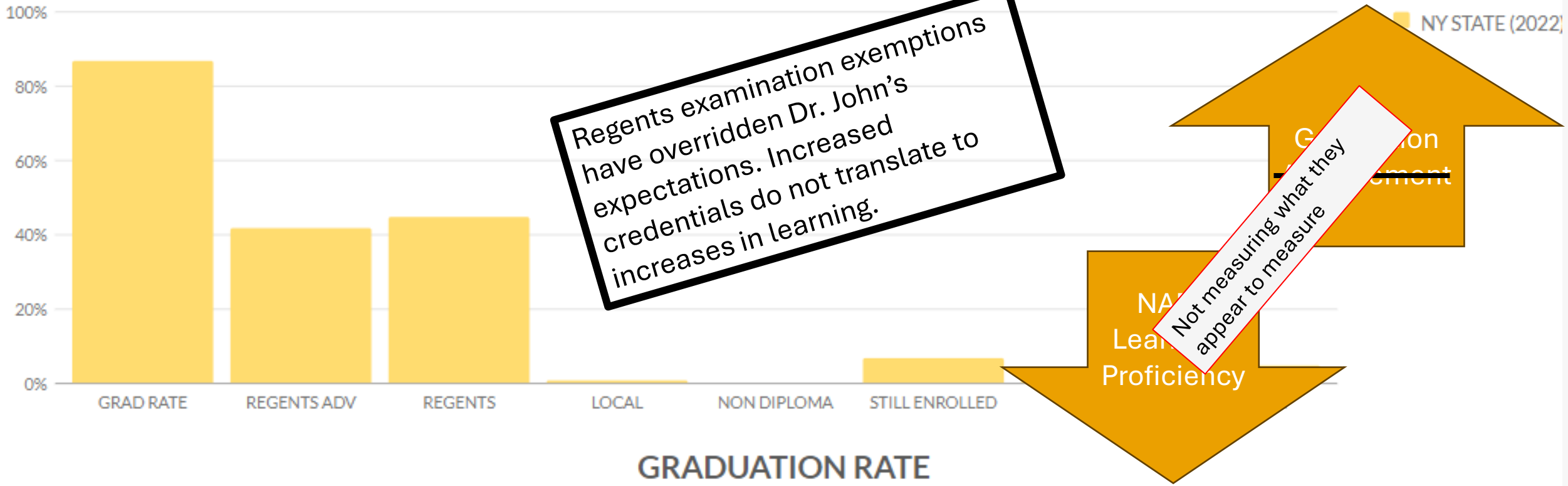
NY STATE GRADUATION RATE DATA
4 YEAR OUTCOME AS OF AUGUST 2019



GRADUATION RATE

Subgroup	Total	GRAD RATE		REGENTS WITH ADVANCED DESIGNATION		REGENTS DIPLOMA		LOCAL DIPLOMA		NON DIPLOMA CRED		STILL ENROLLED		GED TRANSFER		DROPOUT	
		#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%
All Students	207,826	173,276	83%	69,751	34%	90,882	44%	12,643	6%	1,200	1%	19,057	9%	1,419	1%	12,776	6%

NY STATE GRADUATION RATE DATA
4 YEAR OUTCOME AS OF AUGUST 2022

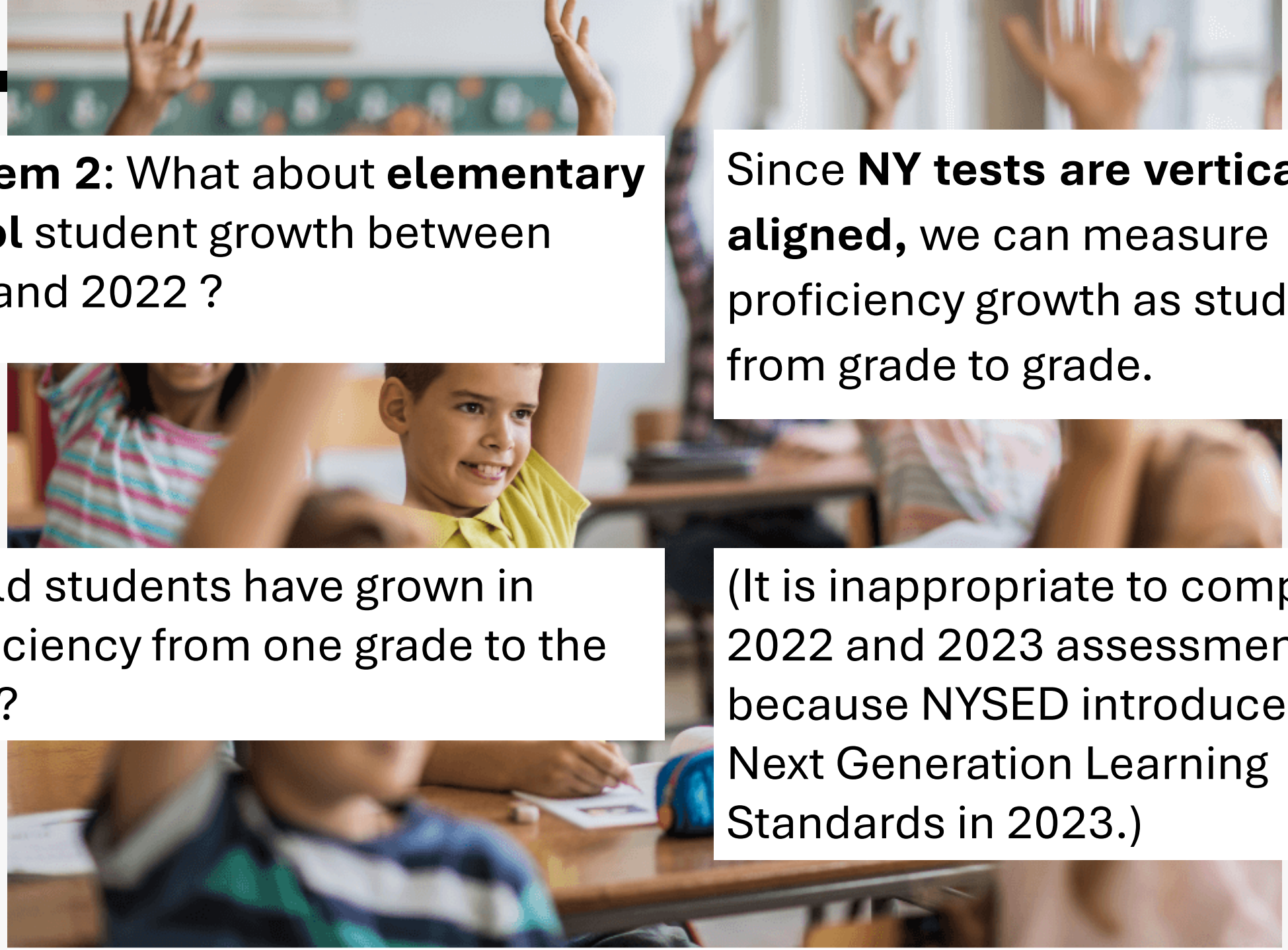


Regents examination exemptions have overridden Dr. John's expectations. Increased credentials do not translate to increases in learning.

Not measuring what they appear to measure

GRADUATION RATE

Subgroup	Total	GRAD RATE		REGENTS WITH ADVANCED DESIGNATION		REGENTS DIPLOMA		LOCAL DIPLOMA		NON DIPLOMA CRED		STILL ENROLLED		GED TRANSFER		DROPOUT	
All Students 2019	207,826	173,276	83%	69,751	34%	90,882	44%	12,643	6%	1,200	1%	19,057	9%	1,419	1%	12,776	6%
All Students	202,097	175,886	87%	84,544	42%	90,113	45%	1,229	1%	1,010	0%	14,383	7%	1,116	1%	9,605	5%



Problem 2: What about **elementary school** student growth between 2021 and 2022 ?

Since **NY tests are vertically aligned,** we can measure proficiency growth as students go from grade to grade.

Could students have grown in proficiency from one grade to the next?

(It is inappropriate to compare 2022 and 2023 assessments because NYSED introduced the Next Generation Learning Standards in 2023.)

At the Nassau RIC, we have a report that compares **test takers going from one grade to the next.**

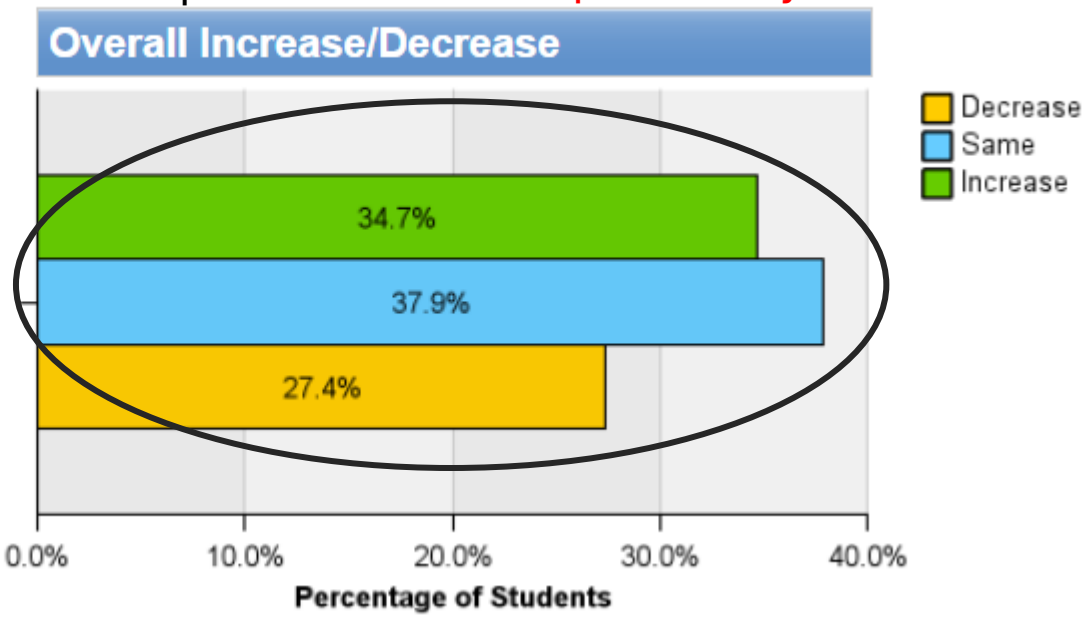
Called the “Half Level Change—This Year vs. Last Year Report,” **it measures changes in Performance Level for students moving from grade to grade**.

Half Level Performance Level Change Report for the Same Students

Count of Students			2020/2021 Grade 5 ELA						
			Level 1	Level 2 L	Level 2H	Level 3 L	Level 4 L	Level 4H	Total
2021/2022	Grade 6 ELA	Level 1	4	2		2			8
		Level 2 L	3	4					7
		Level 2H	2	4		6	2		14
		Level 3 L	1	2	3	10	2	1	19
		Level 3H	1	2	2	8	4	2	19
		Level 4 L		1	4	9	27	13	54
		Level 4H					1	2	3
		Total	11	15	9	35	36	18	124

16 became proficient

10 lost proficiency



6/124=5% increase in proficiency

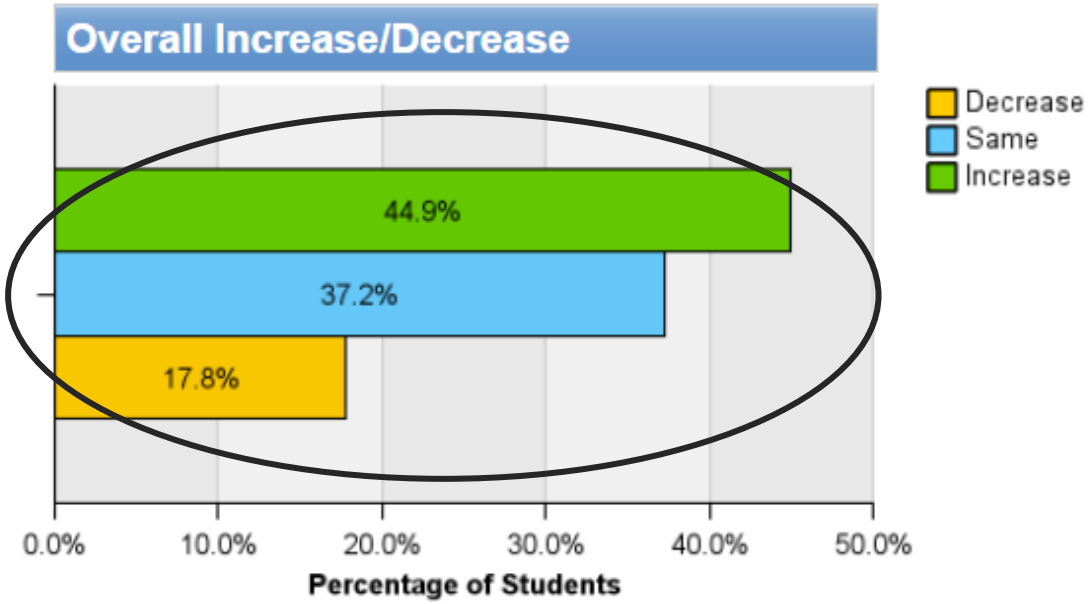
Half Level Performance Level Change Report for the Same Students

Count of Students			2020/2021 Grade 5 ELA						
			Level 1	Level 2 L	Level 2H	Level 3 L	Level 4 L	Level 4H	Total
2021/2022	Grade 6 ELA	Level 1	4	3	2				9
		Level 2 L	3	3		1	1		8
		Level 2H		6	1	1	5		13
		Level 3 L	1	8	4	9	4	1	27
		Level 3H		7	2	11	6		26
		Level 4 L		14	6	38	69	20	147
		Level 4H				6	5	6	17
		Total	8	41	15	66	90	27	247

42 became proficient

8 lost proficiency

34/247=14% increase in proficiency



District after district shows students **grew** in proficiency from Grade 5 to Grade 6 from 2021 to 2022.

Is this growth real or is something flawed in the statistical model?

Fat Tony knows, “When coins aren’t “fair,” statistical models don’t work!”

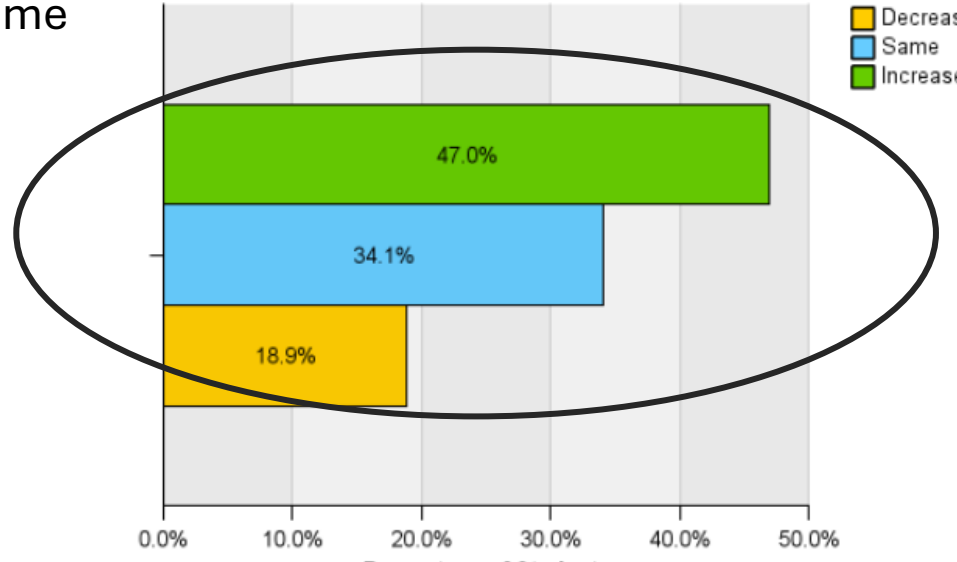
And Sam Smart knows, “If tests aren’t vertically aligned, real student growth is not what is being measured!”

Here is some additional evidence Sam found from before the pandemic.

Half Level Change-This Year vs Last Year (Nassau Totals)

Count of Students			2017/2018 Grade 5 ELA							
			Level 1	Level 2 L	Level 2H	Level 3 L	Level 3H	Level 4 L	Level 4H	Total
2018/2019	Grade 6 ELA	Level 1	853	283	121	58	13	8		1,336
		Level 2 L	244	254	178	106	37	19		838
		Level 2H	104	183	165	142	39	32	1	666
		Level 3 L	92	224	316	288	139	125	4	1,188
		Level 3H	17	60	90	143	56	75	1	442
		Level 4 L	28	155	346	673	504	943	90	2,739
		Level 4H		2	13	44	88	334	98	579
		Total	1,338	1,161	1,229	1,454	876	1,536	194	7,788

1,343 became proficient



455 lost proficiency

888/7788=11% increase in proficiency

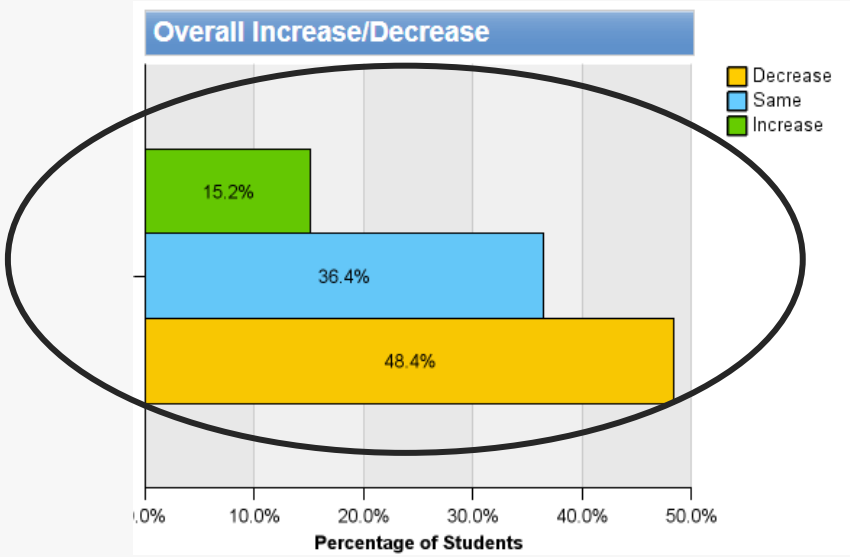
Half Level Change-This Year vs Last Year (Nassau Totals)

Count of Students			2017/2018 Grade 6 ELA						
			Level 1	Level 2 L	Level 2H	Level 3 L	Level 3H	Level 4 L	Level 4H
2018/2019	Grade 7 ELA	Level 1	649	205	120	49	38		
		Level 2 L	118	167	174	171	72	88	
		Level 2H	44	80	116	172	137	189	
		Level 3 L	11	49	104	235	235	509	
		Level 3H	2	12	35	85	100	389	
		Level 4 L	5	2	12	63	88	755	
		Level 4H				2	12	160	10
		Total	829	515	561	777	682	2,104	366

How accurate were the Parent Reports that showed students gained proficiency in grade 6?

How accurate were the Parent Reports that showed students lost proficiency in grade 7?

ecame
nt



938
lost proficiency

706/5834=12% decrease in proficiency

Related Report Links

Nassau County 2022 Proficiency

Subject	Test	# L1	# L2	# L3	# L4	Total	L3+L4	% L1	% L2	% L3	2019 Region L3+L4%	L3+L4%
ELA	Grade 3 ELA	949	3,093	5,760	1,141	10,943	6,901	9%	28%	53%	66%	63%
	Grade 4 ELA	1,325	3,179	3,439	2,776	10,719	6,215	12%	30%	32%	62%	58%
	Grade 5 ELA	1,794	3,455	2,870	2,433	10,552	5,303	17%	33%	27%	51%	50%
	Grade 6 ELA	1,378	1,554	2,562	4,513	10,007	7,075	14%	16%	26%	60%	71%
	Grade 7 ELA	1,387	2,641	3,019	2,315	9,362	5,334	15%	28%	32%	52%	57%
	Grade 8 ELA	1,079	2,232	2,447	2,327	8,085	4,774	13%	28%	30%	59%	59%

Related Report Links

Nassau County 2023 Proficiency

Subject	Test	# L1	# L2	# L3	# L4	Total	L3+L4	% L1	% L2	% L3	% L4	L3+L4%
ELA	Grade 3 ELA	1,940	2,703	4,162	2,546	11,351	6,708	17%	24%		2%	59%
	Grade 4 ELA	1,296	2,737	3,900	2,840	10,773	6,740	12%			%	63%
	Grade 5 ELA	1,505	2,799	4,127	2,282	10,713	6,409	1				60%
	Grade 6 ELA	1,529	2,400	3,457	2,664	10,050	6,121	15%			27%	61%
	Grade 7 ELA	1,650	2,073	3,304	2,360	9,387	5,664	18%		35%	25%	60%
	Grade 8 ELA	967	1,747	2,900	2,472	8,086	5,372	12%	22%	36%	31%	66%

Vertical Alignment Restored

Grade	Total	Not Tested		Tested		Level 1		Level 2		Level 3		Level 4		Proficient (Levels 3 & 4)	
	#	#	%	#	%	#	%	#	%	#	%	#	%	#	%
Grade 3	182,578	25,392	14%	157,186	86%	29,344	19%	55,062	35%	63,354	40%	9,426	6%	72,780	46%
Grade 4	185,949	27,643	15%	158,306	85%	38,732	24%	53,851	34%	39,207	25%	26,516	17%	65,723	42%
Grade 5	187,945	30,129	16%	157,816	84%	46,476	29%	51,812	33%	34,641	22%	24,887	16%	59,528	38%
Grade 6	190,059	35,588	19%	154,471	81%	37,308	24%	29,681	19%	36,928	24%	23,873	15%	87,482	57%
Grade 7	194,984	41,825	21%	153,159	79%	32,208	21%	47,510	31%	44,416	29%	28,825	19%	73,441	48%
Grade 8	198,929	52,550	26%	146,379	74%	27,873	19%	45,675	31%	42,314	29%	30,517	21%	72,831	50%

2022

Not measuring what they appear to measure

Proficiency range of 19% (yoyo effect)

Grade	Total	Not Tested		Tested		Level 1		Level 2		Level 3		Level 4		Proficient (Levels 3 & 4)	
	#	#	%	#	%	#	%	#	%	#	%	#	%	#	%
Grade 3	183,472	25,059	14%	158,413	86%	46,622	29%	40,677	26%	47,107	30%	24,007	15%	72,114	45%
Grade 4	183,783	27,658	15%	156,125	85%	34,565	22%	45,392	29%	47,049	30%	29,119	19%	76,168	49%
Grade 5	187,792	30,231	16%	157,561	84%	41,134	26%	45,390	29%	48,811	31%	22,226	14%	71,037	45%
Grade 6	188,442	33,969	18%	154,473	82%	40,022	26%	43,668	28%	42,449	27%	28,334	18%	70,783	46%
Grade 7	191,473	40,074	21%	151,399	79%	40,005	26%	38,727	26%	45,058	30%	27,609	18%	72,667	48%
Grade 8	196,313	50,676	26%	145,637	74%	25,376	17%	39,869	27%	48,470	33%	31,922	22%	80,392	55%

2023

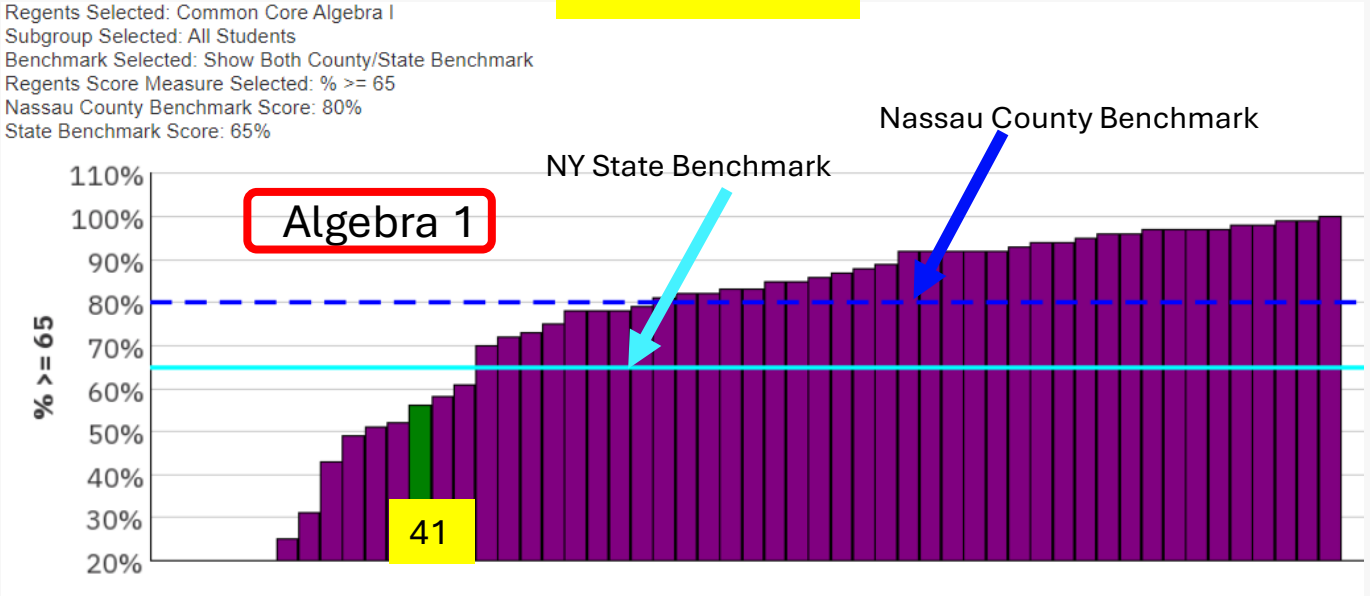
Vertical Alignment Restored

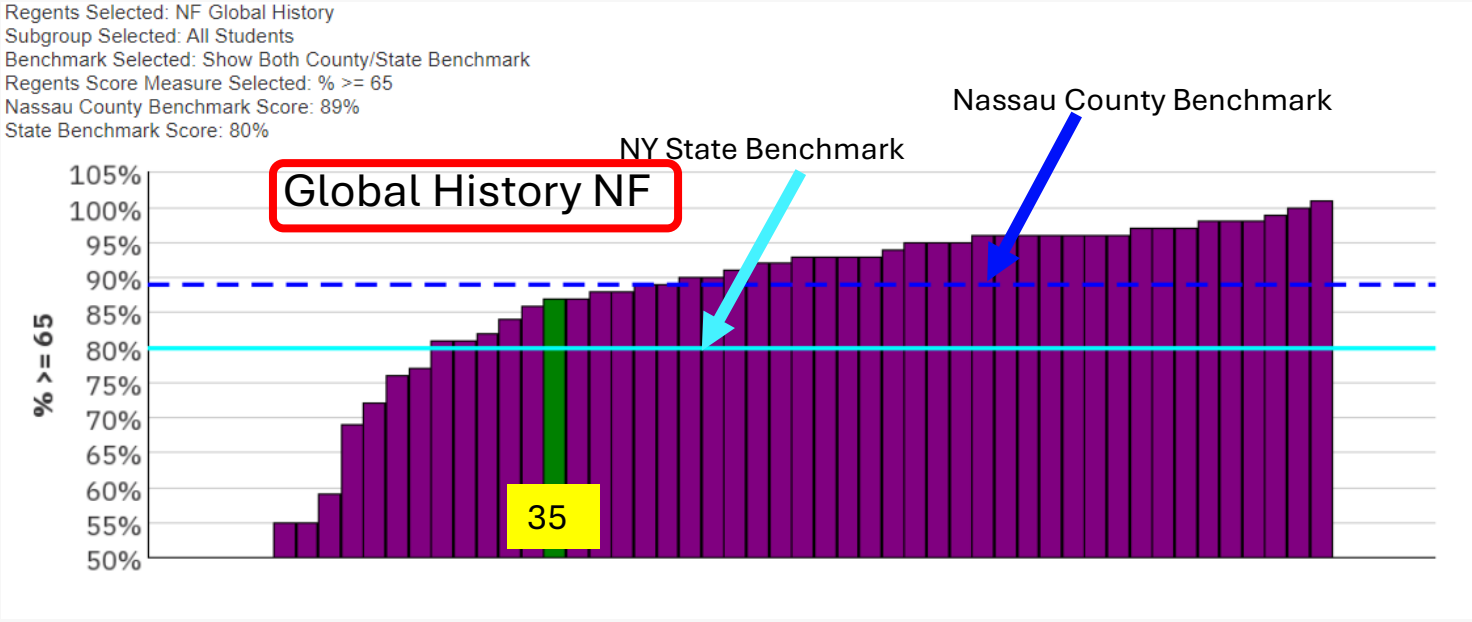
Proficiency range of 3%

Problem 3: How would Dr. John react to these charts showing results on the **Algebra 1** and **Global History NF** exams in 2022?

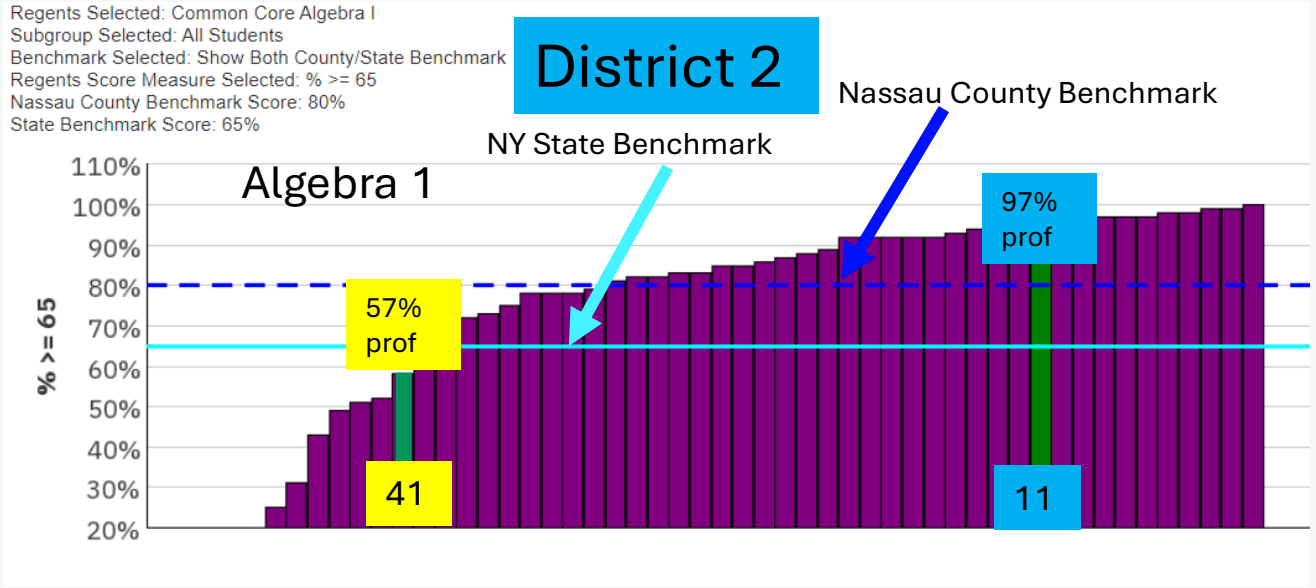
(Based on a Nassau RIC summary of School Report Card data)

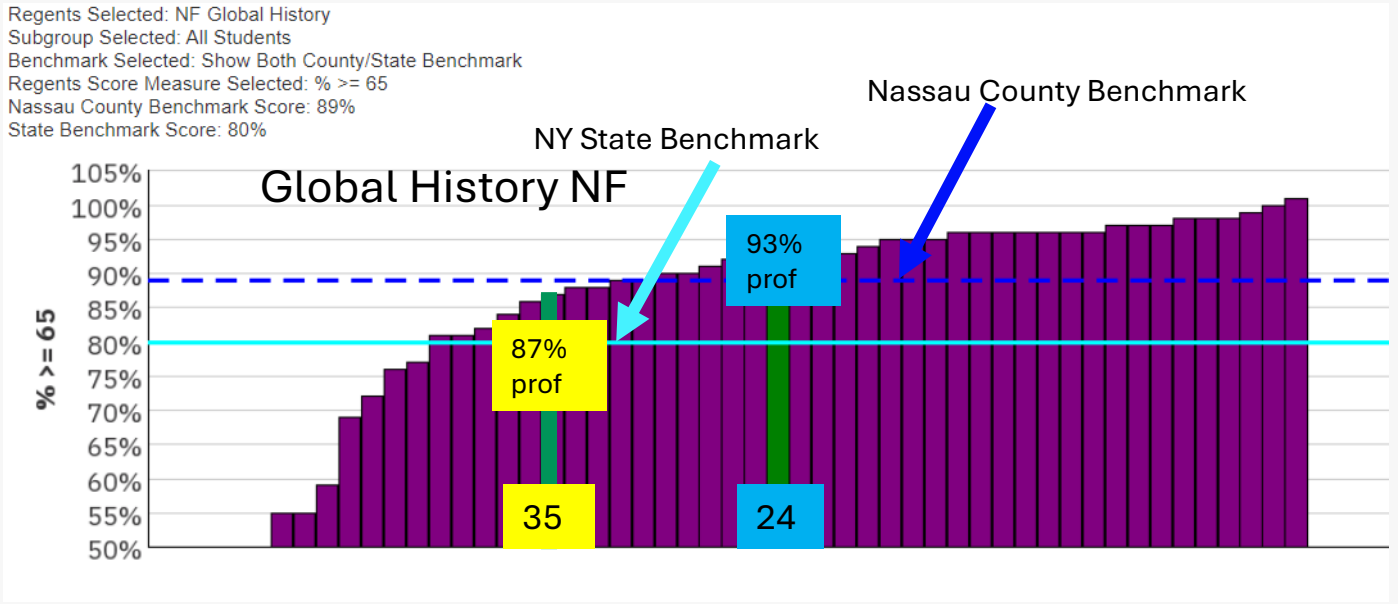
District 1





Problem 3. How would Dr. John react to these charts showing results on the Algebra 1 and Global History NF exams in 2022?





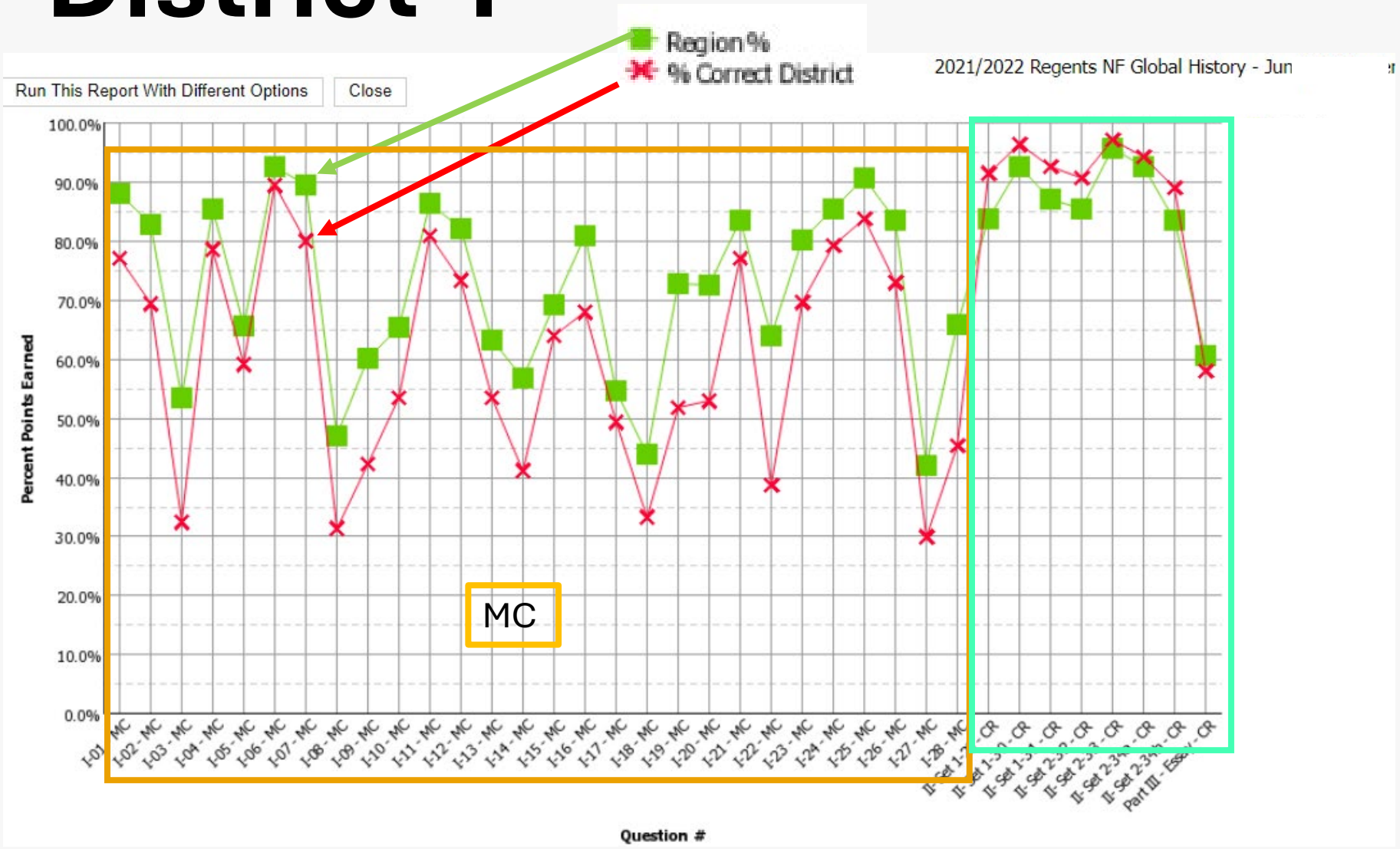
District 1 has very high percentages of Economically Disadvantaged and ELL students. District 2 has low percentages.

Based on that demographic data, how might our two experts respond?

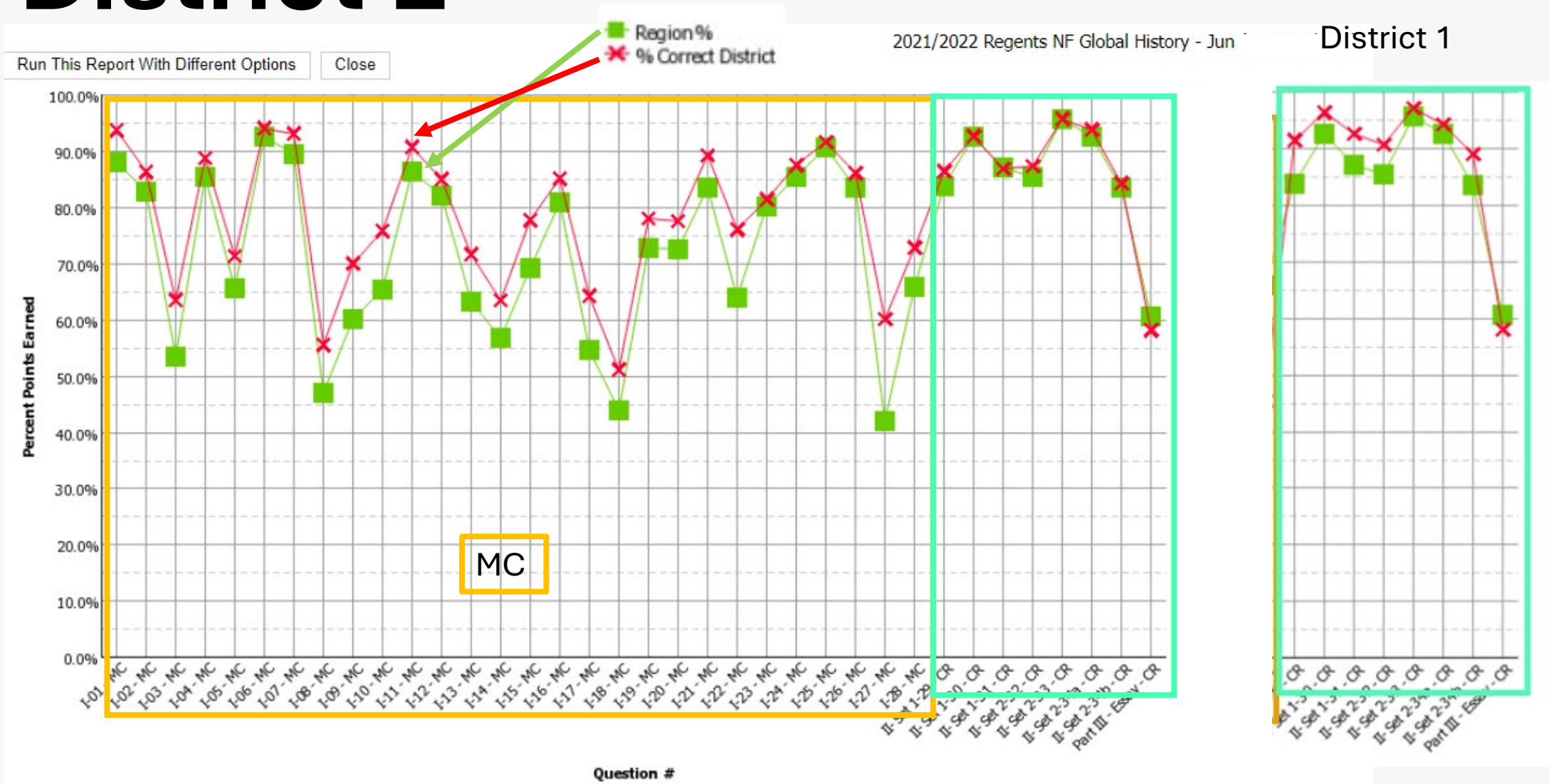
Dr. John lauds District 1 for its Global History instruction and suggests that District 2 might need to revise its Global History instructional practices.

Sam Smart defers judgment until Gap Reports for each district can be reviewed.

District 1

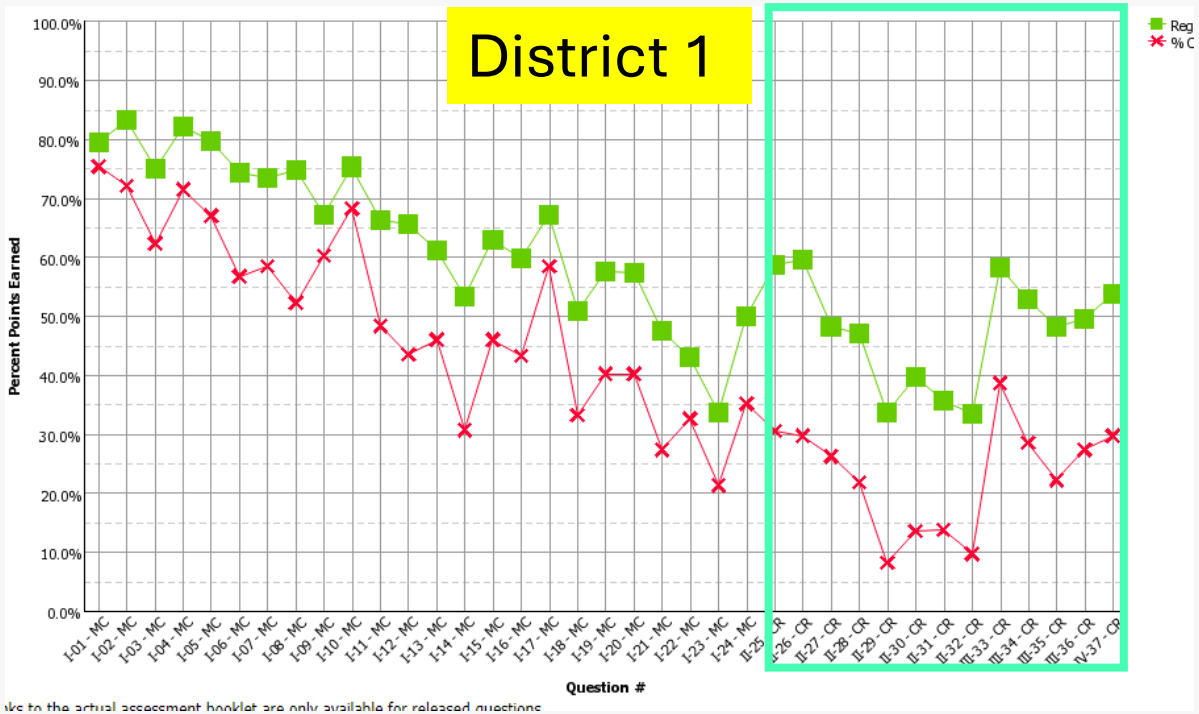

Region %
% Correct District
2021/2022 Regents NF Global History - Jun
Run This Report With Different Options
Close
100.0%
90.0%
80.0%
70.0%
60.0%
50.0%
40.0%
30.0%
20.0%
10.0%
0.0%
Percent Points Earned
MC
Question #

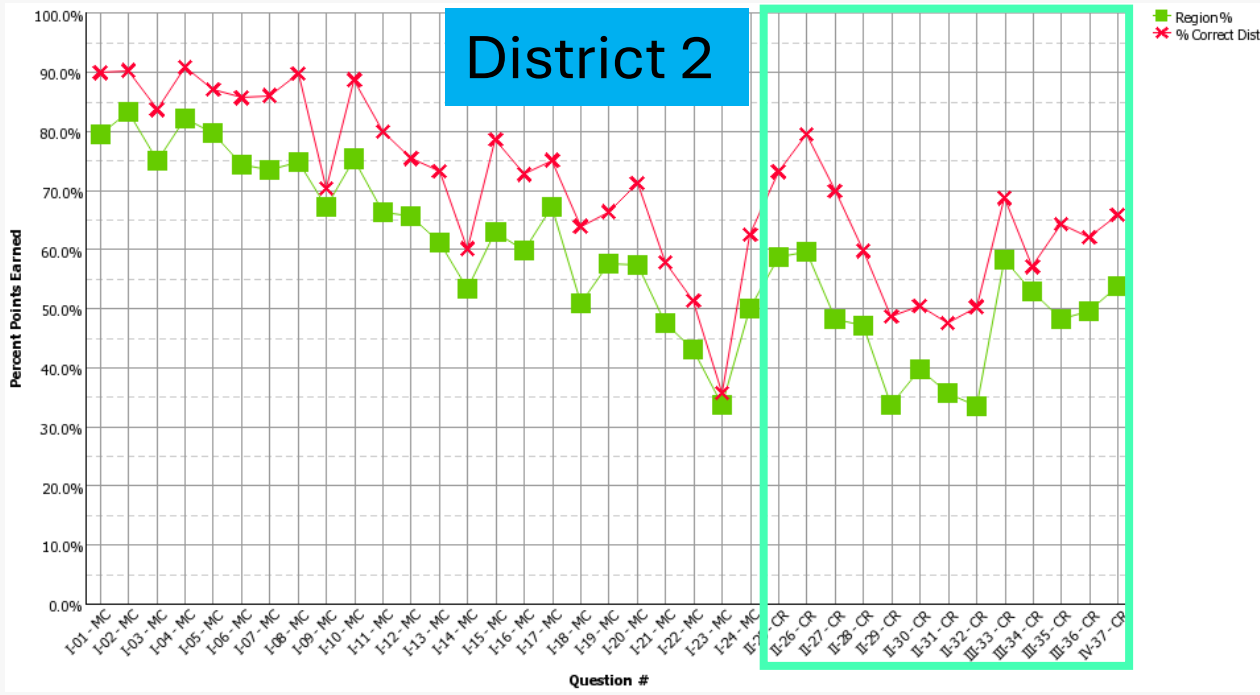
District 2



District 1

Dr. John wonders about the instructional emphases in the two districts' Global History programs. **Sam Smart still wants to see the Algebra I Gap charts.**





Dr. John suggests that District 1 might be emphasizing writing responses in **Global History** while District 2 is focusing on multiple choice questions. In **Algebra**, the differences in CR results can be explained by demographic differences between the two districts.

Sam Smart says he's seen this many times before.

Not measuring what they appear to measure

Sam Smart wonders if the Global History graders in District 1 were perhaps more generous than the graders in District 2 when interpreting the state rubric. **Sam cites the "Four R's syndrome,"** the **r**eading, w**r**iting, **r**ubric **r**eference syndrome.

To ensure consistency in the future, Sam recommends a committee from both schools blindly regrade sample essays from both schools to reaffirm their interpretations of the state grading rubrics.

Sam also wonders about NYSED's Blue Ribbon Panel's recommendation for new graduation measures--

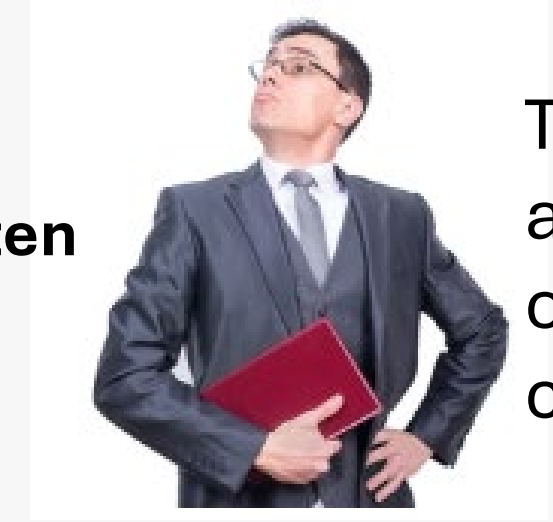
specifically, the **"state-developed rubric(s) for performance-based assessments"** that will be allowed as an option to satisfy the diploma assessment requirements.

Will the rubric be sufficiently unambiguous? Will the new measure be applied with equal rigor by all?

Problem 4. Does how we present school data always get the outcomes we say we want?

“Teach the best, forget the rest. You want a friend, buy a dog.”

A teacher in my school jokingly, but proudly, often boasted of his teaching philosophy.



To support his boast, he crowed about his extraordinary results on the Advanced Placement course he taught.

Did our School Report Card inadvertently stress **high performance** over **high participation** when it introduced AP exams on the School Report Card in 2022?

Consider how it displays results on Advanced Placement exams.

AP Course Participation

Course	Subject Area	12th Grade		11th Grade		10th Grade		9th Grade		Not HS	
		# in Course	% of Total Grade Level Subgroup	# in Course	% of Total Grade Level Subgroup	# in Course	% of Total Grade Level Subgroup	# in Course	% of Total Grade Level Subgroup	# in Course	% of Total Grade Level Subgroup
AP Computer Science A	Computer Sciences	2,503	1%	2,231	1%	628	0%	77	0%	1	0%
AP Computer Science Principles	Computer Sciences	3,965	2%	3,423	1%	3,269	1%	636	0%	7	0%
AP Eng Lang & Comp	ELA	5,007	2%	28,412	15%	1,851	0%	251	0%	63	0%
AP Eng Lit & Comp	ELA	23,525	12%	4,735	2%	576	0%	37	0%	21	0%
AP Art-History of Art	Fine and Performing Arts	610	0%	387	0%	175	0%	27	0%	0	
AP Music Theory	Fine and Performing Arts	619	0%	466	0%	155	0%	8	0%	0	

AP Assessments Proficient (3 & Higher)

Assessment	Subject Area	12th Grade		11th Grade		10th Grade		9th Grade		Not HS #	
		Total # Tested	% Total Tested with a Score of 3 or Higher	Total # Tested	% Total Tested with a Score of 3 or Higher	Total # Tested	% Total Tested with a Score of 3 or Higher	Total # Tested	% Total Tested with a Score of 3 or Higher	Total # Tested	% Total Tested with a Score of 3 or Higher
AP Computer Science A	Computer Sciences	1,684	59%	1,898	79%	491	72%	17	94%	2	–
AP Computer Science Principles	Computer Sciences	2,491	46%	2,676	51%	2,940	73%	508	58%	0	
AP Language and Comp	ELA	2,747	42%	25,001	54%	1,214	34%	103	23%	0	
AP Literature and Comp	ELA	16,057	72%	3,417	65%	373	51%	10	60%	4	–
AP Art History	Fine and Performing Arts	402	50%	296	49%	178	56%	13	30%	0	

Each listing uses alphabetical order, but they are different.

There are 42 Courses listed, but only 38 Assessments listed.

AP Assessments Proficient (3 & Higher)

Course	Assessment	Subject Area	12th Grade		11th Grade		10th Grade		
			Total # Tested	% Total Tested with a Score of 3 or Higher	Total # Tested	% Total Tested with a Score of 3 or Higher	Total # Tested	% Total Tested with a Score of 3 or Higher	Tot Tes
AP Computer Science A	AP Computer Science A	Computer Sciences	1,684	59%	1,898	79%	491	72%	17
AP Computer Science Principles	AP Computer Science Principles	Computer Sciences	2,491	46%	2,676	51%	2,940	73%	508
AP Eng Lang & Comp	AP Language and Comp	ELA	2,747	42%	25,001	54%	1,214	34%	103
AP Eng Lit & Comp	AP Literature and Comp	ELA	16,057	72%	3,417	65%	373	51%	10
AP Art-History of Art	AP Art History	Fine and Performing Arts	402	50%	296	49%	178	56%	13
AP Music Theory	AP Music Theory	Fine and Performing Arts	429	62%	327	60%	108	65%	7
AP Studio Art- 2D	AP Studio Art 2d Design	Fine and Performing Arts	1,522	84%	413	83%	48	75%	1
AP Studio Art- 3D	AP Studio Art 3d Design	Fine and Performing Arts	109	77%	39	76%	7	85%	0
AP Studio Art-Draw Portfolio	AP Studio Art Drawing	Fine and Performing Arts	914	89%	234	74%	32	65%	2
AP Calculus AB	AP World History	Global Studies	194	57%	481	48%	24,081	69%	538
	AP Calculus AB	Mathematics	12,412	53%	1,709	52%	48	79%	6

Let's compare the **AP Course** data with the **AP Assessment** data for World History.

Course	Subject Area	12th Grade		11th Grade		10th Grade		9th			
		# in Course	% of Total Grade Level Subgroup	# in Course	% of Total Grade Level Subgroup	# in Course	% of Total Grade Level Subgroup	# in Course			
AP World History	Social Studies	84	0%	268	0%	13,973	7%	3,065	1%		
AP World History: Modern	Social Studies	254	0%	558	0%	12,330	6%	1,694	0%	9	0%
		338		826		26,303		4,759			

Assessment	Subject Area	12th Grade		11th Grade		10th Grade		9th Grade		Not HS #	
		Total # Tested	% Total Tested with a Score of 3 or Higher	Total # Tested	% Total Tested with a Score of 3 or Higher	Total # Tested	% Total Tested with a Score of 3 or Higher	Total # Tested	% Total Tested with a Score of 3 or Higher	Total # Tested	% Total Tested with a Score of 3 or Higher
AP World History	Global Studies	194	57%	481	48%	24,081	69%	538	61%	5	80%

Not measuring what they appear to measure

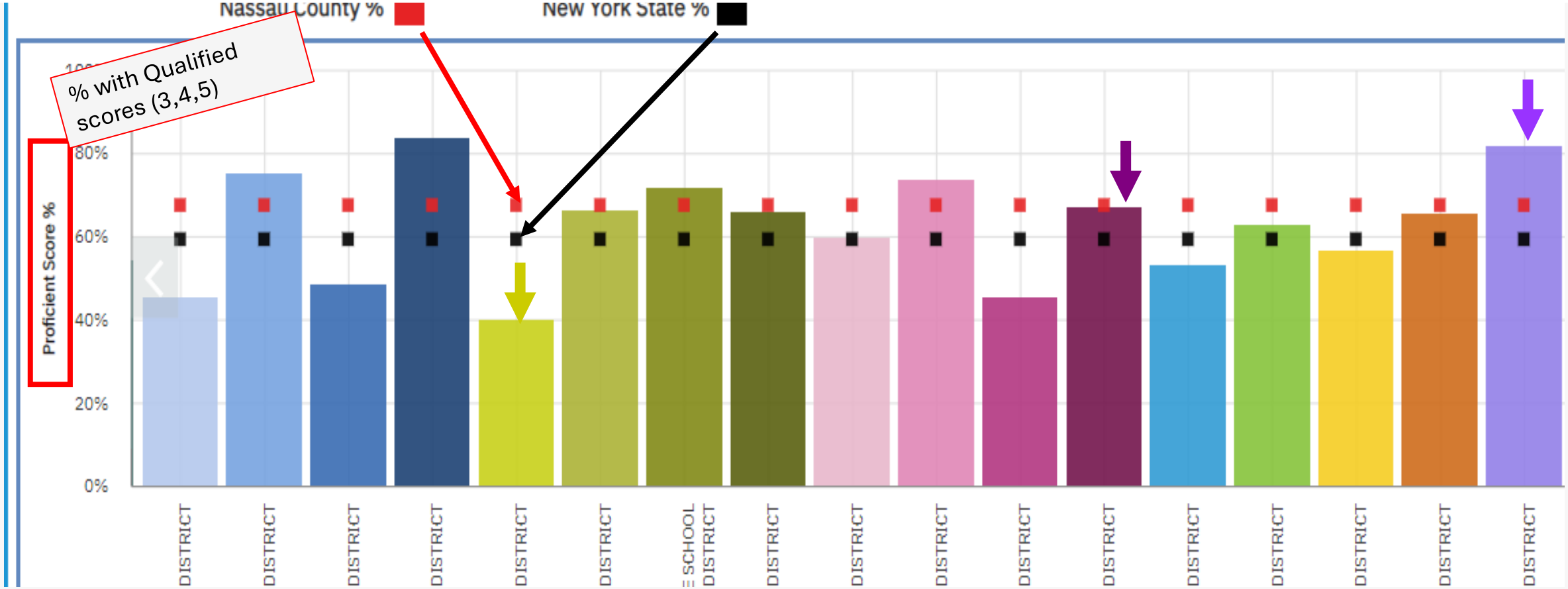
Great First Effort

Comparison of Advanced Placement Scores Across Districts i-Report

Purpose:

This report displays the percent of proficient scores on the Advanced Placement (AP) data for each Nassau County public school district. Proficient scores are defined as AP scores that ar equal to 3 or higher. Note that the data reflect the number of assessments, which is not necessarily the same as the number of students. Also, the data are based on the ALTREG and CCR assessment data submitted to NYSED and may not match other IDW reports based on Nassau BOCES IDW AP scanning project. Data are sourced from the NYSED State Report Card database.

Please use caution when comparing districts on this report. The numbers and percents reflect students who took AP assessments and do not take into account the proportion of a district's enrollment. Therefore, a district that only has their top students taking an AP assessment will have inflated proficiency numbers compared with a district that has a larger percentage of their enrollment taking the same test.

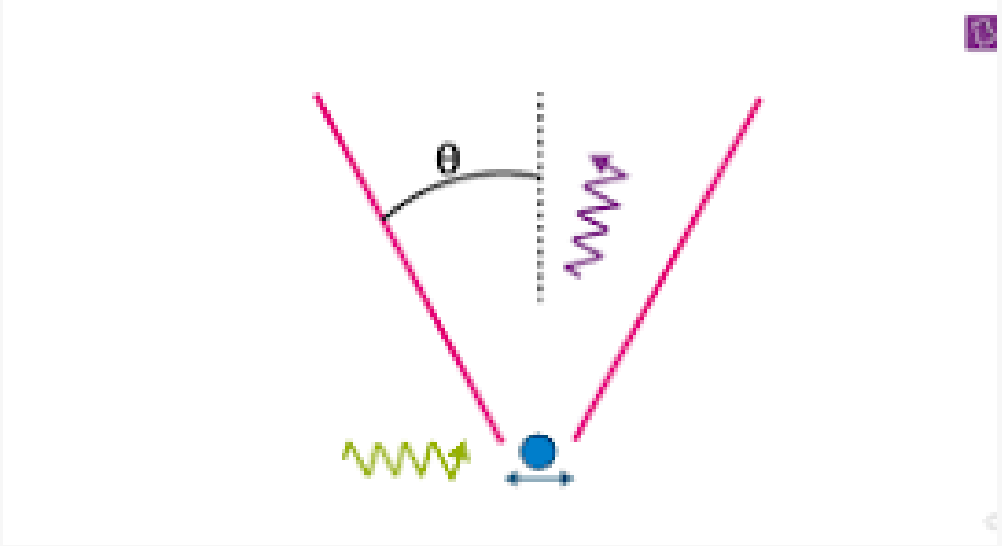


[redacted]

		2017		2018		2019		2020		2021		2022		2023	
		School#	Region	School#	Region	School#	Region	School#	Region	School#	Region	School#	Region	School#	Region
Score	1	***87***	68.4	***93***	66.6	***93***	67.6	***67***	60.5	***196***	98.4	***129***	92.7	***95***	78.0
	2	***266***	171.3	***266***	177.6	***290***	180.3	***205***	146.1	***350***	172.9	***347***	162.4	***267***	155.6
	3	***419***	232.7	***446***	243.5	***475***	246.1	***422***	246.9	***472***	216.4	***447***	219.5	***535***	241.9
	4	***428***	180.9	***443***	198.9	***463***	193.6	***402***	216.6	***373***	174.2	***401***	188.4	***546***	210.0
	5	***315***	110.5	***373***	135.8	***303***	131.5	***312***	150.3	***260***	105.9	***339***	128.8	***422***	145.5
# Schools Surveyed		54		53		55		55		60		60		60	
Total Tests Taken*		***1515***		***1621***		***1624***		***1408***		***1651***		***1663***		***1865***	
Participation Index (Enrollment Tested)**		***3.81***	Tests taken per average grade enrollment (400 students per grade).												
Total Scoring 3,4,5		***1,162***	524.1	***1,262***	578.2	***1,241***	571.2	***1,136***	613.8	***1,105***	496.5	***1,187***	536.6	***1,503***	597.4
Quality Index (% Test Takers 3,4,5)		***76.7%***	Percent proficient (3,4,5).												
Quality / Participation (Enrollment 3,4,5)**		***2.92***	Students earning grades of 3,4,5 per average grade enrollment.												

In analyzing school data, two principles must always be considered.

Heisenberg's Uncertainty Principle



Werner Heisenberg's uncertainty principle states that we cannot know both the position and speed of a particle, such as a photon or electron, with perfect accuracy; the more we nail down the particle's position, the less we know about its speed and vice versa.

The act of measuring it actually changes it.

Like trying to feel the shape of a snowflake.

Is there a similar principle affecting school measurement?

Campbell's Law

Donald Campbell

The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor.

falsify

contaminate

Consider teacher observations by supervisory personnel.

Can we all remember how testing was turned upside down about a dozen years ago?

SED used the following study to justify using standardized tests to measure teacher performance.

48 pages long + 30 pages of tables

MEASURING THE IMPACTS OF TEACHERS II: TEACHER VALUE-ADDED AND STUDENT OUTCOMES IN ADULTHOOD

Raj Chetty John N. Friedman Jonah E. Rockoff

http://www.nber.org/papers/w19424

(Abstract)

Are teachers' impacts on students' test scores ("value-added") a good measure of their quality? This question has sparked debate partly because of a lack of evidence on whether high value-added (VA) teachers who raise students' test scores improve students' long-term outcomes. Using school district and tax records for **more than one million children**, we find that students assigned to high-VA teachers in primary school are more likely **to attend college, earn higher salaries, and are less likely to have children as teenagers.** Replacing a teacher whose VA is in the bottom 5% with an average teacher would increase the present value of students' lifetime income by approximately **$250,000 per classroom.**

But footnote number 45 revealed how limited the conclusion was.

[45] As we noted above, even in the low-stakes regime we study, some unusually high VA teachers have test score impacts consistent with test manipulation. If such behavior becomes more prevalent when VA is used to evaluate teachers, the predictive content of VA as a measure of true teacher quality could be compromised.

Campbell's Law in a nutshell!!!

Here are two troubling manifestations of these “corruption pressures.”

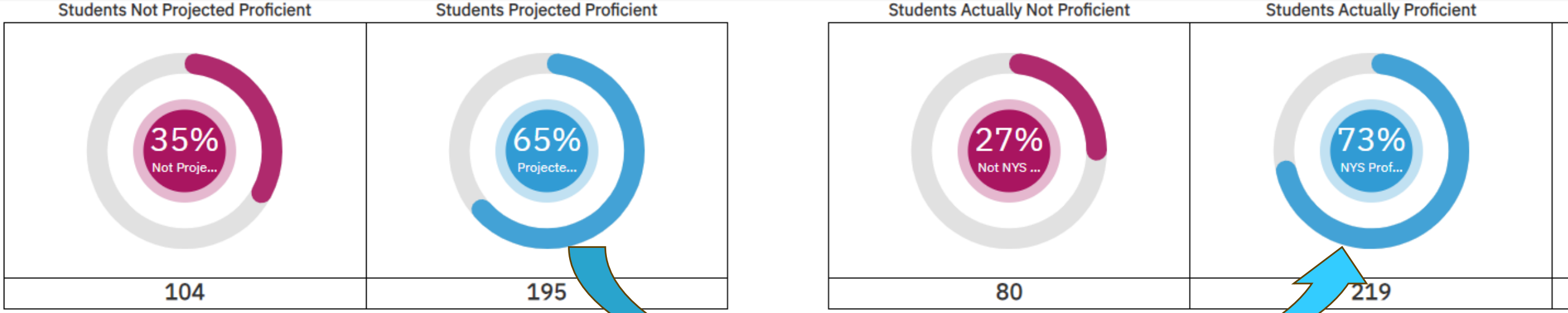
Manifestation 1. Based partly on high opt-out (test-refusal) rates, districts now offer third-party tests like the NWEA to project likely scores on the Grade 3-8 State Assessments and also give immediate and more frequent feedback.

We have found these projections to be reasonably accurate.

Typical District's Sixth Grade Scores

Projected Proficiency

Actually Proficient



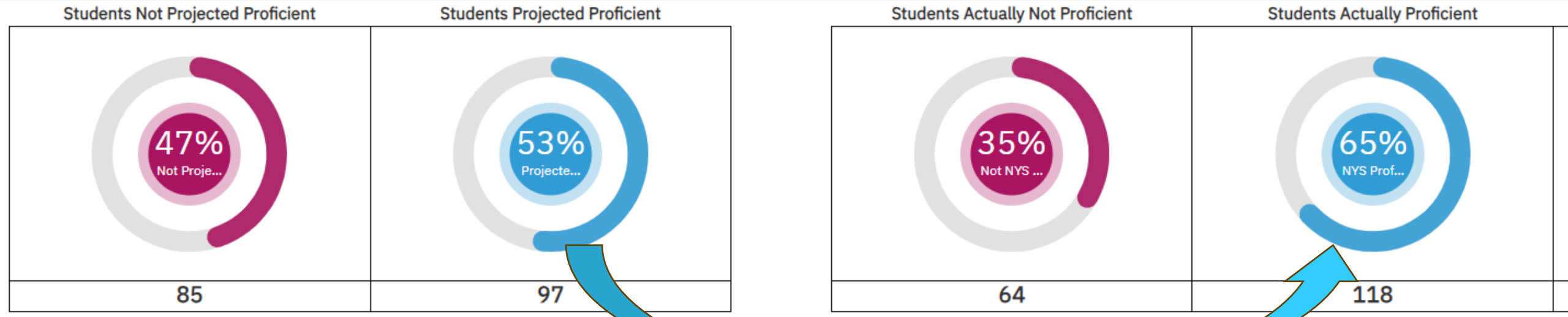
8% difference in proficiency

23% Refusals

Typical District's Sixth Grade Scores

Projected Proficiency

Actually Proficient



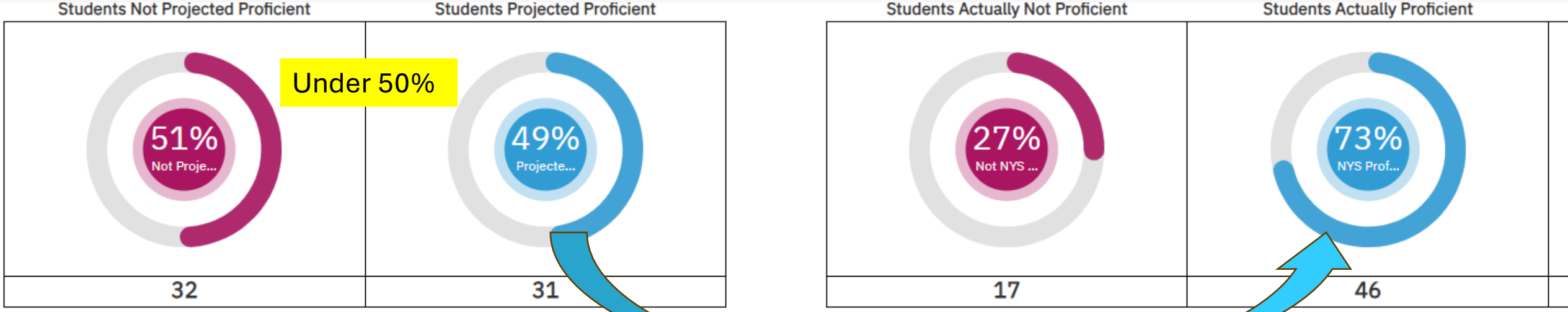
12% difference in proficiency

35% Refusals

Troublesome District's Sixth Grade Scores

Projected Proficiency

Actually Proficient



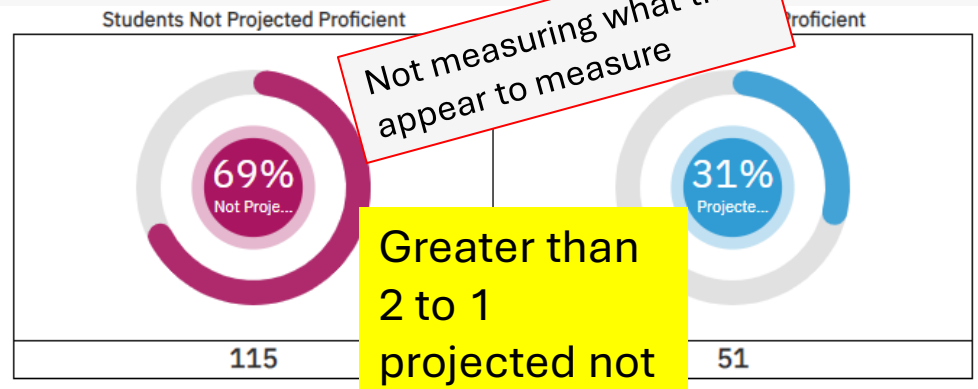
Under 50%

24% difference in proficiency

65% Refusals

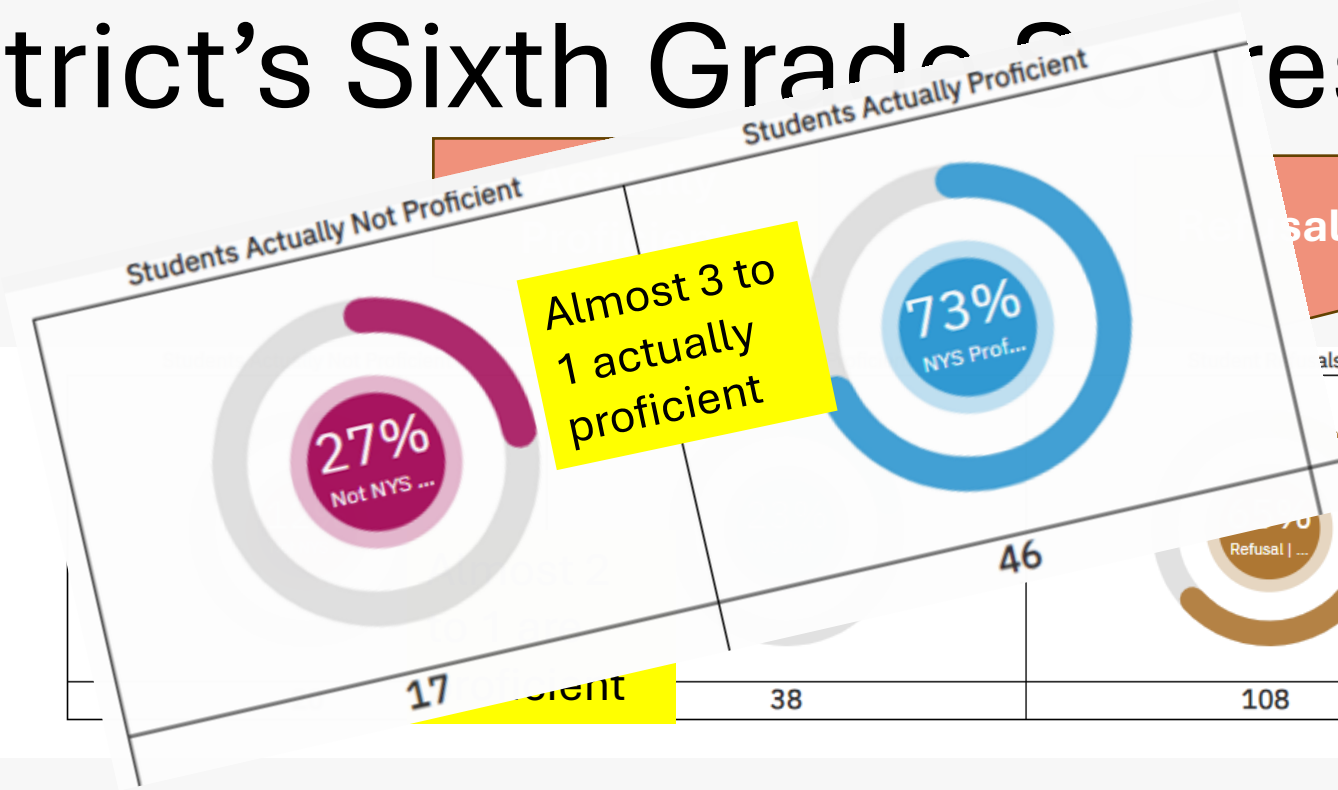
Troublesome District's Sixth Grade S…es

Projected Proficiency



Not measuring what they appear to measure

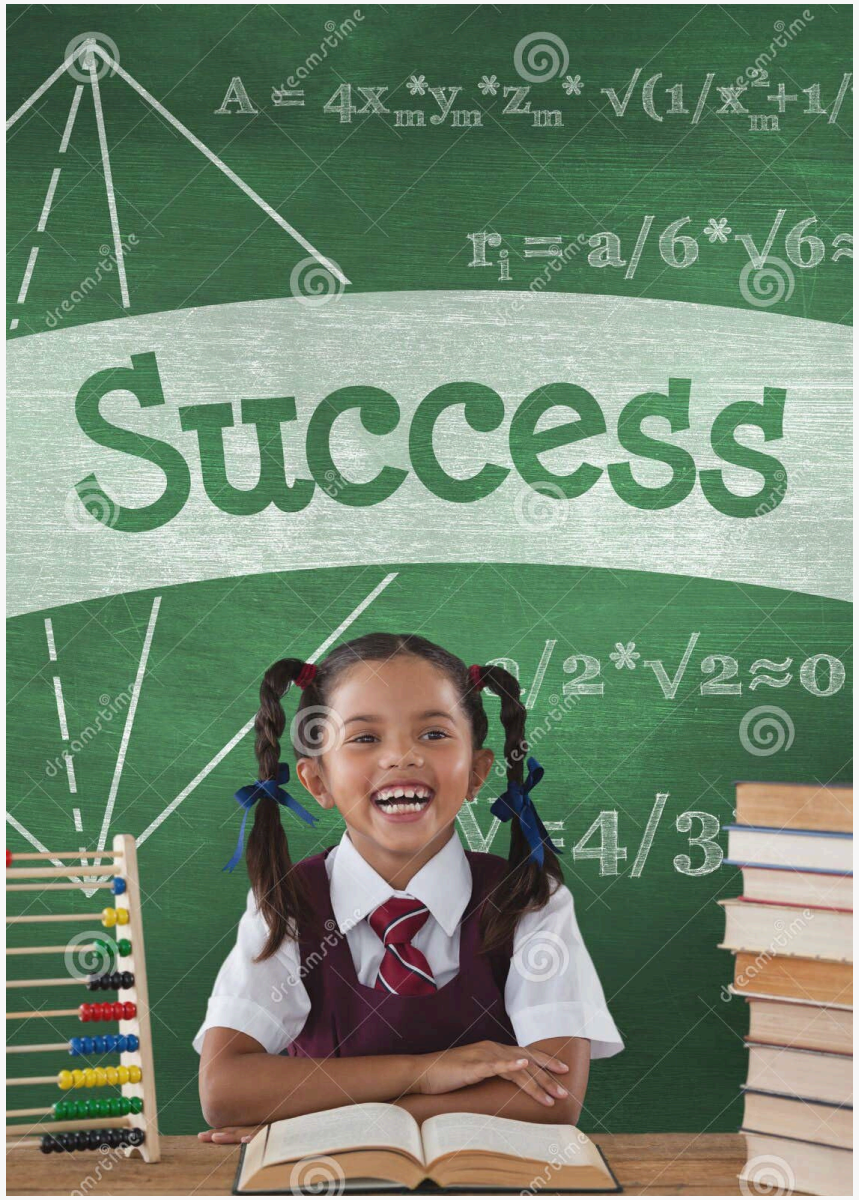
Greater than 2 to 1 projected not proficient



Almost 3 to 1 actually proficient

School officials explained that the "spirit of refusing" among parents and students caused many students to completely discount third-party computer-based tests.

Some students who "blew-off" the computer-based tests (saw them as "corrupt" or inconsequential) took the paper and pencil tests more seriously. The third-party tests "melted the snowflake."

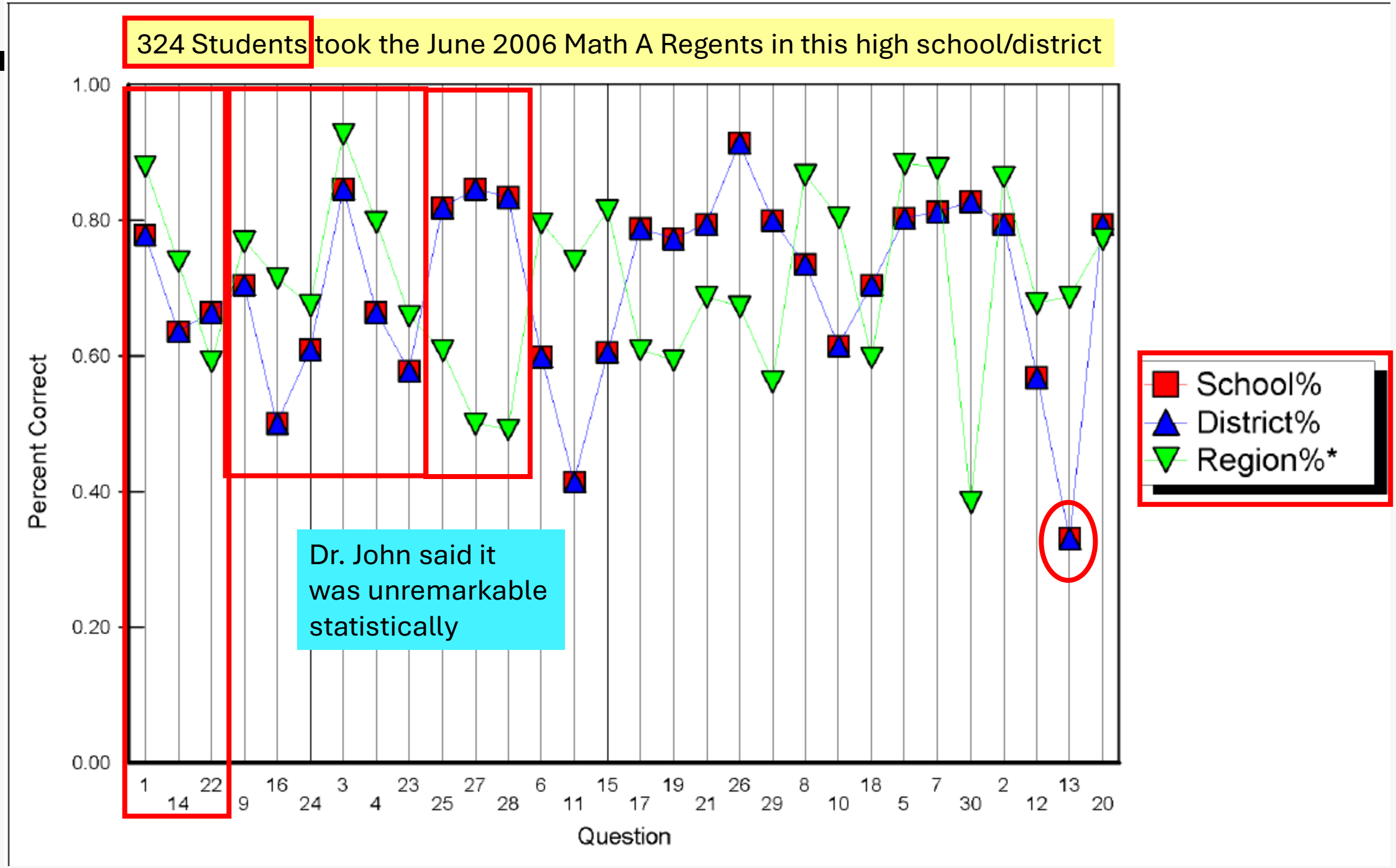


Manifestation 2. Even before APPR and Value Added, there was still a focus on test scores. Publishing scores was sufficient to cause issues.

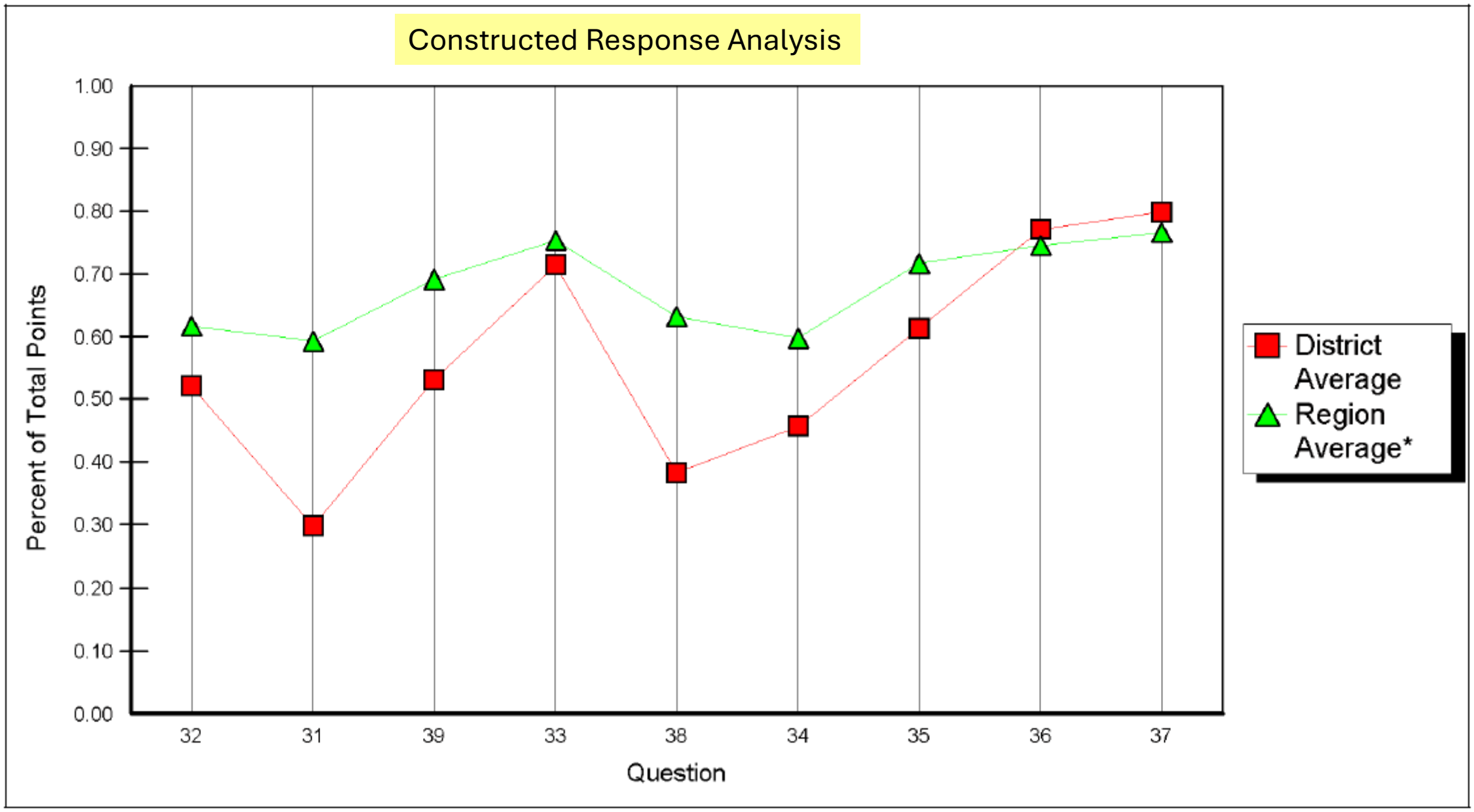
The Nassau RIC created insightful reports to help districts improve instruction.

Analyzing Math A Regents scores, I saw this high school's Gap Report for the June 2006 Regents.

Key Idea/Standard	Question	School%	District%	Region%*
1 Mathematical Reasoning	1	77.8%	77.8%	87.9%
	14	63.6%	63.6%	73.9%
	22	66.4%	66.4%	59.1%
2 Number and Numeration	9	70.4%	70.4%	76.8%
	16	50.0%	50.0%	71.4%
	24	60.8%	60.8%	67.4%
3 Operations	3	84.6%	84.6%	92.6%
	4	66.4%	66.4%	79.6%
	23	57.7%	57.7%	65.8%
	25	81.8%	81.8%	60.7%
	27	84.6%	84.6%	50.0%
	28	83.3%	83.3%	49.1%
4 Modeling/Multiple Repre	6	59.9%	59.9%	79.5%
	11	41.4%	41.4%	74.0%
	15	60.5%	60.5%	81.3%
	17	78.7%	78.7%	60.8%
	19	77.2%	77.2%	59.2%
	21	79.3%	79.3%	68.6%
	26	91.4%	91.4%	67.2%
	29	79.9%	79.9%	56.2%
5 Measurement	8	73.5%	73.5%	86.6%
	10	61.4%	61.4%	80.4%
	18	70.4%	70.4%	59.7%
6 Uncertainty	5	80.2%	80.2%	88.2%
	7	81.2%	81.2%	87.7%
	30	82.7%	82.7%	38.4%
7 Patterns/Functions	2	79.3%	79.3%	86.3%
	12	56.8%	56.8%	67.7%
	13	33.0%	33.0%	68.6%
	20	79.3%	79.3%	77.1%


324 Students took the June 2006 Math A Regents in this high school/district
1.00
0.80
0.60
0.40
0.20
0.00
Percent Correct
1 14 22 9 16 24 3 4 23 25 27 28 6 11 15 17 19 21 26 29 8 10 18 5 7 30 2 12 13 20
Question
School%
District%
Region%*
Dr. John said it was unremarkable statistically

Constructed Response Analysis



Language
- Chinese
- Haitian Creole
- Korean
- Russian
- Spanish

The University of the State of New York
REGENTS HIGH SCHOOL EXAMINATION

MATHEMATICS A

Thursday, June 15, 2006 1:15 p.m.

Answer Sheet

{Student ID}

{dist}

{student ID}

Student Name: {Name}

School: {School Name}

Teacher: {Teacher Name} Course/ Section: {student course-section}

Your answers to Part I should be recorded on this answer sheet.

Part I

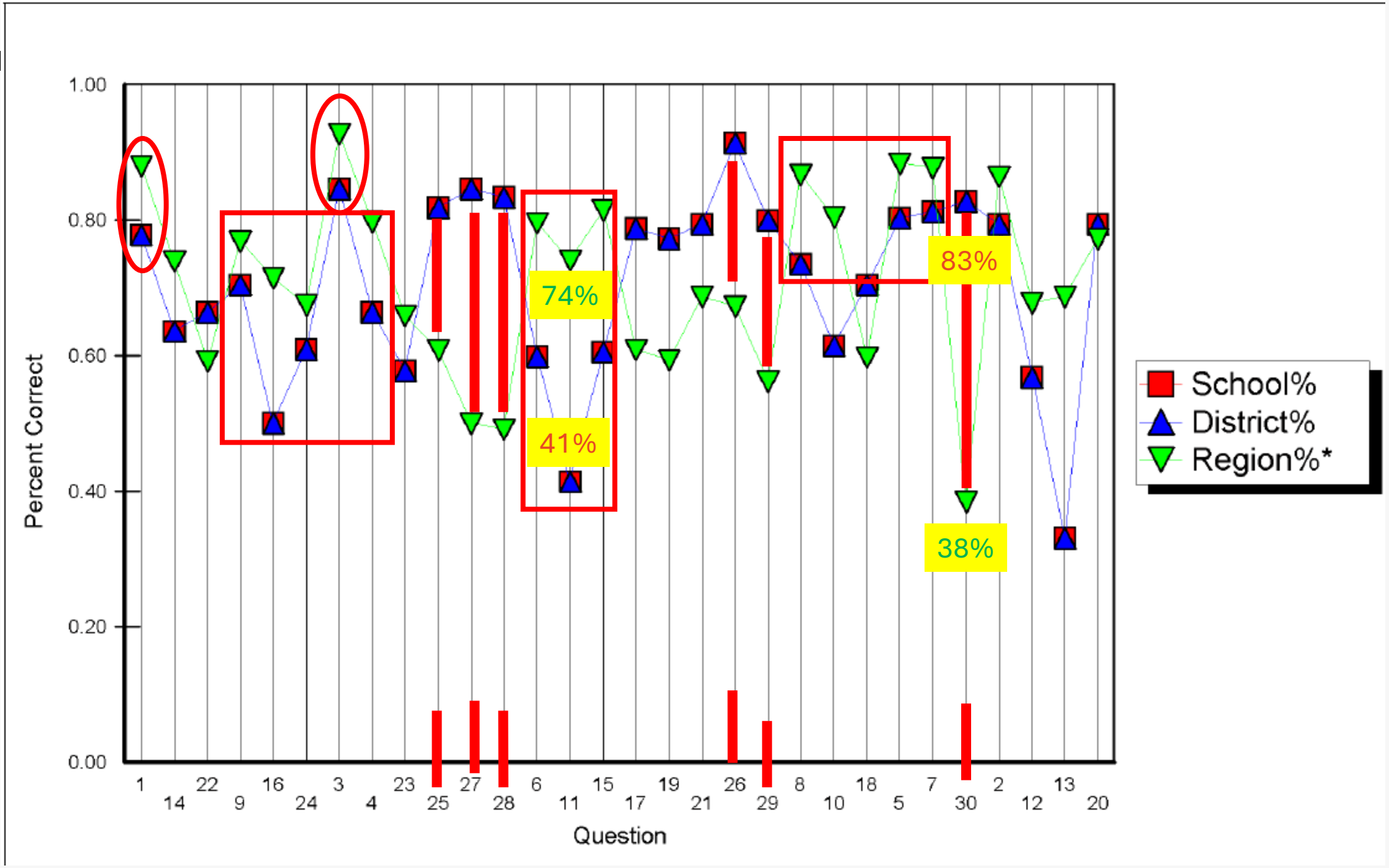
Answer all 30 questions in this part

Use #2 Pencil

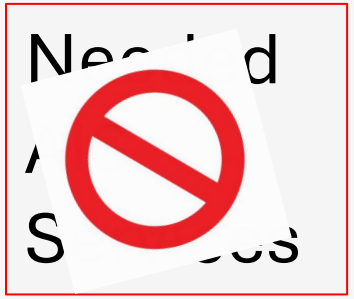
Shade the circles completely.

1.	①	②	③	④	9.	①	②	③	④	17.	①	②	③	④
2.	①	②	③	④	10.	①	②	③	④	18.	①	②	③	④
3.	①	②	③	④	11.	①	②	③	④	19.	①	②	③	④
4.	①	②	③	④	12.	①	②	③	④	20.	①	②	③	④
5.	①	②	③	④	13.	①	②	③	④	21.	①	②	③	④
6.	①	②	③	④	14.	①	②	③	④	22.	①	②	③	④
7.	①	②	③	④	15.	①	②	③	④	23.	①	②	③	④
8.	①	②	③	④	16.	①	②	③	④	24.	①	②	③	④

25.	①	②	③	④
26.	①	②	③	④
27.	①	②	③	④
28.	①	②	③	④
29.	①	②	③	④
30.	①	②	③	④


1.00
0.80
0.60
0.40
0.20
0.00
Percent Correct
74%
41%
83%
38%
1 14 22 9 16 24 3 4 23 25 27 28 6 11 15 17 19 21 26 29 8 10 18 5 7 30 2 12 13 20
Question
School%
District%
Region%*

Sam Smart overruled Dr. John, and BOCES turned the case over to NYSED. When erasure analysis confirmed Sam's conjecture, NYSED took further action.



Not measuring what they appear to measure

Campbell's Law

Donald Campbell

The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor.

Let’s remember…

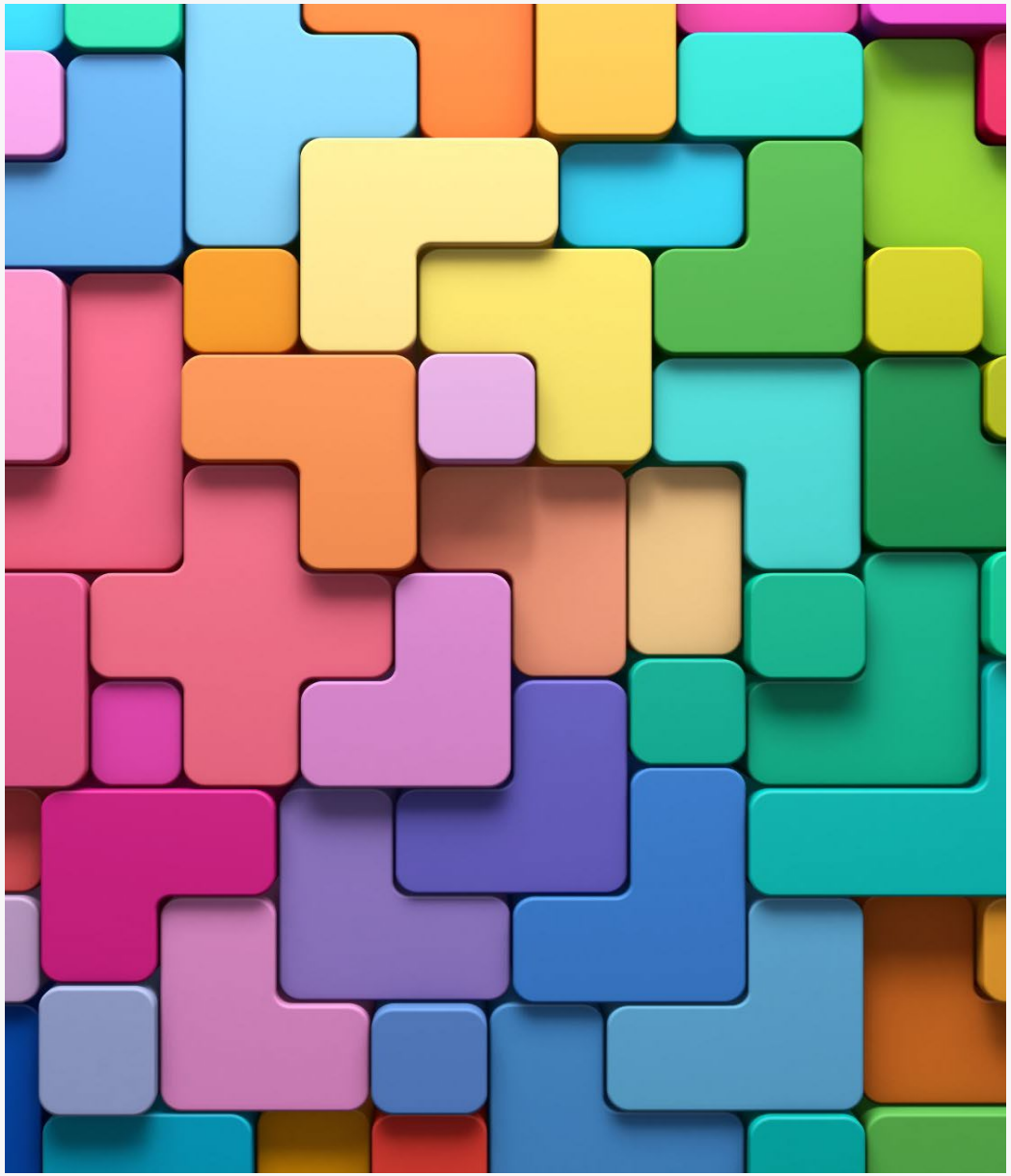
- In economics, Dr. John’s scientific and mathematical expertise must be balanced by Fat Tony’s real-life experiences.

- And when looking at school data, Dr John’s mathematical and scientific expertise must be balanced by Sam Smart’s school experience.

Completing the Data Puzzle

What happens when the pieces don't quite fit? Do our statistical models always work?

The experienced educator must always be an active partner in any data analysis!



Fred Cohen
Nassau BOCES
fcohen@nasboces.org