Illinois Accountability Technical Advisory Committee (TAC)

April Webinar: Defining and Scoring ESSA Indicators

Center for Assessment

April 30, 2018

April Meeting Agenda

- Final comments on April meeting notes
- Finalize TAC's recommendations related to scoring and reporting of each indicator in IL's CSP for inclusion in the Phase 2 report
- Public Comment
- Review future meeting schedule
- Adjourn

Center for Assessment

ESSA Indicators: Goal of Discussion

For each indicator recap and finalize the following:

- Proposed indicator definition (may be identical to that provided in CSP or modified/clarified based recommendations by TAC)
  - Highlight areas where additional data or analyses are needed to operationalize one or more components
- Proposed framework for reporting school level performance on a 0-100 scale.
  - Highlight data necessary to finalize the framework and key factors that should be considered.
- Recommended business rules or reporting specifications
- Analyses to support validation
- Remaining questions/issues

Center for Assessment

General TAC Recommendations Related to Scoring

- *To the extent possible* use a similar framework to support scoring of school performance on 0-100 scale across indicators.
  - **Meet/Exceed long term goal, interim target or other state defined expectation for performance** - assign full points (i.e., 100)
  - **Do not meet annual expectation**: Assign points representing the proportion of the target achieved within the given year (1-99)
    - Percent of interim target met (e.g., academic achievement)
    - Proportion of points obtained within an "effective range" defined by a state specified maximum and minimum value
  - **Performance below a state-defined minimum annual threshold** (0 or some other minimum value).
- Do not penalize small decreases in performance for schools that are performing well.
- Incentivize and reward improvement in schools that are performing above and below expectations.

Center for Assessment

Core Questions Related to Scoring

- How to establish the boundaries of an effective range?
  - Long term goals (when available)
  - Normative procedures
  - Judgmental procedures– state defined
- Should the same range be used for all schools or differ by school type (elementary, middle, high school)?
- How should points for performance within the effective range be calculated?
- Should 3-year interim targets (when provided) be fixed or distributed across years?
- Should student level targets for progress be reset each year (e.g., ELL)?
- Should standards associated with expected performance be baselined or re-normed each year? (e.g., growth)
- What, if any business rules need to be put in place to account for the impact of small N-count on scoring?

Center for Assessment

Academic Achievement Indicator

Indicator Definition: the percentage of students meeting or exceeding standards on the required applicable assessment.

Indicator Weight in Summative Determination:

	2017-2018		2018-2019		2019-2020	
	3-8	HS	3-8	HS	3-8	HS
Math	10%	10%	10%	10%	7.5%	7.5%
ELA	10%	10%	10%	10%	7.5%	7.5%
Science	0%	0%	0%	0%	5%	5%

Long term goal: 90% of students meeting or exceeding expectations in 15 years.

Scoring Recommendation: (Phase 1 Report)

School Outcome	Points
Met long term goal or interim target	100
Decrease percent of students not proficient by **10%** or more	70
Proficiency rate as proportion of target achieved	Varies (0-99)

Lingering Questions
1. What degree of decrease in percent of students "not proficient" should be awarded? Should it be the same for all schools?
2. How should proportional points be calculated given the interim target is defined on a 3-year cycle?

Center for Assessment

Academic Achievement: Lingering Questions

1. **What degree of decrease in percent of students "not proficient" should be awarded?** Depends on:
   – The degree of change that is feasible to expect in a given year AND
   – The degree of change the state believes warrants the points it intends to award (e.g., 70).

❖ **What data and analyses should be considered to support these decisions?**
   ✓ data reflecting the degree of change that can be expected from one year to the next (overall, by school type and by school size)

❖ **Options for discussion:**
   ✓ Identify the degree of change representing the 90th percentile in base year and use this as criterion for earning 70 points
   ✓ Identify a less rigorous criterion (75th percentile) if you want more schools to earn points through this option and/or potentially reduce points to 50.

Center for Assessment

Distribution of School Change in Proficiency Rate: ELA

2015 to 2016 (grade 3 to grade 4)

Basic Statistical Measures			
Location		**Variability**	
Mean	1.068705	**Std Deviation**	10.61776
Median	0.000000	**Variance**	112.73686
Mode	0.000000	**Range**	84.02839
		Interquartile Range	12.81915

Quantiles (Definition 5)	
Level	**Quantile**
100% Max	44.89796
99%	28.71287
95%	19.44444
90%	14.28571
75% Q3	7.50000
50% Median	0.00000
25% Q1	-5.31915
10%	-12.04819
5%	-15.58442
1%	-25.58140
0% Min	-39.13043

82% of schools with a increase of less than 10%

2016 to 2017 (grade 4 to grade 5)

Basic Statistical Measures			
Location		**Variability**	
Mean	-0.15653	**Std Deviation**	10.01902
Median	0.00000	**Variance**	100.38077
Mode	0.00000	**Range**	108.24623
		Interquartile Range	11.92308

Quantiles (Definition 5)	
Level	**Quantile**
100% Max	61.90476
99%	25.71429
95%	16.21622
90%	12.06897
75% Q3	5.76923
50% Median	0.00000
25% Q1	-6.15385
10%	-12.12121
5%	-16.32653
1%	-24.00000
0% Min	-46.34146

87% of schools with an increase of less than 10%

Center for Assessment

Distribution of School Change in Proficiency Rate : Math

2015 to 2016 (grade 3 to grade 4)

Basic Statistical Measures			
Location		Variability	
Mean	-4.42217	Std Deviation	10.63831
Median	-4.00000	Variance	113.17359
Mode	0.00000	Range	97.41834
		Interquartile Range	12.01885

Quantiles (Definition 5)	
Level	Quantile
100% Max	39.72603
99%	23.80952
95%	12.50000
90%	8.33333
75% Q3	1.49254
50% Median	-4.00000
25% Q1	-10.52632
10%	-17.21311
5%	-22.00000
1%	-32.65306
0% Min	-57.69231

2016 to 2017 (grade 4 to grade5)

Basic Statistical Measures			
Location		Variability	
Mean	-0.87031	Std Deviation	9.63641
Median	0.00000	Variance	92.86045
Mode	0.00000	Range	96.48649
		Interquartile Range	10.55108

Quantiles (Definition 5)	
Level	Quantile
100% Max	60.00000
99%	25.00000
95%	14.28571
90%	10.20408
75% Q3	4.30108
50% Median	0.00000
25% Q1	-6.25000
10%	-12.96296
5%	-16.36364
1%	-24.21053
0% Min	-36.48649

93 % of schools with an increase of less than 10%

90 % of schools with an increase of less than 20%

Center for Assessment

IL TAC Webinar - April 2018

Potential Implications of 10% Criterion

	Percentage of Schools that Demonstrated a Change in Proficiency Rate of 10% or Greater (Based only on performance in these grades/content areas)	
	2015-2016 (Grade 3 to Grade 4)	2016-2017 (Grade 4 to Grade 5)
ELA	18%	13%
Math	7%	10%

- This example is limited as it only looks at change in a schools performance with respect to one assessment within a given year (e.g. Grade 4 ELA/Grade 4 Math).
- In practice would look at the change in proficiency rate for all students administered an ELA/Math assessment in the school.

Questions:
1. Do you believe a 10% criterion is reasonable? What additional data do you believe ISBE should consider in making this determination?
2. Should the criterion be the same for ELA and Math?

Center for Assessment

Academic Achievement: Lingering Questions

2. How should proportional points be calculated?

- School score is equivalent to percentage of the next 3-year interim target represented by the current indicator score: (Proficiency rate/State-wide Interim Target)*100= Score
  - May see a significant score drop every 4th year when a new 3-year target is established.

- Distribute 3-year target across years and then calculate proportional points.

2019 G3-8 State ELA Interim Target = 46.5	A 2018 – 10%	B 2018 – 32%	C 2018 – 44.5%	E 2018- 89.2%
2018 Target 46.5	(10/46.5)* 100 22	(32/46.5)*100= 69	(44.5/46.5)*100 = 90	100
2018 Target: 43.1	23	74	100	100

Center for Assessment

Academic Achievement Goals

- CSP reflects a long term goal of 90% for all students and each sub-group.

- Understanding that these are extremely rigorous goals, points are awarded to schools based on the degree to which interim progress goals are met, as previously discussed.

- What are the pros and cons associated with setting such high long term goals and interim targets?

- What are the potential implications given the procedures recommended for awarding points to schools?

Growth

- **Indicator Definition**: mean Student Growth Percentile (SGP)

- **Indicator Weight in Summative Determination:** 50% for elementary and middle schools; 25% for math and 25% for ELA

- **Lingering Questions:**

    1. How should growth be translated to a 0-100 scale?

    2. Should SGPs be baselined or re-normed each year?

    3. Should categories used for letter grades be normed each year?

Scaling Growth

- Options for scaling (i.e. allocating points on a 0-100 scale).
  - Option 1: Simply use MGP
    - Pro: Most straightforward to implement and easy to understand
    - Con: the distribution will be constrained and it will not operationalize the intended weights as effectively
  - Option 2: Transform mean MGP to new scale based on fixed high and/or low values for all schools/groups
    - Pro: relatively straightforward to understand and implement, all MGP map to a single scale value
    - Con: variability among distributions could create 'clustering' at high/low values
  - Option 3: Transform mean SGP to new scale based on variable values for schools/groups (e.g. set high value to 90th percentile for that distribution)
    - Pro: least likely to 'cluster'
    - Con: most complicated to implement and understand; different points would associated with different mean SGP values,

Scaling Growth

- Options
  - Scale 1: fixed low and high values
  - Scale 2: fixed high value only

Scale Options

Center for Assessment
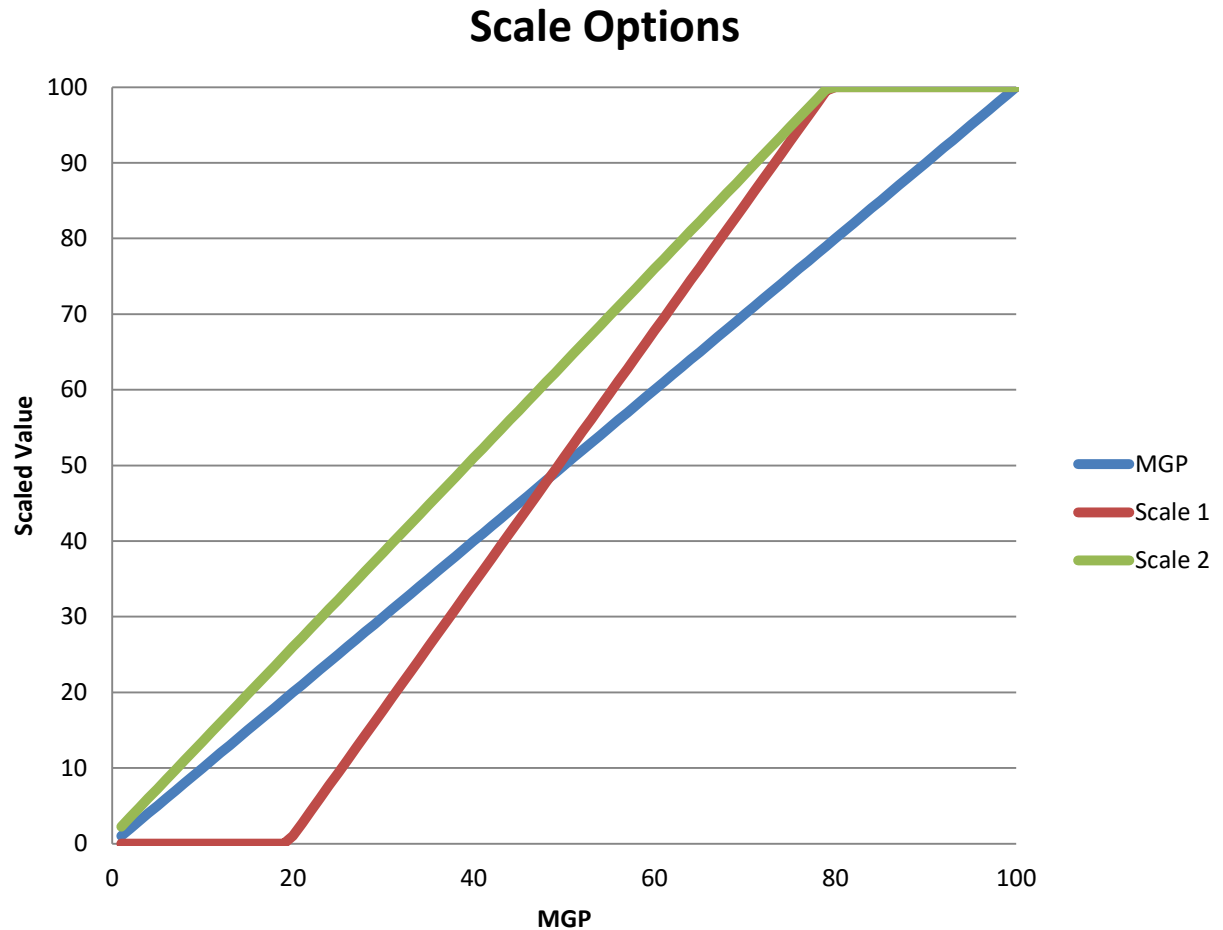
Baseline Growth

- It is possible to establish a baseline for growth such that scale score changes in each year will be 'fixed' to SGPs associated with the baseline distribution
- Pro:
  - If the state is improving in terms of growth rate, the distribution in each year will reflect this improvement. (Conversely, declines would be reflected as well.)
- Cons:
  - Requires multiple years of data to create growth baseline reference
    - Generally at least 2 years is preferred
    - This interacts with priors (e.g. 2 cohorts with 2 priors each = 4 years)
  - Will break from PARCC (unless PARCC also reports baseline reference SGP)
- Recommendation: analyze cohort/scale stability for at least 2 years before baselining.

Center for Assessment

Norm Distribution for Letter Grades?

- At the last meeting, the TAC recommended basing growth letter grades on quintiles for all schools and 'like-school' groups

- Should these quintiles be established each year such that 20% of the schools in each group receive each letter grade?

- Or, should the 'cuts' for the quintiles be fixed and used in each subsequent year, allowing the percent of schools receiving each letter grade to vary?

Center for Assessment

ELL Progress Toward ELP Indicator

Indicator Definition: percentage of ELL students making on-target annual progress toward English Language Proficiency.

- Specifications for individual targets will be established in consideration of existing or additional ACCESS 2.0 data
- May be that the number of years a student has to exit varies depending on baseline proficiency level, grade or both.

Indicator Weight in Summative Determination: 5% for 3-8 and HS

Long term goal: 90% of students making on-target annual progress toward English Language Proficiency in 15 years.

Illustrative Example:

Student Outcome	Points
Prior to year specified as target for ELP	
Meet overall ELP exit criterion of 4.8	100
Meet annual progress target	100
Points calculated as proportional to the annual target that was achieved.	0-99
On or after year specified as target for ELP	
Do not meet exit criterion	0
Meet ELP exit criterion	100

Lingering Questions:

1. How calculate progress for an individual student? (using PL or SS metric).
2. Should the annual target be recalculated each year?
3. Should there be a bonus for students who meet the exit criterion prior to the specified window?

Center for Assessment

How calculate points for student progress?

Example - Student A: In Grade 5 in 2018 (Note: Scaled scores are hypothetical*)

2018 interim target (on-track to meet ELP goal) = 4.0 PL/355 scaled score

2018– 3.8PL /320 scaled score
2017 (grade 4) – 2.5 PL/ 260 scaled score

Option 1: Student score is equivalent to percentage of interim target represented by the 2018 proficiency level (3.8/4.0)*100= 95

Option 2: Student score is equivalent to percentage of interim target represented by the 2018 scaled score (320/355) *100 = 91

Option 3: Score is equivalent to the percentage of points earned (on PL or SS metric) given those necessary to meet interim target in the given year.

- PL: [3.8-2.5 (PL points earned)/4-2.5 (PL points to meet target)]*100 = 86 points

- SS: [320-260 (SS points earned)/355-260 (SS point to meet target)]*100= 63

Center for Assessment

Recalculation of Annual Growth Targets

Student A – Assume 4 years to exit

Student A	Observed Performance	Fixed Targets (given BL) =.85 gain each year	Recalculated Targets after 2018 performance (.8 gain each year)	Recalculated after 2019 (.95 gain each year)
2017 (BL)	1.4			
2018	2.4	**2.25**		
2019	2.9	**3.10**	3.2	
2020	3.4	**3.95**	4.0	3.85
2021		**4.8**	4.8	4.8

Center for Assessment

Bonus points for meeting ELP early?

- **Should there be a bonus for students who meet the exit criterion prior to their specified exit year?**

- What data would you need to see to be convinced that this is a fair option?

  – Students at all grade levels are equally likely to earn bonus points given the manner in which the years-to-exit criterion is established?

  – What else?

- What are the potential positive/negative consequences of this approach?

Center for Assessment

Scoring School-Level ELP

Description of Potential Recommendation

- Calculate average points earned for all ELL students in a school
- Generate state distribution of school mean ELL score
  - Do this separately for elementary schools, middle schools and high schools*
- Award points based on a school's location within the distribution.

Exemplar Framework

School Outcome	Points
School mean greater than 90th percentile	100
School mean between 10th and 90th percentile	11-99 points (interpolated)
School mean less than 10th percentile	10 points

Note: this framework does not consider the percentage of students in the school actually making "target" progress toward ELP. Hypothetically, a school could receive 100 points and not achieve the state –defined targets related to % of students showing target progress toward proficiency.

Center for Assessment

Next Steps for ELP

- What analyses need to be conducted to evaluate whether or not the procedures defined to score schools on this indicator are fair and appropriate?
  - What characteristics and properties must be reflected in the observed results?

Graduation Rate Indicator

Indicator Definition: weighted sum of the four, five and six year adjusted graduation rate.

Weight in Summative Determination:

Cohort	Weight
4-year	30%
5-year	15%
6-year	5%

Long term goal: in 15 years

Cohort	Rate
4-year	90
5-year	92
6-year	92.5

Center for Assessment

Recommendations for Scoring on 0-100 Scale:

- Award full points for meeting state-specified long term goal or interim target
- For schools that don't achieve **the interim target** award points based on percentage of "effective range" attained within the given year.
  - Top of the range would be the long term goal (i.e., 90, 92, 92.5)
  - Bottom of the range could be determined using a normative reference point (e.g., 10the percentile in the distribution) OR a criterion defined by the state (e.g., 70)
- Award 0 points to schools that have a rate below the bottom of the effective range.

Improvement Option:
- Award points for schools that fall below the effective range but have improved.
- Award bonus points for schools above the long term goal that have improved
- Award bonus points for **any** schools that have demonstrated improvement

Graduation Rate Questions

- What recommendations do you have related to defining the effective range for 4, 5 and 6 year graduation rate? Note that some schools scoring below the minimum value may* earn 0 points (i.e., depends on if the state awards for improvement*).
  - Make upper limit of the effective range the long term goal associated with each grad rate?
  - Establish a different lower limit for 4, 5 and 6 year grad rate based on base year grad –rate distribution, or a common LL for all?
- Should the 3-year interim target be the annual criterion for performance or distribute the interim target over three years?
- Should the state award improvement, and if so:
  - should this be for all schools or only schools above/below the established effective range?
  - what degree of improvement should be awarded and how many points would be reasonable given the manner in which schools within the range earn points?

  Option: Assign points based on consecutive years of improvement (of any size?) toward interim target or beyond the long term goal: 1-year = 5 points; 2-years=10 points; 3 years=15 points. Reset every 3 years when a new interim target is defined.

Center for Assessment

Graduation Rate Summary Data

- Is it reasonable to use the long term goal associated with each grad rate as the top of the effective range? (90, 92, 92.5)
- Should the lower limit of the effective range be the same for the 4, 5 and 6 year graduation rate indicator or should it be normatively derived?

4- Year Graduation Rate												
Year	N	Mean	Std Dev	Min	5%	10%	25%	50%	75%	90%	95%	Max
2016	668	84.836	14.388	0	59.5	71.8	81.8	88.5	93.25	95.8	97.5	100
2017	707	87.97	12.36	8	69.1	77.3	85.2	91.1	95.1	97.8	99	100

5- Year Graduation Rate												
Year	N	Mean	Std Dev	Min	5%	10%	25%	50%	75%	90%	95%	Max
2016	663	87.081	11.575	0	69.1	75.8	83.3	89.7	94.2	96.9	98.2	100
2017	686	88.801	13.618	0	70	79.4	86.7	92.3	96	98.3	99.8	100

6- Year Graduation Rate												
Year	N	Mean	Std Dev	Min	5%	10%	25%	50%	75%	90%	95%	Max
2016	661	88.057	10.160	0	71.4	76.2	84.2	90.4	94.6	97.4	98.4	100
2017	667	88.849	13.174	0	72.4	79.3	86.9	92.1	96	97.9	99	100

Center for Assessment

Graduation Rate Questions

- What recommendations do you have related to defining the effective range for 4, 5 and 6 year graduation rate? Note that some schools scoring below the minimum value may* earn 0 points (i.e., depends on if the state awards for improvement*).
  - Make upper limit of the effective range the long term goal associated with each grad rate?
  - Establish a different lower limit for 4, 5 and 6 year grad rate based on base year grad –rate distribution, or a common LL for all?
- Should the 3-year interim target be the annual criterion for performance or distribute the interim target over three years?
- Should the state award improvement, and if so:
  - should this be for all schools or only schools above/below the established effective range?
  - what degree of improvement should be awarded and how many points would be reasonable given the manner in which schools within the range earn points?

  Option: Assign points based on consecutive years of improvement (of any size?) toward interim target or beyond the long term goal: 1-year = 5 points; 2-years=10 points; 3 years=15 points. Reset every 3 years when a new interim target is defined.

Center for Assessment

Graduation Rate Example

LTG (2032): 90% grad rate

		School 4-yr grad rate			
	State (annual)*	A	B	C	D
2016 (Baseline)	**85.5**	84	92	72	67.2
2017	**(85.7)**	85	92	72	67.1
2018	**(86)**	86	93	72.6	67.2
2019	**86.3 (target)**				

*if 3-year target was distributed across years

Scoring (Assume ER (70-90)

School	Points in 2018
A	Proportion of effective range earned in 2018: 86-70/(90-70) = .80 = 80 points • if add 10 bonus points for 2 years consecutive improvement = 90 • If use distributed target, rather than 3-year = 100 + 10 (improvement) = 110
B	100 points for exceeding LTG. • 105 points if award 5 bonus points for 1 year improvement
C	Proportion of effective range earned in 2018: 2.6/20=.10 = 13 points • 18 points if award 5 points for 1 year improvement
D	Improvement 1 year = 5 points

Center for Assessment

Scoring School-Level Chronic Absenteeism

Definition:
Percentage of students who are chronically absent.

Weight in Summative Determination:

3-8: 25%

9-12: 7.5%

Recommendation Framework

- If have a CA rate lower than X% (as defined by state-established range) = 100 points

- If have a CA rate higher than X% = minimum points (e.g. 10)

- If have a CA rate within the defined effective range min and max, interpolate points to be awarded from min pts.-100

Lingering Questions
1. How should the effective range be established?
   - In light of base year distribution?
   - State judgement related to regarding absentee rate that should receive max/min points
2. Should schools receive bonus points for improvement above or below the effective range?

Center for Assessment

Scoring School-Level Climate Survey

Definition:
Percentage of students who participate in survey.

Weight in Summative Determination: 5%

Establish common state-defined annual **expectations** for participation rate that do not change.

Outcome	PTS
96-100% participation	100
90-95%	75
85-90%	50
80-85%	25
Below 80%	0

Lingering Questions
1. Are the specified outcome ranges and associated points reasonable?
2. What additional information would be necessary to inform this decision?
3. What business rules, if any, should be in place to mitigate the impact of N-count.

Center for Assessment

Scoring School-Level 9th Grade On-Track

Definition:
percentage of 9th grade students meeting credit earning requirements

Weight in Summative Determination for HS: 7.5%

- Establish common state-defined criteria for percent of 9th grade students on-track that reflect the state's priorities rather than current distributions.

School Outcome	Points
Meet annual on track expectation target (e.g., 95% of students)	100
Fall within effective range (e.g., 71-94%)	Interpolate points from 1-99 within the effective range
Fall below CE expectations (70% of students)	0

Lingering Questions

1. What is a reasonable expectations for percentage of 9th grade students meeting on-track requirements?
2. Under what rate should schools earn zero points?

Center for Assessment

School-Level Freshman on Track Rates

Year	N	Mean	Std Dev	Min	5%	10%	25%	50%	75%	90%	95%	Max
2016	667	83.569	18.743	0	57.9	69.7	80.5	89.1	93.8	96.8	98.3	100
2017	690	86.325	13.666	0	66.7	73.85	83.2	89.9	93.9	96.85	98.2	100

Center for Assessment

Public Comment