Center for Assessment

ISBE TAC

Achievement Indicator Alternatives: Extended Exploration

André A. Rupp

In-person TAC Meeting, January 22-23, 2025
ISBE Offices Chicago, IL

Slide deck version: 1/15/25; based on final version from 5/29/24 for June 2024 meeting

Overview

- Explore the impact of alternative computation approaches for the achievement indicators for ELA and Math
- Alternative approaches should be designed to remove undesirable effects of current step targets
- Analyses are done with student-level data from the 2022/2023 school year
- Analyses are illustrative as certain computational rules, levels of analysis, or other indicators were excluded
- Extended analyses include investigations of the effect of alternative computation rules on composite index
- All analyses were done in SPSS with syntax saved to allow for replications

TAC Questions

1. Which analytic approaches or design extensions would you recommend beyond what is presented here?
2. What are some potential pitfalls that need to be watched out for? Are there risks to replicate problematic, well-known scientific relationships or trends in this space?
3. What are other productive framings for this work that could yield empirical insight? If it were more valuable to pivot slightly, what would that look like?
4. What kinds of qualitative information could ISBE provide to help make judgments about acceptable use of alternative computation approaches?

Context

Assessments for ELA and Math

ES

Current elementary and middle school ELA and math assessment is **Illinois Assessment of Readiness**
IAR is administered in **grades 3-8**
Reporting is done with **five proficiency levels (1-5)**
A student is considered **"proficient"** if they are classified into **levels 4 or 5**

HS

Current high school ELA and math assessment is **SAT**
SAT is administered once in high school, typically in **grades 9-12** [grade 11 is most common]
Reporting is done with **four proficiency levels**
A student is considered **"proficient"** if they are classified into **levels 3 or 4**

ES + HS

Students with significant cognitive disabilities can take the **DLM** as an **alternate assessment**
DLM is administered in **grades 3-8** as well as **grades 9-12** [grades 9-11 are most common]
Reporting is done with **four proficiency levels**
A student is considered **"proficient"** if they are classified into **levels 3 or 4**

Proficiency Computations

	ES (Grades 3-8)	HS (Grades 9-12)
Students without Disabilities	IAR 5 Levels Proficient ≥ 4	SAT 4 Levels Proficient ≥ 3
Students with Disabilities	DLM 4 Levels Proficient ≥ 3	DLM 4 Levels Proficient ≥ 3

Scoring Schema

	DLM & SAT Performance Levels			
	1	**2**	**3**	**4**
Schema 1	0	0.5	1	1.5

	IAR Performance Levels				
	1	**2**	**3**	**4**	**5**
Schema 1	0	0.33	0.67	1	1.33
Schema 2	0	0.25	0.50	1	1.5

For the two IAR schema in the right table:

- Schema 1 gives equal value to each performance level increase
- Schema 1 rewards levels 2 and 3 more strongly than Schema 2
- Schema 2 rewards level 5 more strongly than Schema 1
- Most students (84%) in a given year are scored by the IAR schema as most students are regular ES students

Annual Proficiency Targets

Grade	ELA	Math	Science
3	3&4	3&4	N/A
4			
5	5&6	5&6	5
6			N/A
7	7&8	7&8	
8			8
9	HS	HS	N/A
10			
11			11
12			N/A

Annual Math Proficiency Targets 5th & 6th Grade

5th & 6th Grade Math	All	White	Black or African American	Hispanic or Latino	Asian	Two or More Races	American Indian or Alaska Native	Native Hawaiian or Other Pacific Islander	Low Income	Child With a Disability (CWD)	English Learner	Former English Learner
2022	17.18	26.16	2.63	6.31	41.23	18.17	11.34	17.48	5.41	7.05	1.57	18.53
2023	23.80	31.96	10.57	13.92	45.67	24.70	18.49	24.07	13.10	14.59	9.61	25.02
2024	30.42	37.76	18.52	21.53	50.10	31.23	25.64	30.67	20.79	22.13	17.65	31.52
2025	37.04	43.57	26.46	29.14	54.53	37.76	32.79	37.26	28.48	29.68	25.69	38.02
2026	43.66	49.37	34.40	36.74	58.97	44.29	39.94	43.85	36.17	37.22	33.73	44.52
2027	50.28	55.18	42.34	44.35	63.40	50.82	47.09	50.44	43.86	44.76	41.77	51.01
2028	56.90	60.98	50.29	51.96	67.83	57.35	54.24	57.04	51.55	52.30	49.81	57.51
2029	63.52	66.78	58.23	59.57	72.27	63.88	61.40	63.63	59.24	59.84	57.84	64.01
2030	70.14	72.59	66.17	67.18	76.70	70.41	68.55	70.22	66.93	67.38	65.88	70.51
2031	76.76	78.39	74.11	74.78	81.13	76.94	75.70	76.81	74.62	74.92	73.92	77.00
2032	83.38	84.20	82.06	82.39	85.57	83.47	82.85	83.41	82.31	82.46	81.96	83.50
2033	90.00	90.00	90.00	90.00	90.00	90.00	90.00	90.00	90.00	90.00	90.00	90.00

ELA & Math: https://www.isbe.net/Documents/2022-ELA-Math-Targets-All.pdf
Science: https://www.isbe.net/Documents/Science-Proficiency-Targets.pdf

Data Structure

Population Size

930,053 students from 3,772 schools (derived through aggregation via RCDTS x ES/HS x SchoolName)

Student-level Information

Grade

Race

EL indicator (yes, no)

Former EL indicator (yes, no)

Disability status indicator (yes, no)

Low-income indicator (yes, no)

ELA scale score and proficiency level

Math scale score and proficiency level

Data were merged with school-level indicator data received in October 2023

190 schools in new file but not old while 138 schools in old file but not new

3,582 schools remaining as their information was in both old and new files

Analyses: Part 1, January 2025

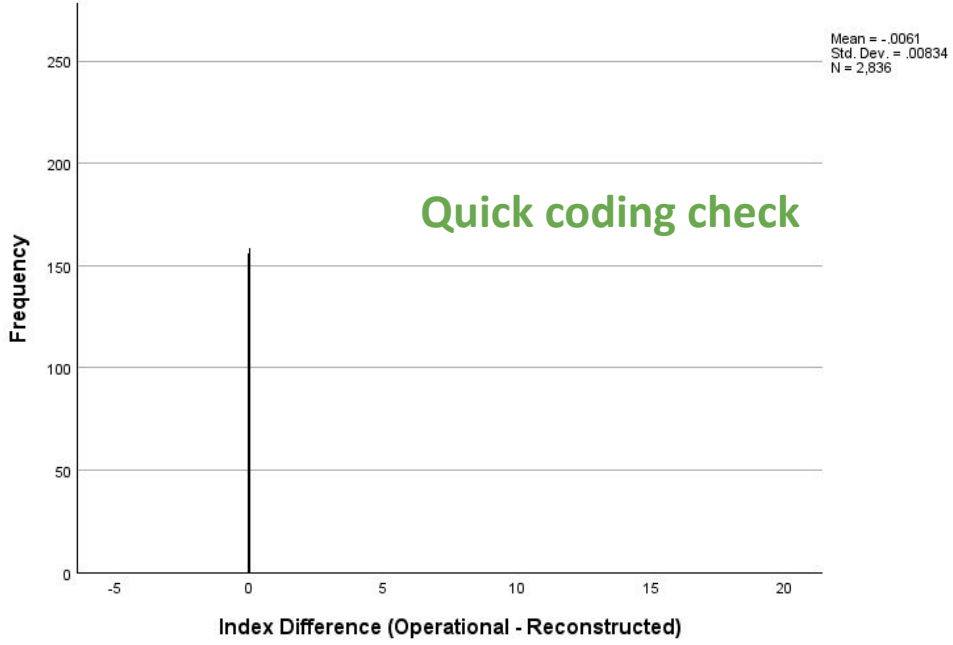
Composite Index Effects

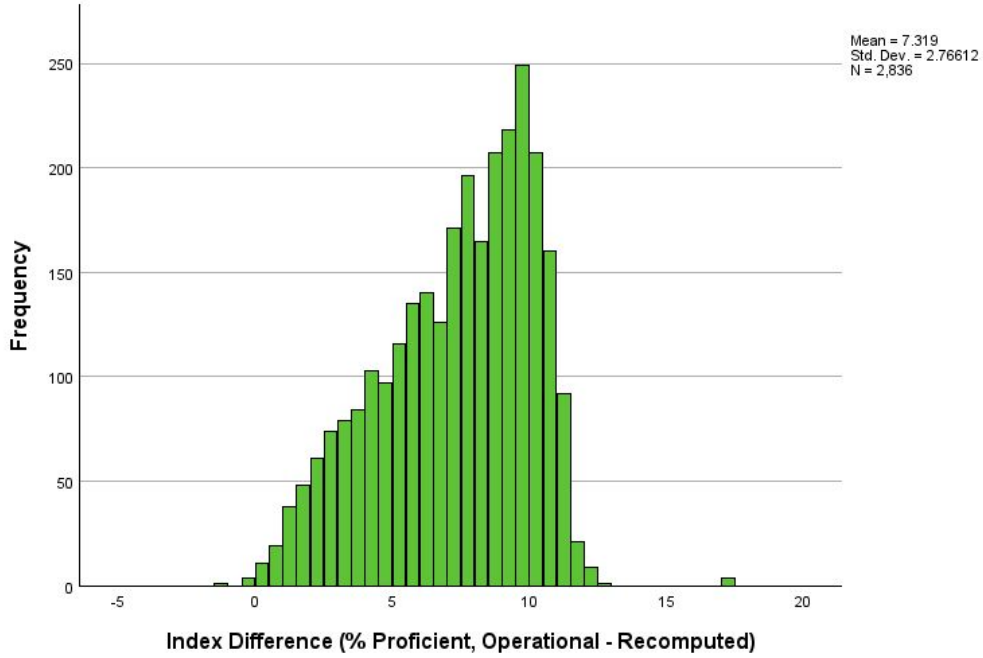
Approach

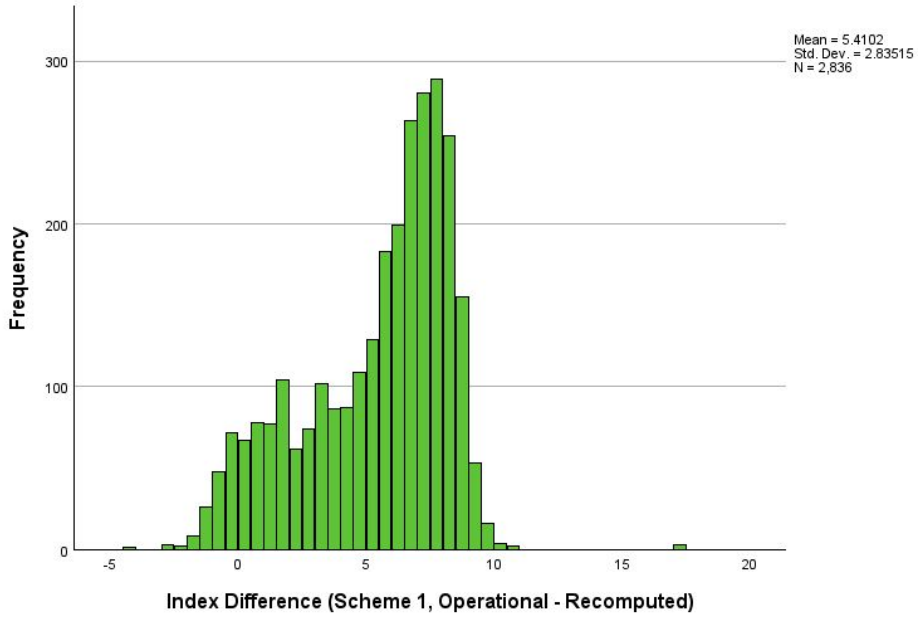
- Compute the achievement indicator three ways:
  - % proficient
  - scoring scheme 1
  - scoring scheme 2

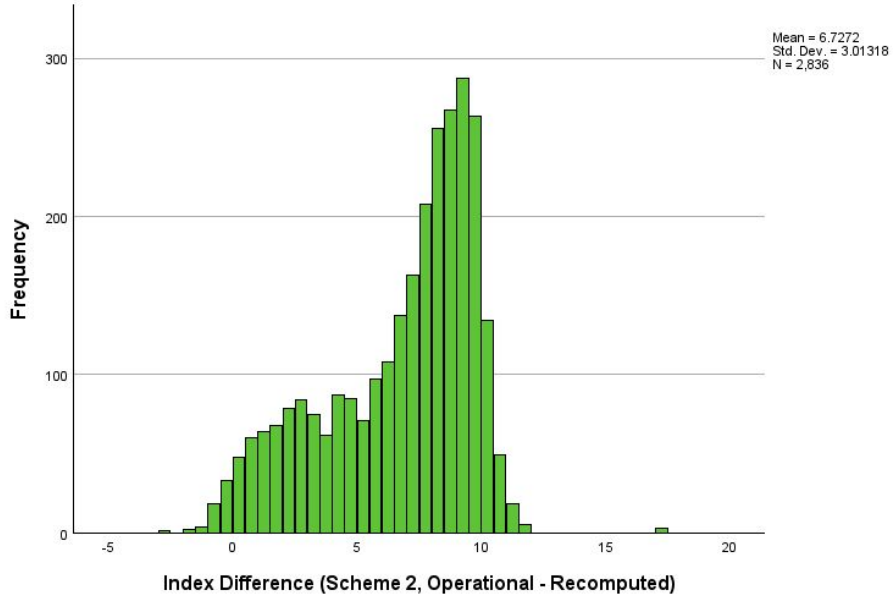
- Compare the values of the resulting composite indices to operational index:
  - globally
  - by current designation

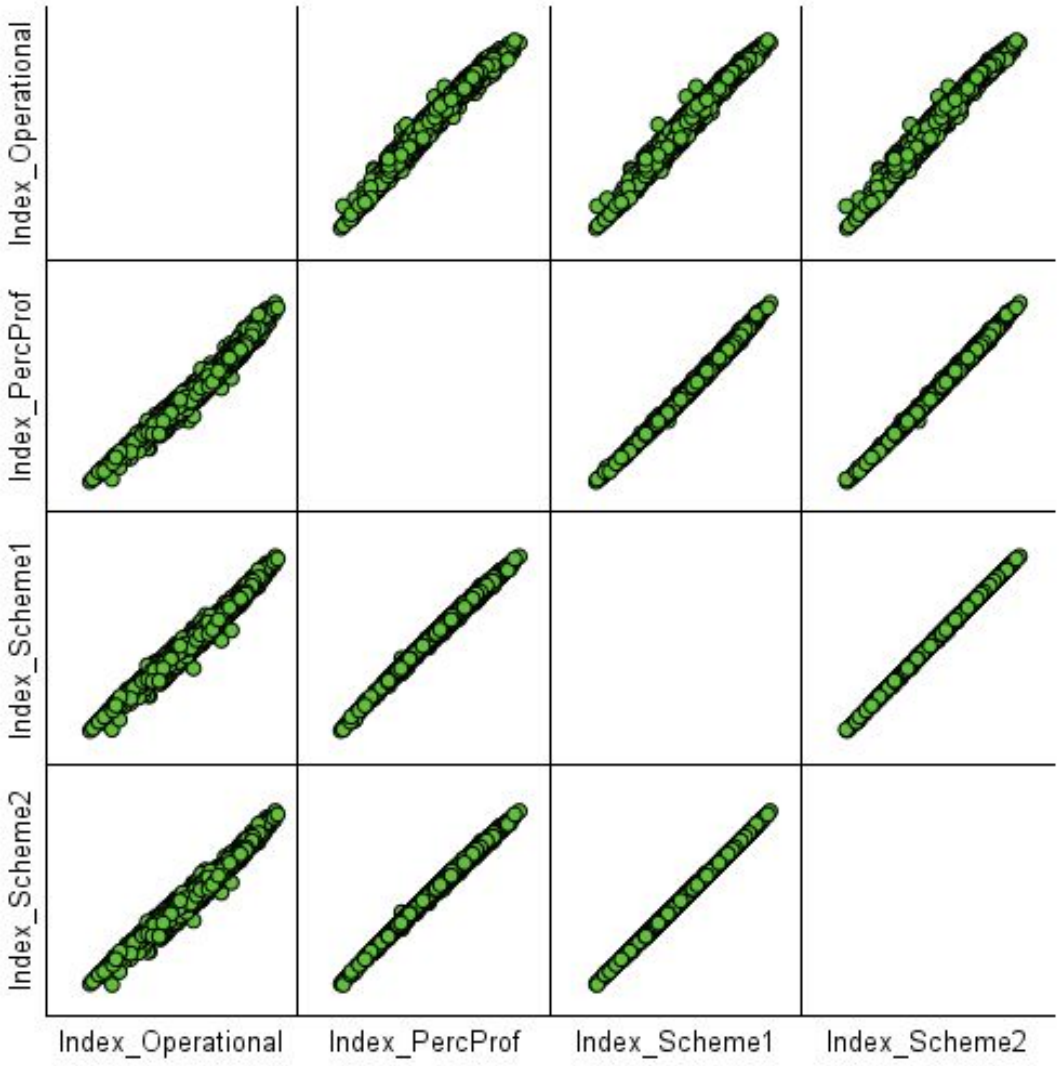
Quick coding check







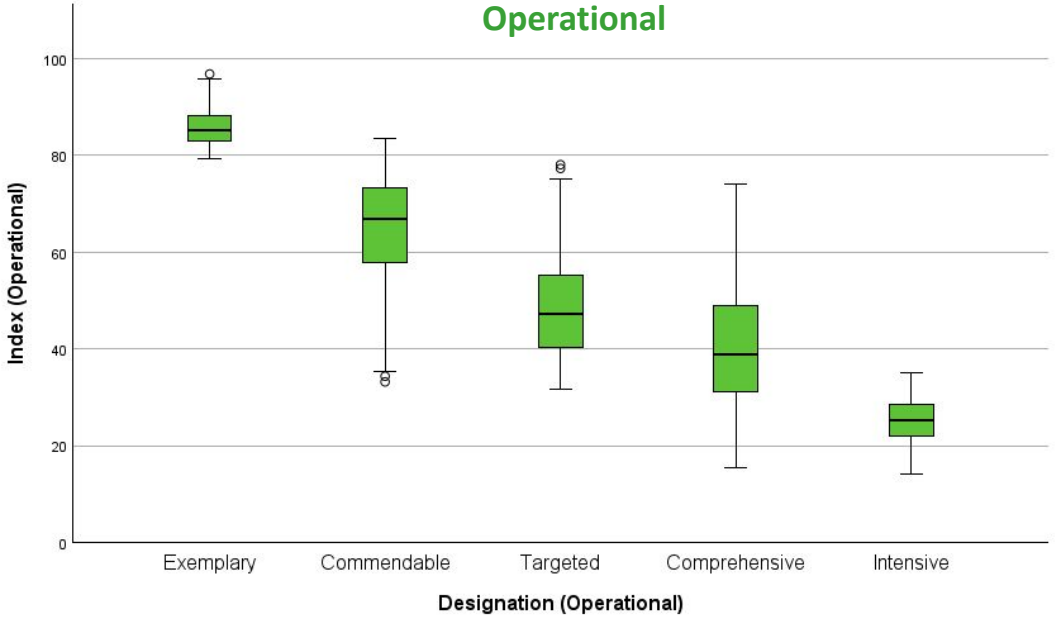




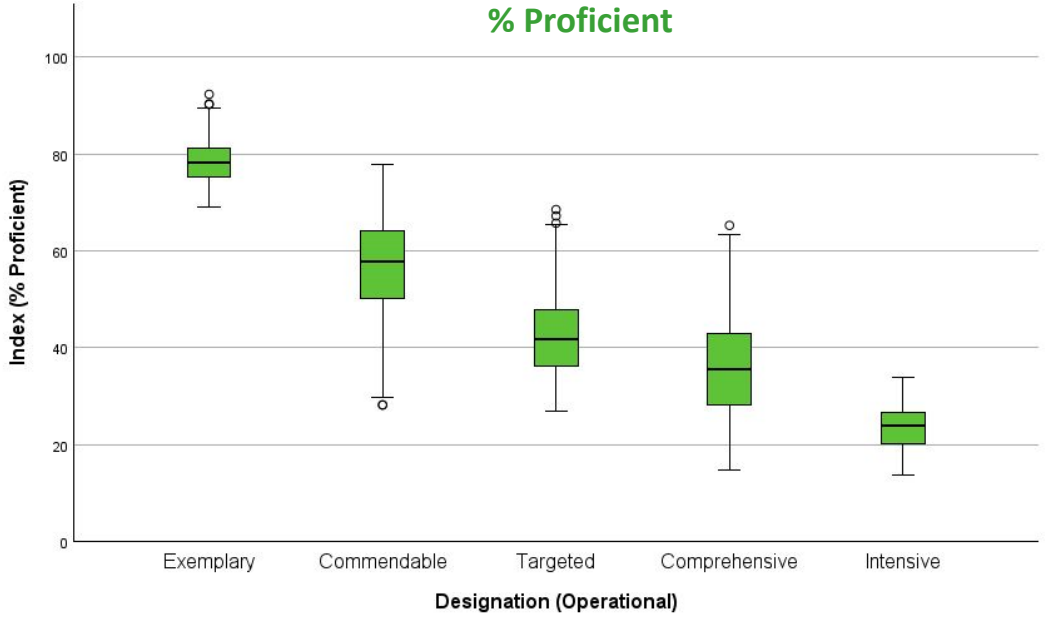
Correlations

	Index_Operational	Index_PercProf	Index_Scheme1	Index_Scheme2
Index_Operational	1.000	.987	.989	.988
Index_PercProf	.987	1.000	.999	.999
Index_Scheme1	.989	.999	1.000	1.000
Index_Scheme2	.988	.999	1.000	1.000

Operational

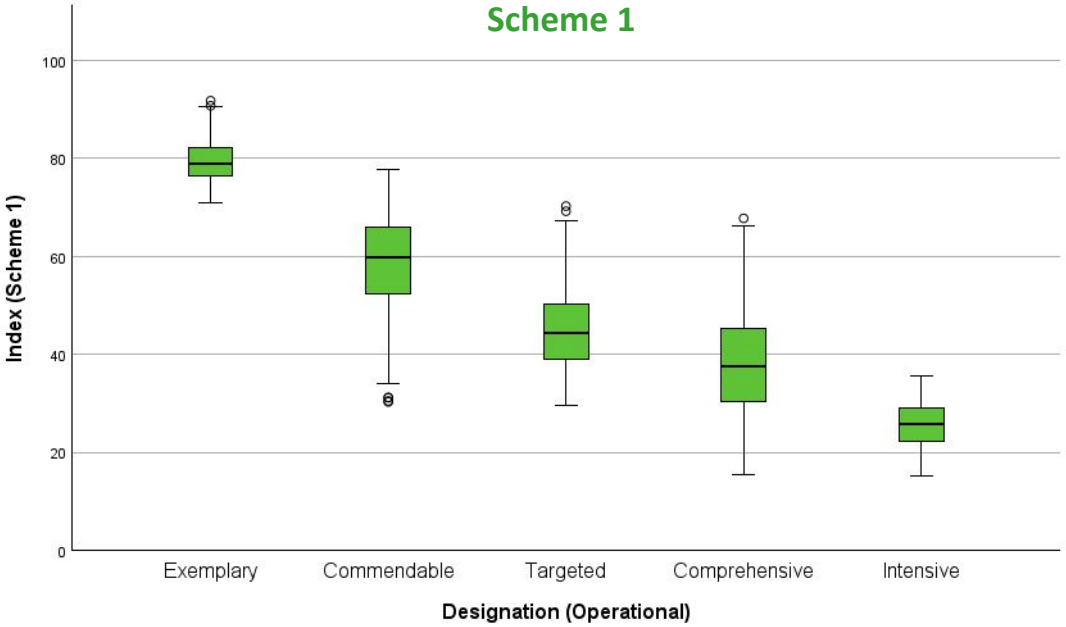


% Proficient

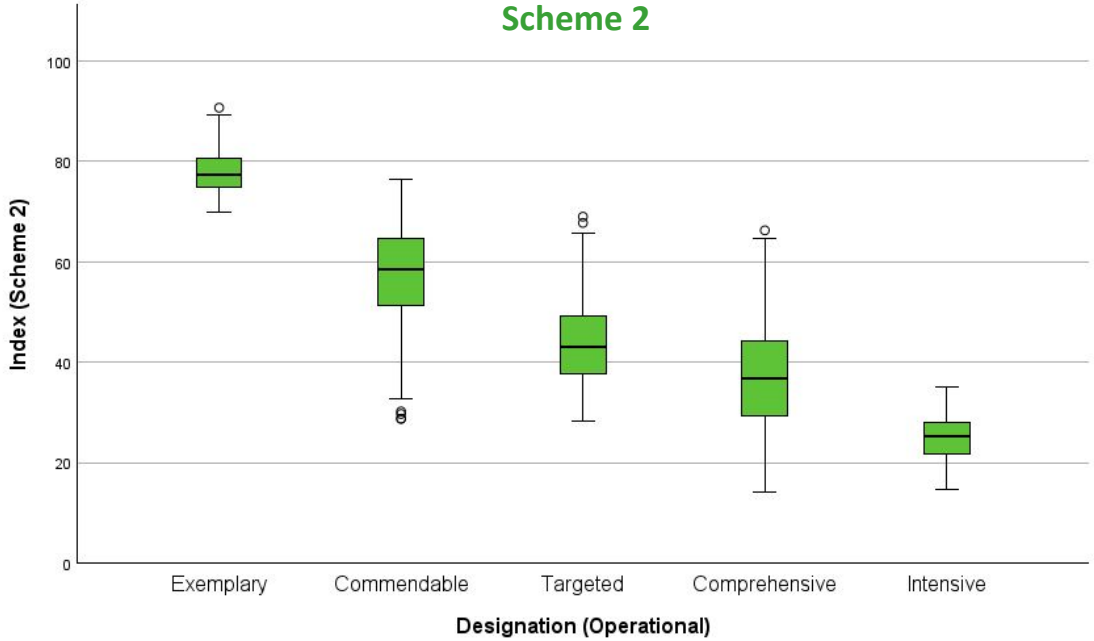


for
ment

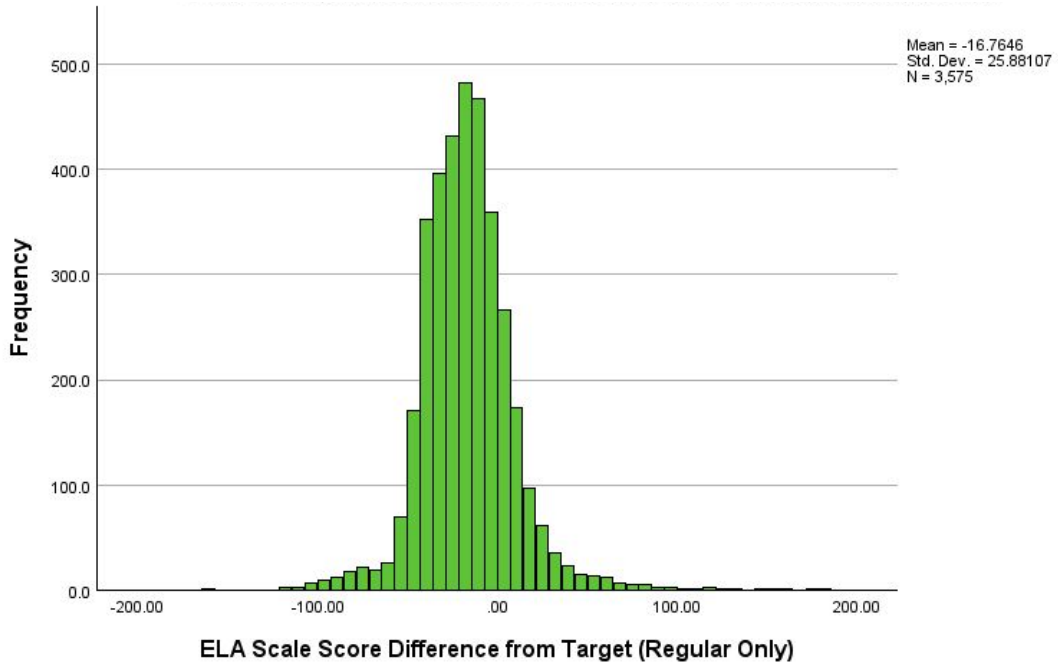
Scheme 1

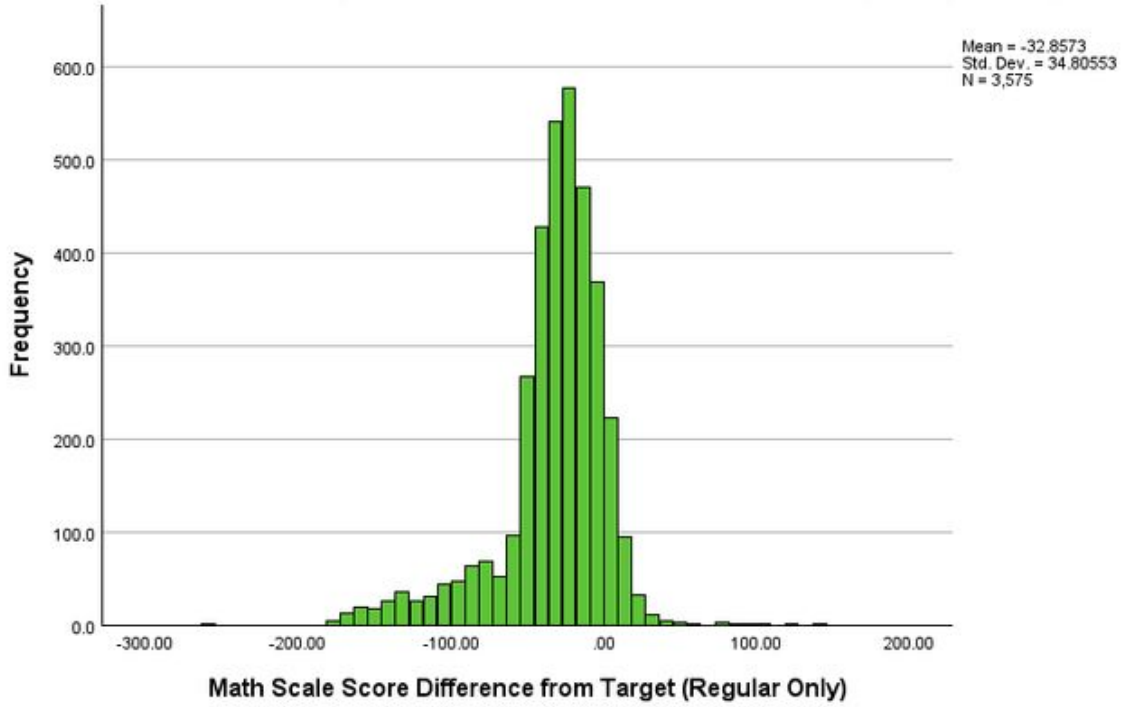


Scheme 2



Scale Score Difference





Exploring the impact of using this with a scale from 0 to 100 requires more complicated scoring such as:

- giving maximum credit at 0 (at best meeting target)
- giving credit in a full range (bonus for exceeding target)

In either case, maybe consider an effective range of -100 to +100 with recoding at boundary values

Observations

- Values of composite indices computed using the simpler / alternative schemes are generally between about 0-10 points lower than the current operational values
- From a rank-ordering perspective, however, the composite index values are all very closely related
- There is negligible impact on the separation of the schools across the different designation categories - using current operational classifications.

Analyses show that, at this level of analysis, simpler/alternative scoring schemes lead to similar outcomes as more complex target-based approaches. However, targets may still be useful for conversations with district leaders to discuss improvement plans and goals.

Center for Assessment

ISBE TAC

Achievement Indicator Alternatives: Extended Exploration

André A. Rupp

In-person TAC Meeting, January 22-23, 2025
ISBE Offices Chicago, IL

Slide deck version: 1/15/25; based on final version from 5/29/24 for June 2024 meeting

Addendum
Analyses: Part 2, June 2024

Original Indicator-level Analyses

Data Preparation

- Merge data from current and fall data set
- Remove schools with < 10 students (distracting in graphical analyses)
- Reference official ISBE business rules from website whenever needed
- Perform various QC checks to ensure no errors were made
- Use SPSS scripting language to allow for retracing of steps

Core Analyses

- Limit analyses to ALL STUDENTS group and exclude subgroup analyses

- Compute alternative achievement level scores:
  1. Raw percent proficient (0-100) [**coarse-grained**]
  2. Scored percent proficient using Schema 1 (0 - 150) [**mid-grained 1**]
  3. Scored percent proficient using Schema 2 (0 - 150) [**mid-grained 2**]
  4. Scale score distance from scale score threshold for 'Proficient' [**fine-grained**]

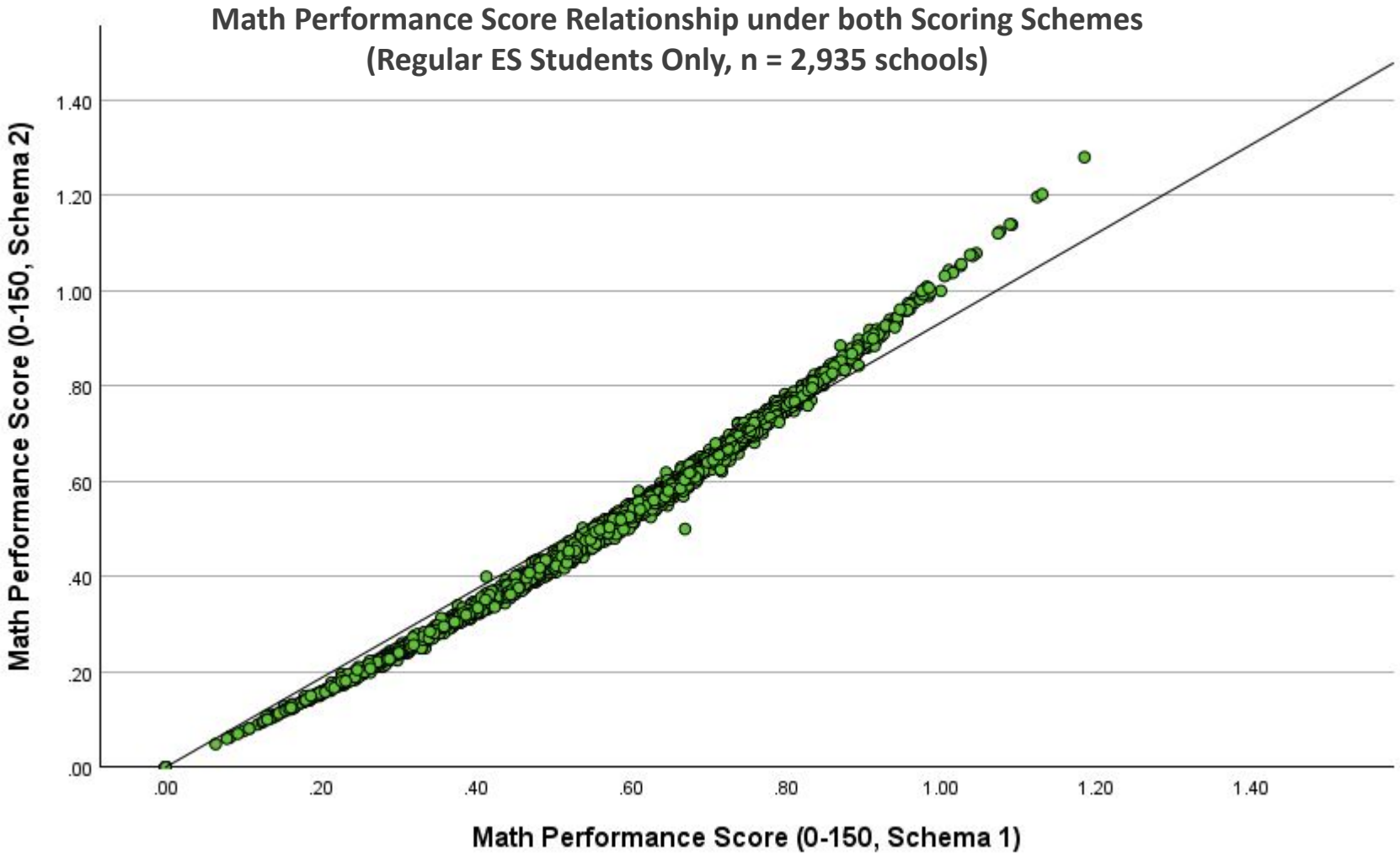
- Compare resulting achievement indicator values to:
  1. Operational indicator values
  2. Each other (pairwise)

- Compare indicator distributions using schools' current designations even if they would have to be recomputed

Outcomes

Math

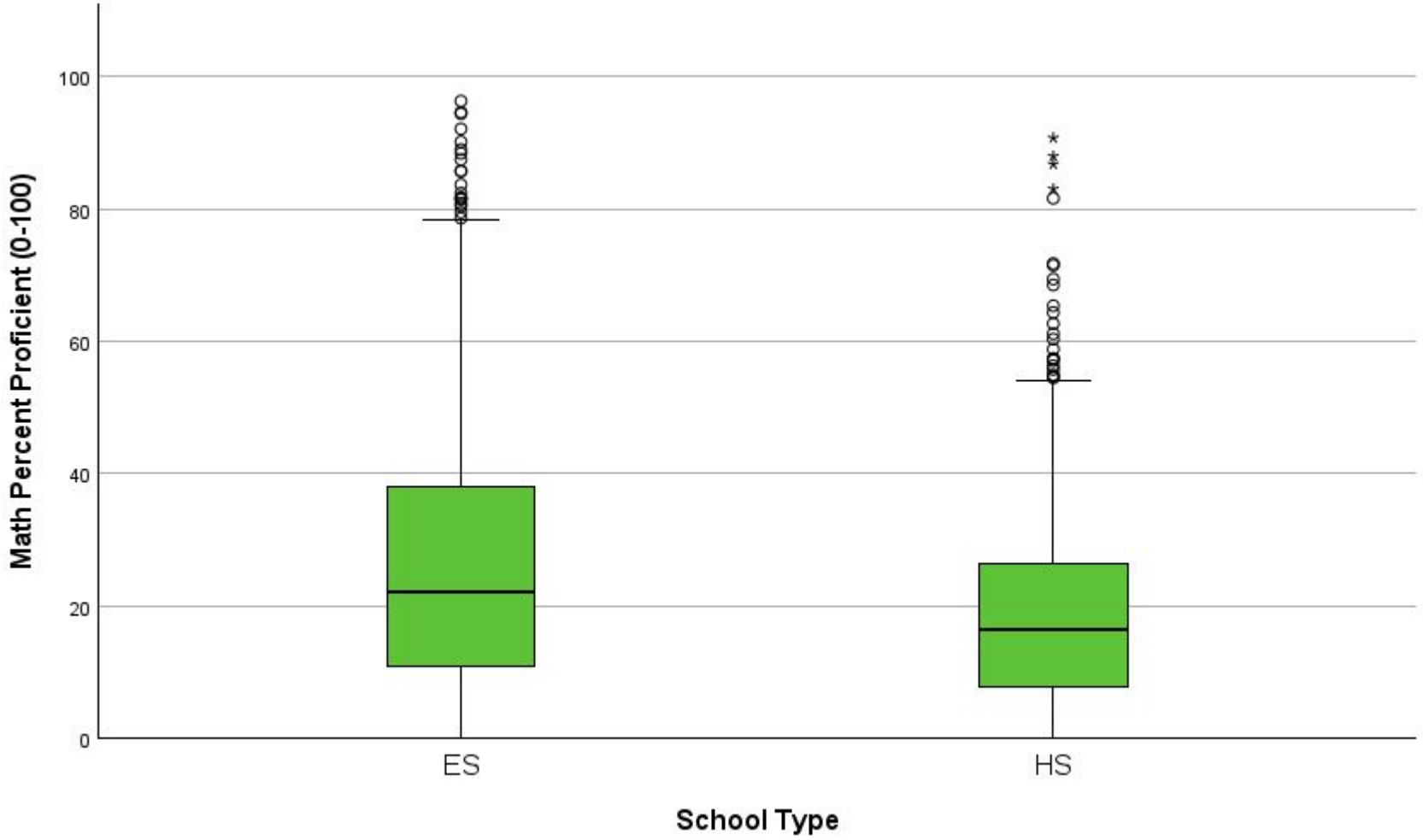
Math


Math Performance Score Relationship under both Scoring Schemes
(Regular ES Students Only, n = 2,935 schools)
Math Performance Score (0-150, Schema 2)
1.40
1.20
1.00
.80
.60
.40
.20
.00
.00
.20
.40
.60
.80
1.00
1.20
1.40
Math Performance Score (0-150, Schema 1)

Math

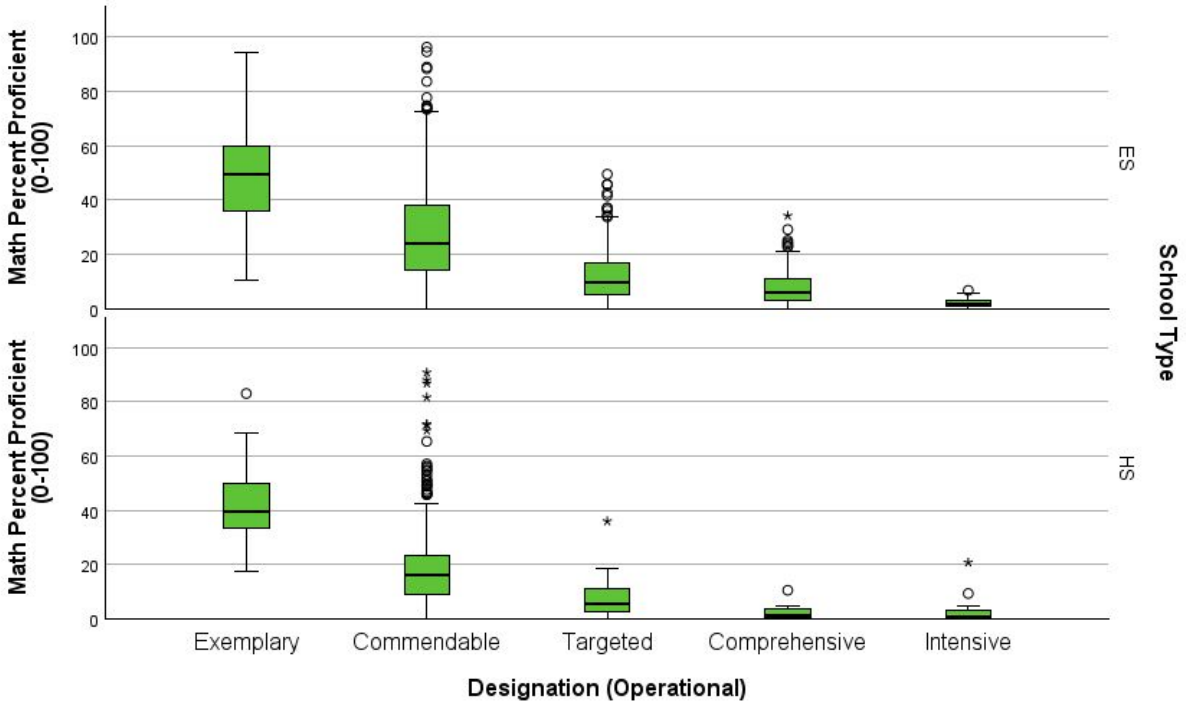
Center for Assessment

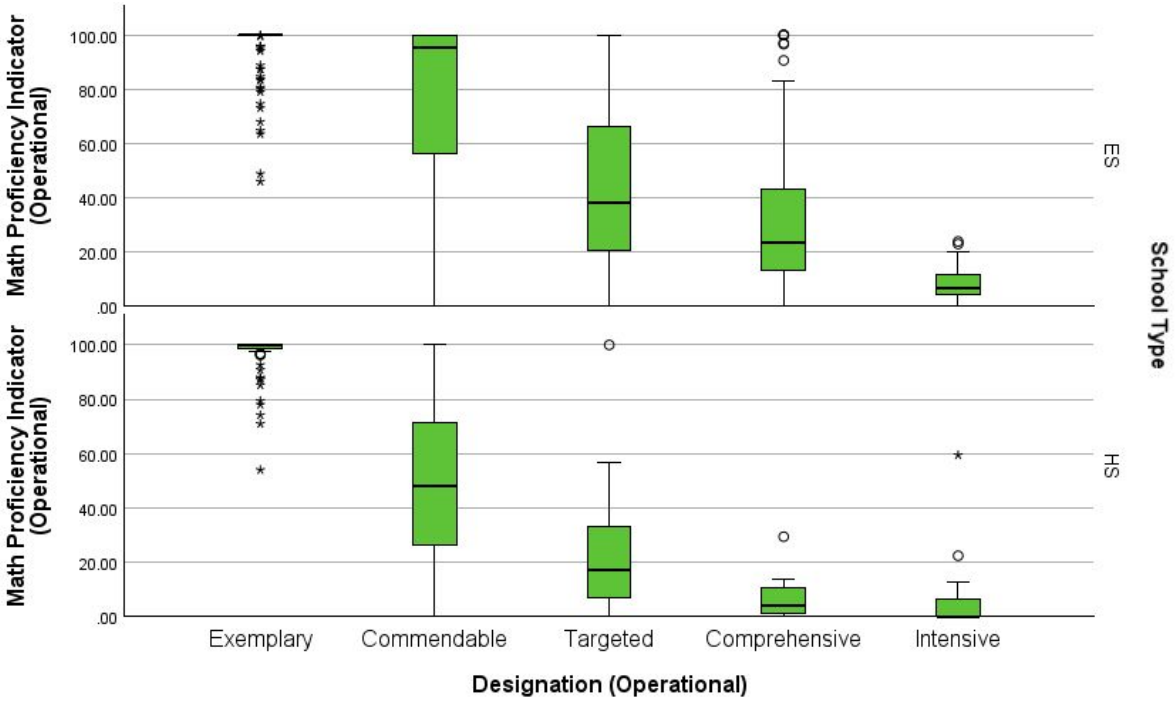
Math Percent Proficient Distribution
(All Students, by School Type, *n* = 3,573 schools)



Math

Math Percent Proficient by Operational Designation
(All Students, by School Type, *n* = 3,573 schools)





Computation

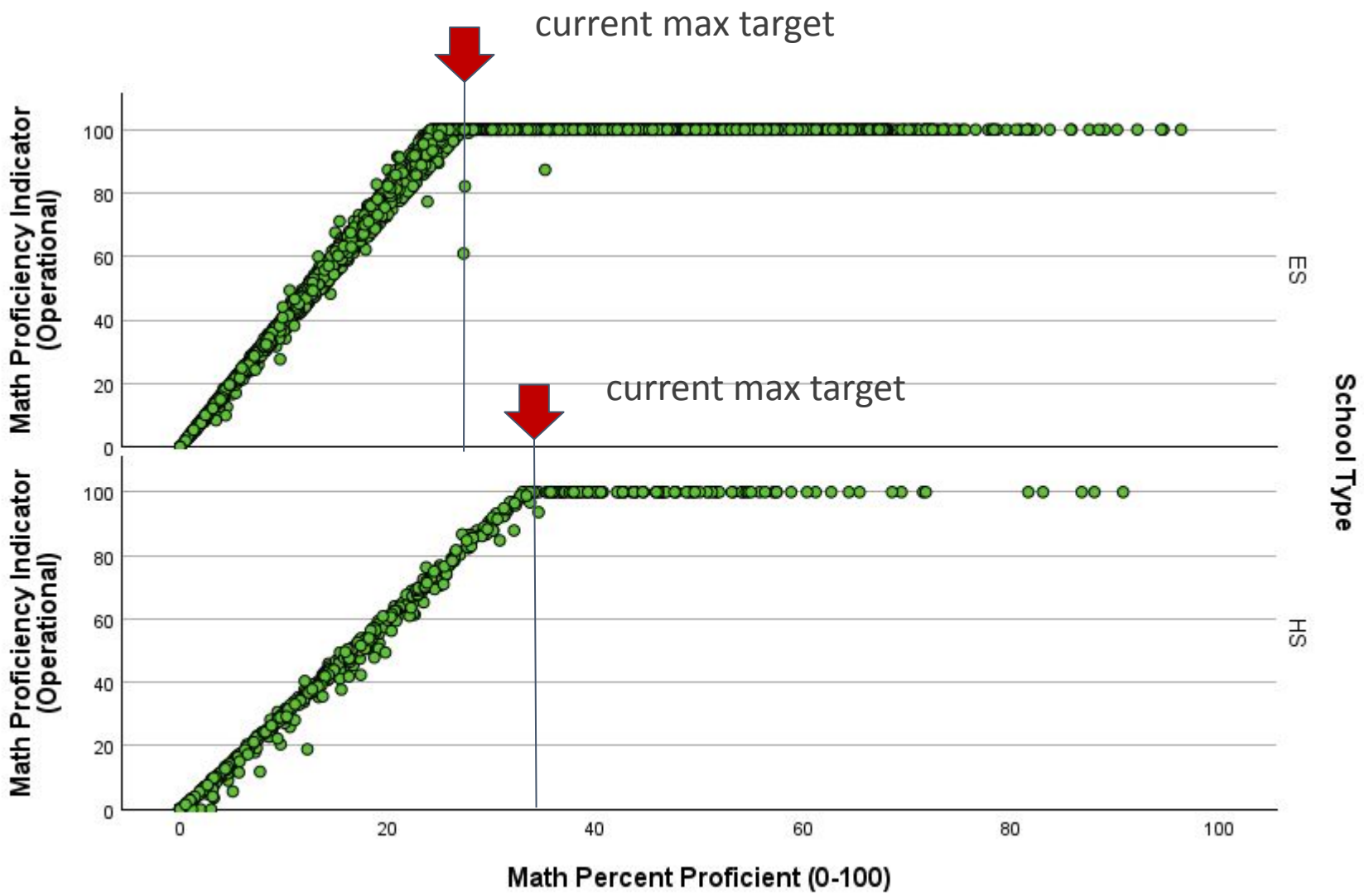
$0 \leq (\%MathProf / \%MathTargetProf) < 1$
$(\%Points\ MathProf / \%MathTargetProf) \geq 1$

→ $((\%MathProf / \%MathTargetProf)*100)$ Points
100 Points

Note. Operational classifications would be recomputed if an alternative indicator value were used during composite index computations.

Math

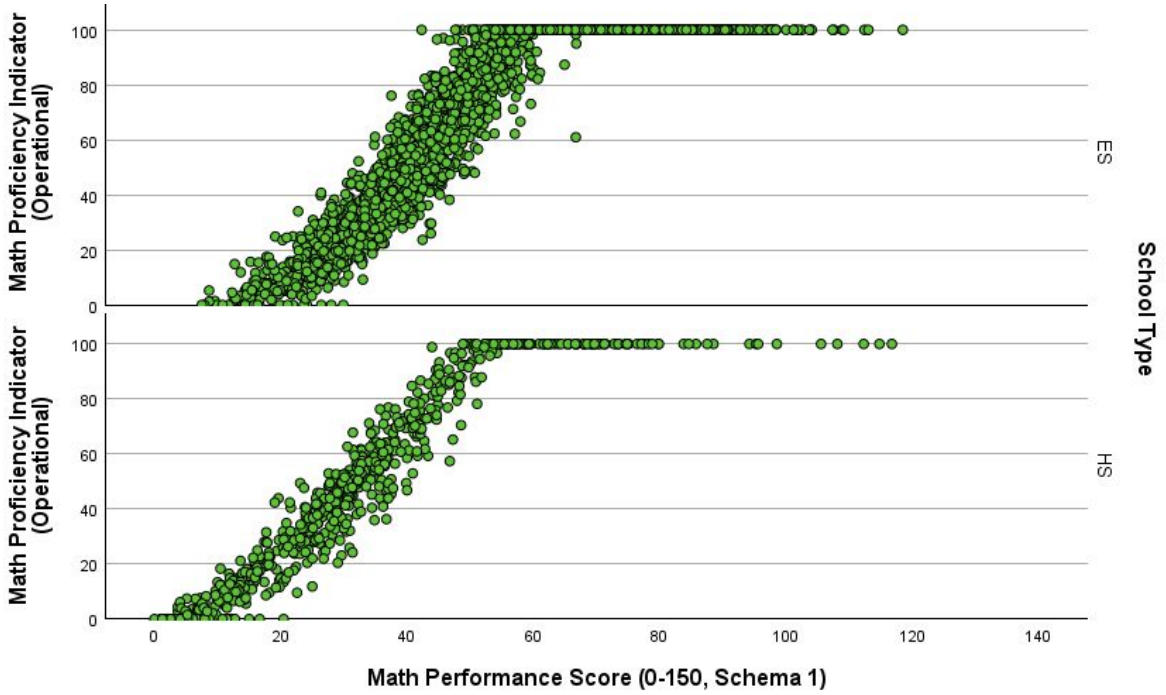
Math Percent Proficient by Operational Indicator Scores
(All Students, by School Type, *n* = 3,573 schools)



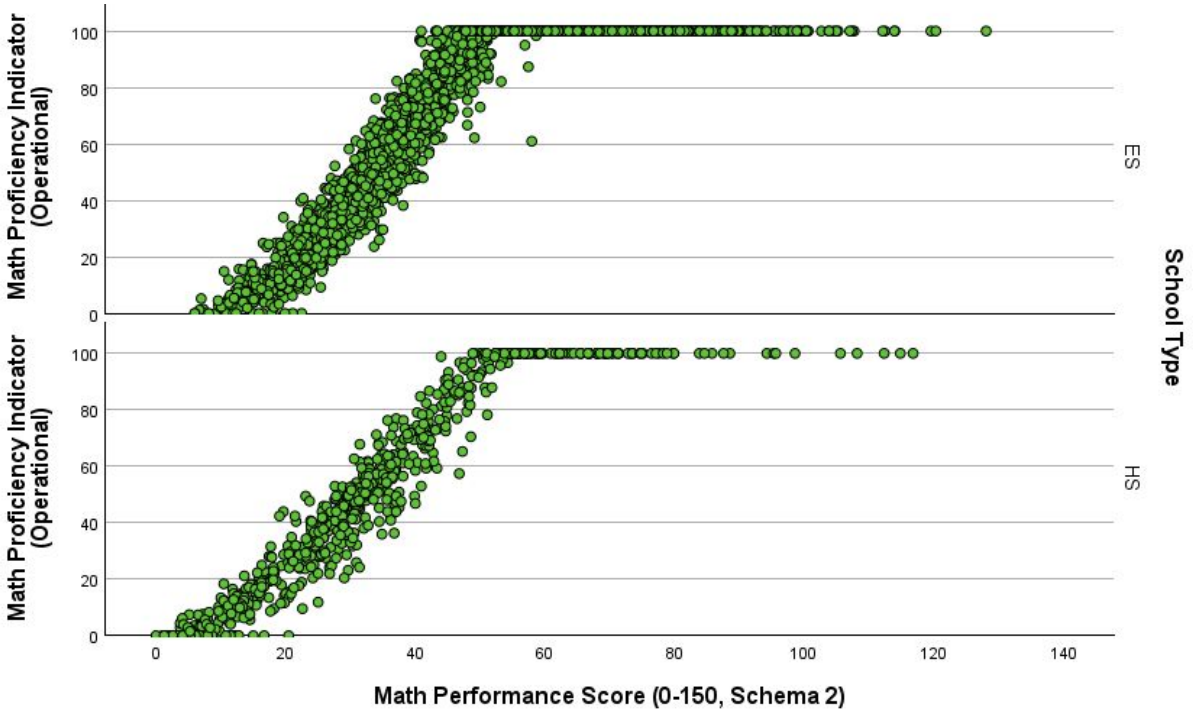
Math

ELA Performance Scores by Operational Indicator Scores
(All Students, by School Type, *n* = 3,573 schools)

Schema 1

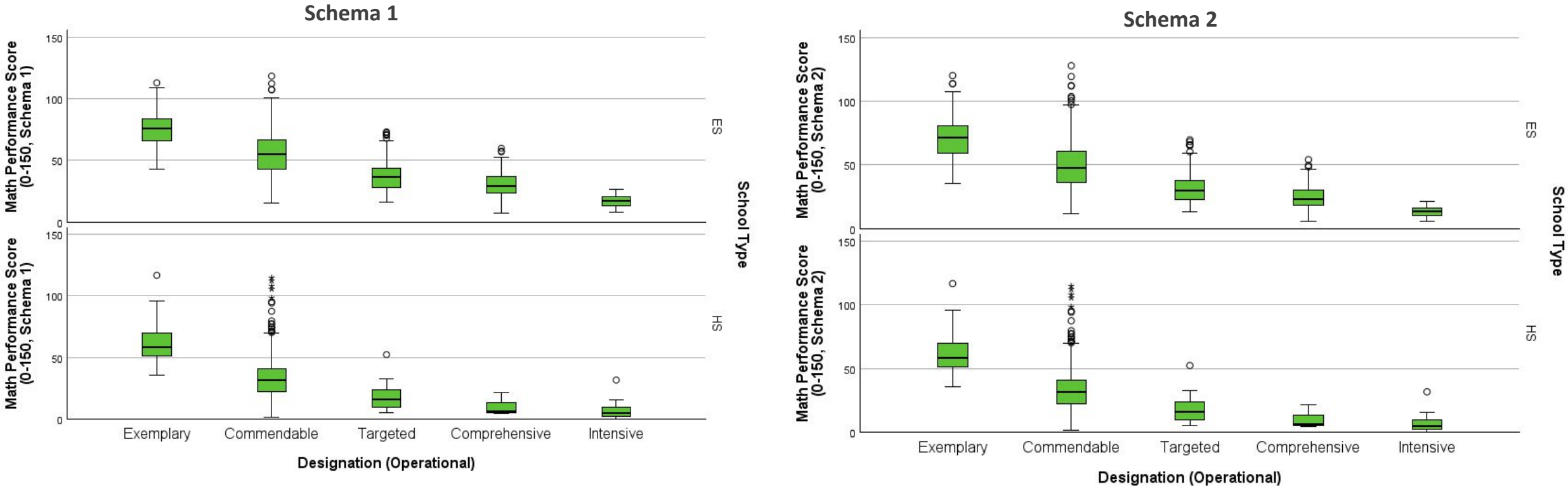


Schema 2



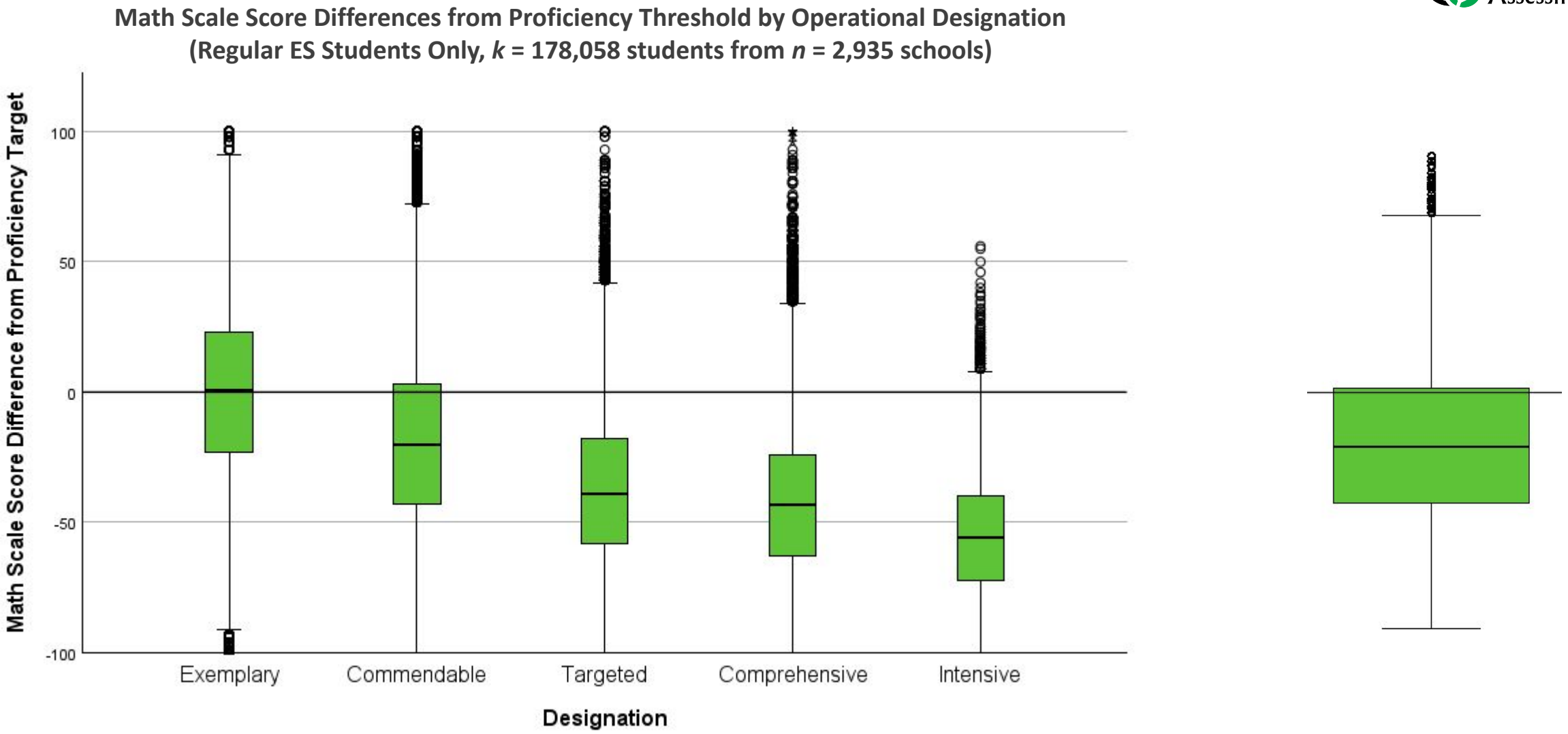
Math

Math Performance Scores by Operational Designation
(All Students, by School Type, *n* = 3,573 schools)



Note. Operational classifications would be recomputed if an alternative indicator value were used during composite index computations.

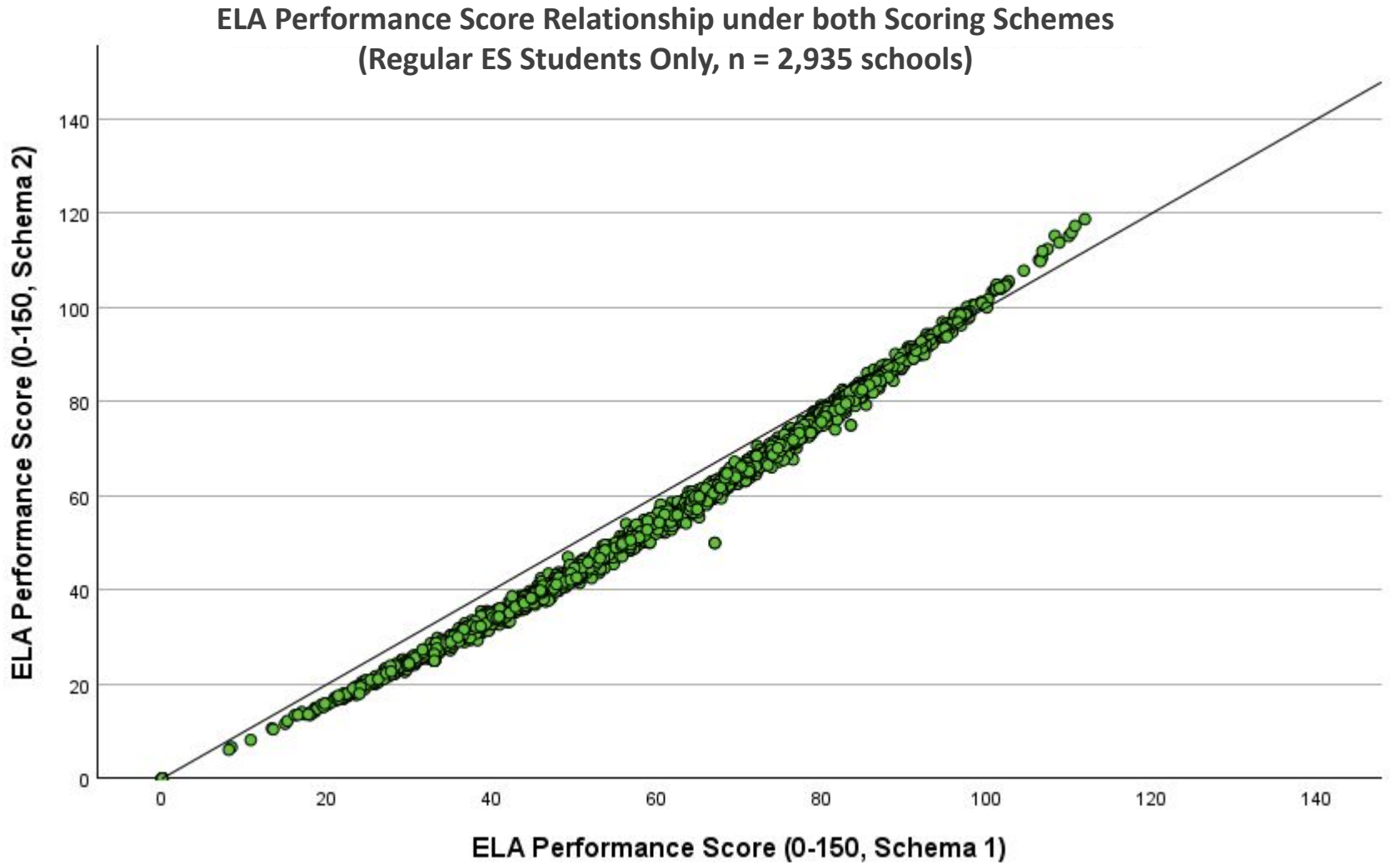
Math



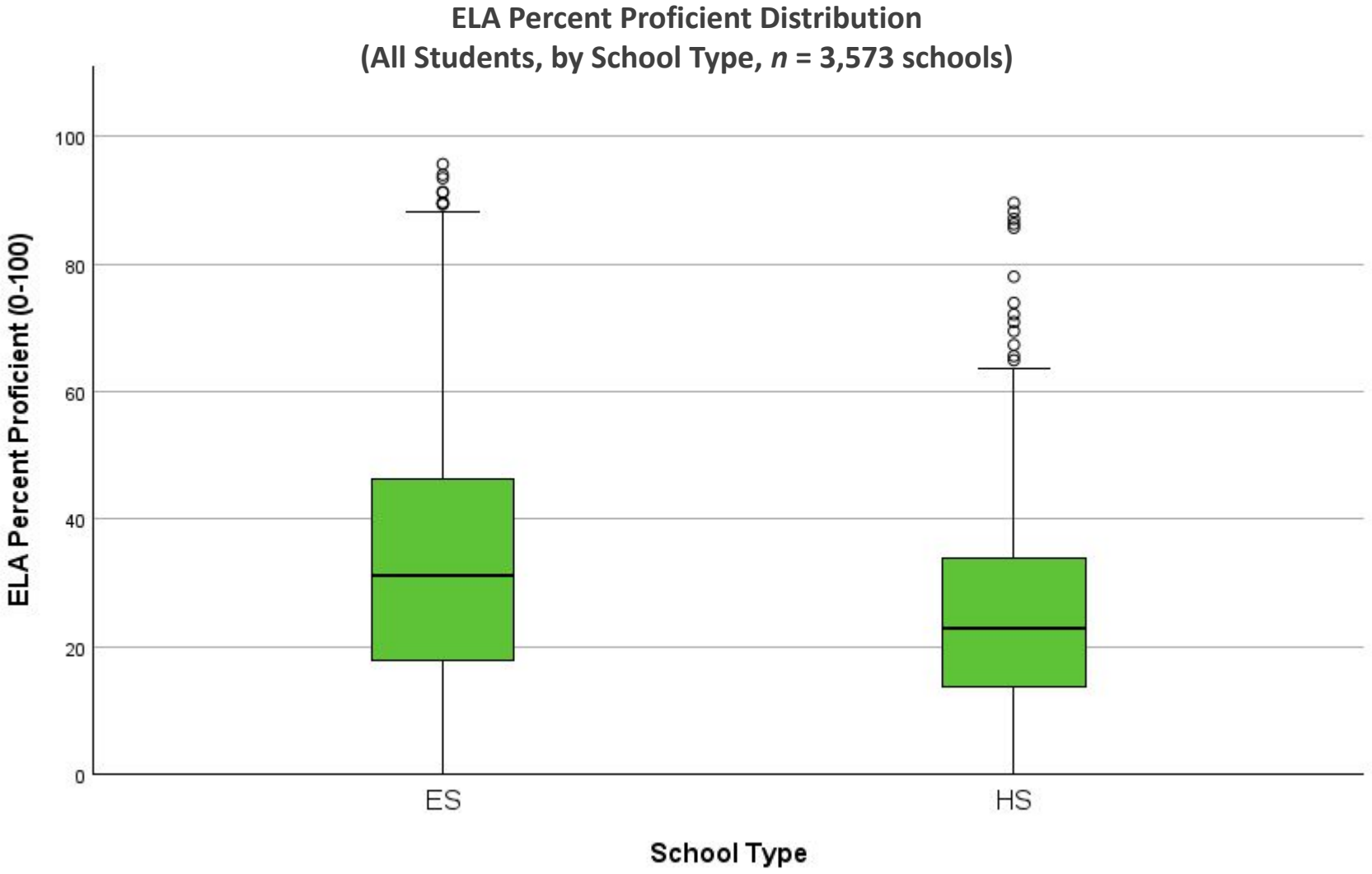
Note. Operational classifications would be recomputed if an alternative indicator value were used during composite index computations.

ELA

ELA

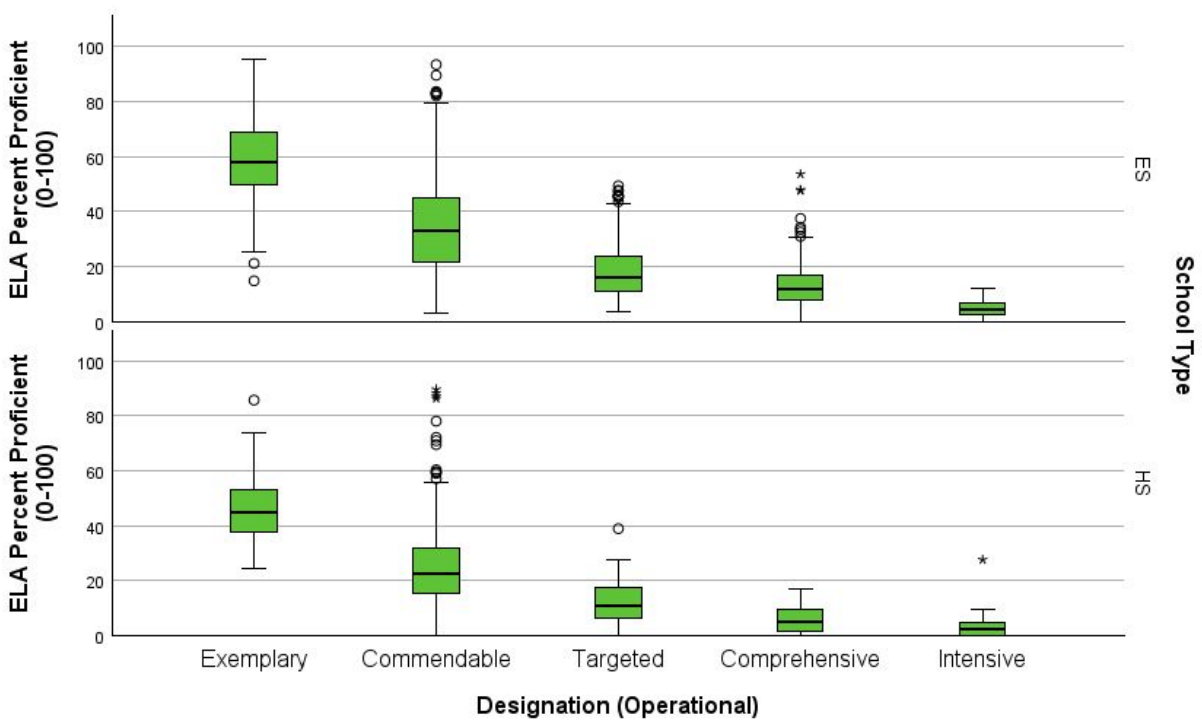

ELA Performance Score Relationship under both Scoring Schemes
(Regular ES Students Only, n = 2,935 schools)
ELA Performance Score (0-150, Schema 2)
0
20
40
60
80
100
120
140
ELA Performance Score (0-150, Schema 1)

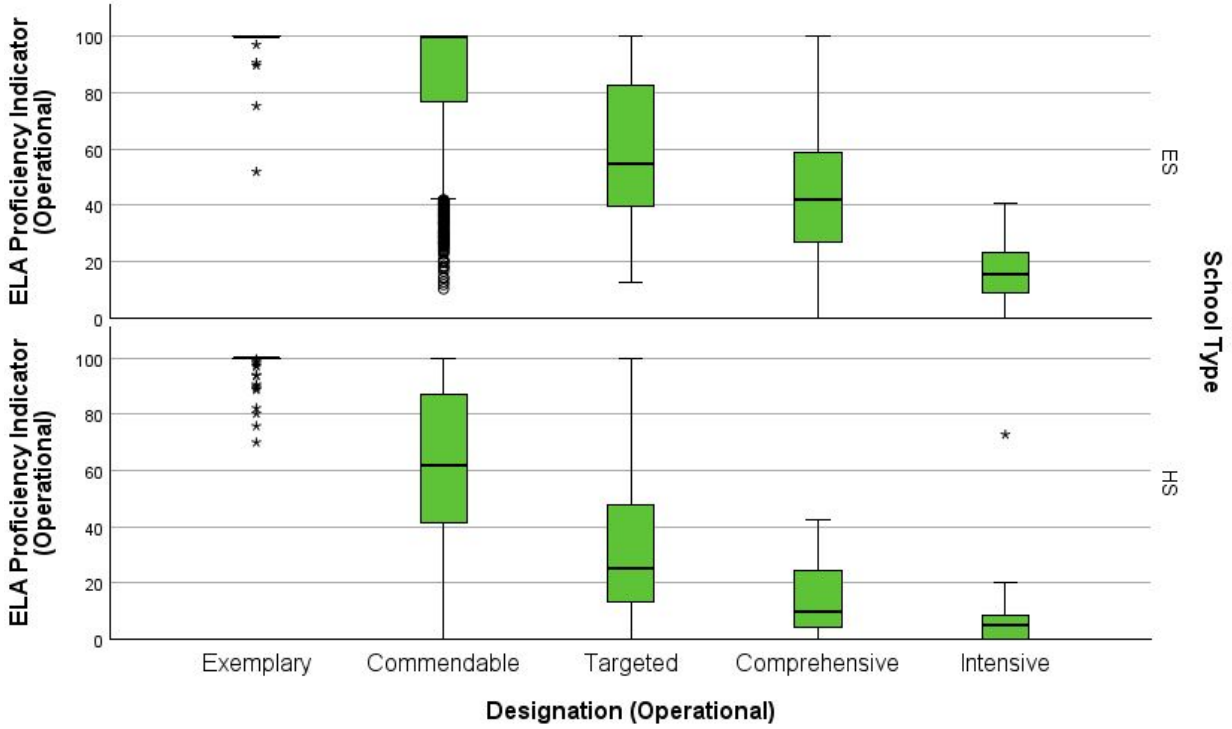
ELA


ELA Percent Proficient Distribution
(All Students, by School Type, n = 3,573 schools)
ELA Percent Proficient (0-100)
100
80
60
40
20
0
ES
HS
School Type

ELA

ELA Percent Proficient by Operational Designation (All Students, by School Type, *n* = 3,573 schools)





Computation

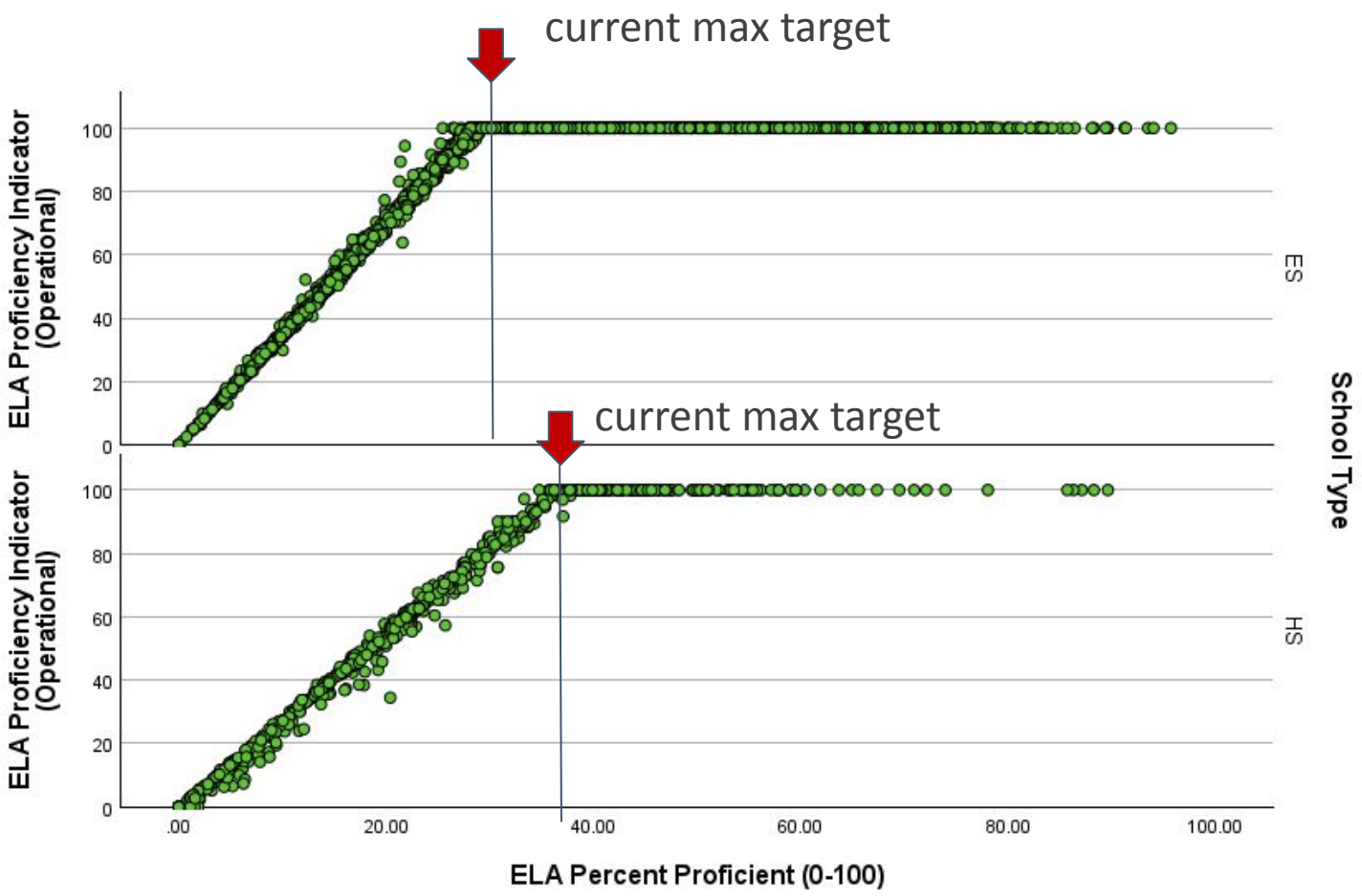
0 ≤ (%ELAProf / %ELATargetProf) < 1 → ((%ELAProf / %ELATargetProf)*100) Points

(%Points ELAProf / %ELATargetProf) ≥ 1 → 100 Points

Note. Operational classifications would be recomputed if an alternative indicator value were used during composite index computations.

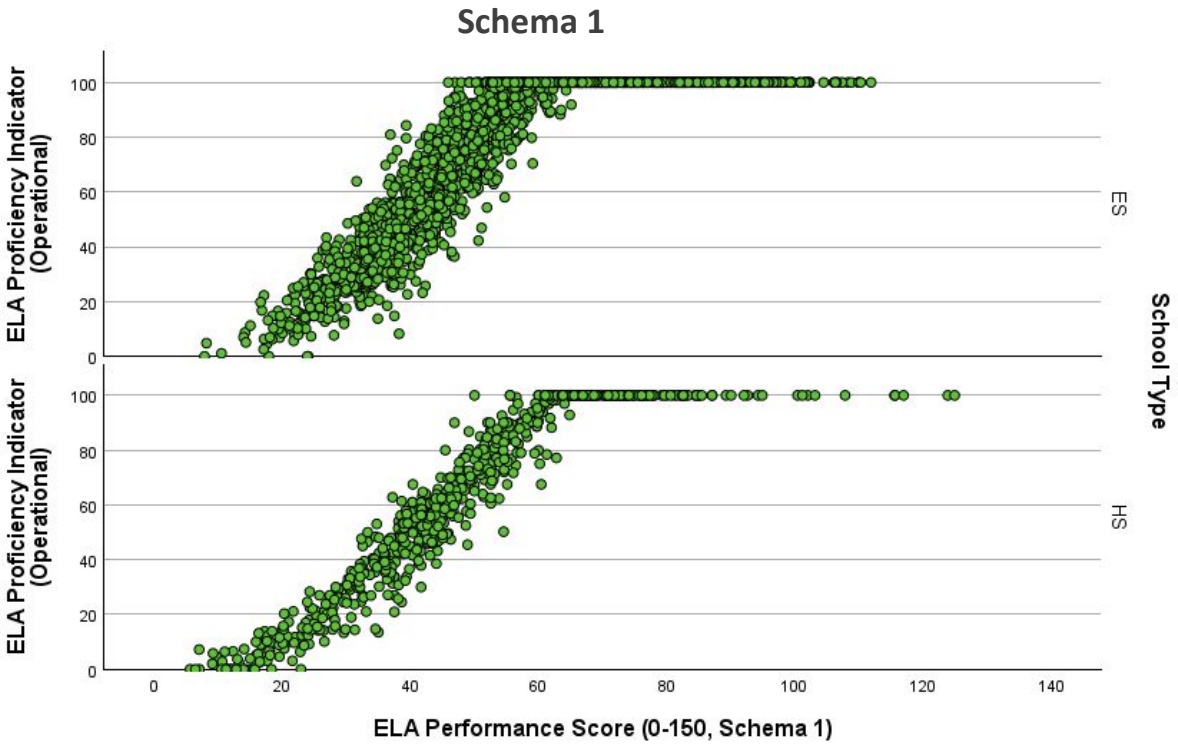
ELA

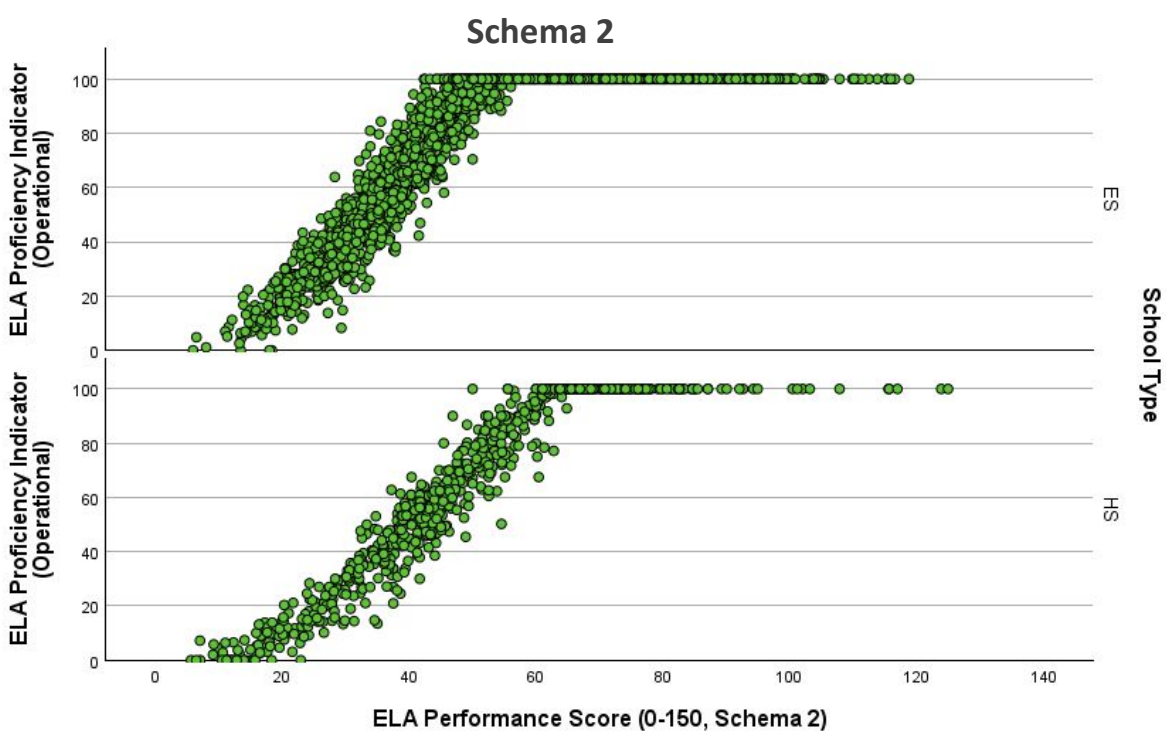
ELA Percent Proficient by Operational Indicator Scores
(All Students, by School Type, *n* = 3,573 schools)



ELA

ELA Performance Scores by Operational Indicator Scores
(All Students, by School Type, *n* = 3,573 schools)

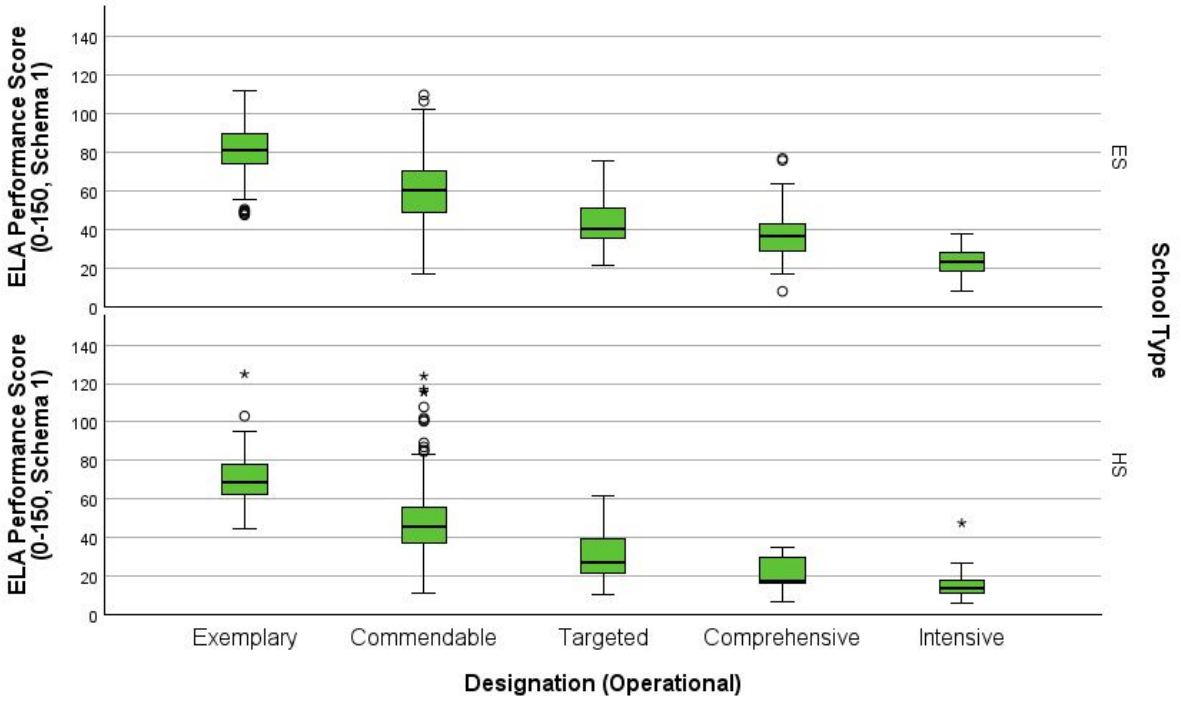




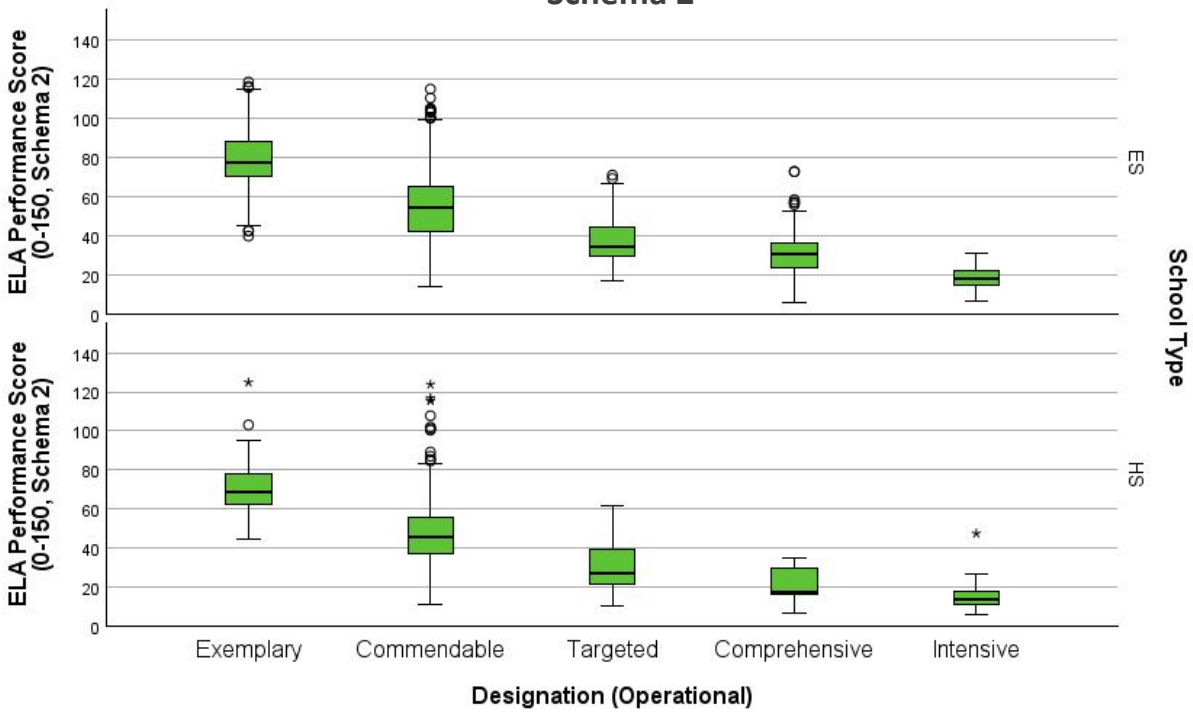
ELA

ELA Performance Scores by Operational Designation (All Students, by School Type, *n* = 3,573 schools)

Schema 1



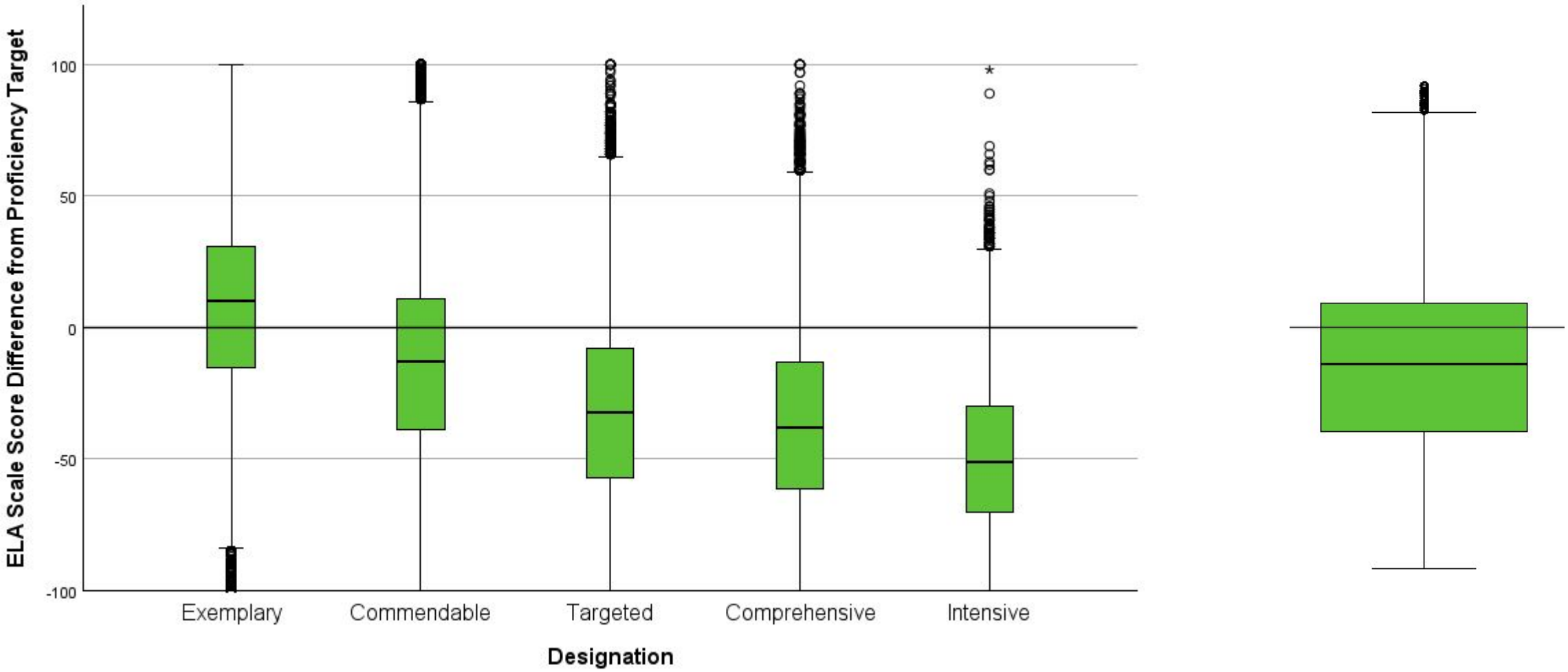
Schema 2



Note. Operational classifications would be recomputed if an alternative indicator value were used during composite index computations.

ELA

ELA Scale Score Differences from Proficiency Threshold by Operational Designation (Regular ES Students Only, *k* = 178,058 students from *n* = 2,935 schools)



Note. Operational classifications would be recomputed if an alternative indicator value were used during composite index computations.

Learnings & Next Steps

Observations

- Analyses at this aggregate level without multi-year data and specific use cases of schools who seem to be currently adversely affected by the computations is somewhat challenging; insights remain limited and are rather coarse-grained.
- Additionally, no analyses were done as of yet to look at subgroup performance, changes in relationships across multiple indicators, or at considering multiple composite index computation approaches in conjunction with achievement indicator computation changes
- Beyond statistical analyses it would also be helpful to further visualize certain distributions across the schools in heat-map type of analyses (e.g., visualizing and sorting schools by % of students in the different performance levels for instance).
- At this initial, exploratory level most visuals simply reflect the statistical properties baked into the design of the scoring approaches
- Given the relatively poor performance of students in most schools with regard to % proficient, scoring scheme 2 yields lower scores than scoring scheme 1 by design even though the resulting scores are similar (r/rho = .99)
- Despite some visual differences, key distinctions across designation groups are not too strikingly different under different schemas. For instance, the approx. lowest 25% of the Commendable schools perform as poorly as approx. up to 75% of the Targeted schools under most schemes.
- Curvilinear downward trends in medians / means across the groups are similar across scoring scoring schemes but significant amounts of distributional overlap remains.

TAC Questions

1. What are productive framings for this work that are most likely to yield empirical insight? If it were more valuable to pivot slightly, what would that look like?
2. Which descriptive and analytic methods would you recommend?
3. Which effect sizes or ways of understanding effects would seem most helpful to make meaningful and informed decisions about impact?
4. What kinds of qualitative information could ISBE provide to shore up the empirical patterns?
5. What are some potential pitfalls that need to be watched out for? Are there risks to replicate problematic, well-known scientific relationships or trends in this space?
6. What is the expected “lift” for this work and how can ISBE meaningfully tackle the required set of analyses given in-house capacity and operational demands?

Center for Assessment

ISBE TAC

Achievement Indicator Alternatives: An Initial Exploration

André A. Rupp, Chris Domaleski, & Will Lorié

In-person TAC Meeting, June 4-5, 2024
ISBE Offices Chicago, IL

Slide deck version: 5/29/24