

An Analysis of Historical Illinois Assessment Data to inform the use of Skip-Year SGP Analyses

Damian W. Betebenner & Adam VanIwaarden

The National Center for the Improvement of Educational Assessment

Dover, New Hampshire

June 2, 2020

Background

The COVID-19 global pandemic is having far-reaching effects in all facets of education in the United States. The impact on student education has been discussed extensively in general terms in the media relative to school closures and the suspension of statewide student assessment in 2020. States, including Illinois, who utilize student academic growth as a substantial portion of school accountability calculations are actively considering how to approach the calculation and use of student academic growth in 2021. In this brief, we discuss the current situation and present some preliminary student growth percentile (SGP) analyses for Illinois to help inform future decisions about possible uses of student growth in 2021 (Betebenner, 2008).

Possible Assessment Administration Scenarios

Depending upon the type of assessment, there are several ways in which the pandemic has impacted student assessment in Spring 2020. Here are some of the feasible assessment administration scenarios states are taking in the face of COVID-19 outbreaks across the nation.

- No testing
- Full (normal) testing (e.g., WIDA-ACCESS English Language Proficiency Testing was administered in several states the winter before school closures)
- Partial testing (test forms administered to a portion of students before school closures)
- Abbreviated testing (shortened test forms administered)
- Delayed “normal” testing (administered at the end of the delayed school year or beginning of next school year)

States administer several assessments each year, so it is possible for several of the above scenarios to play out in a single state. Within a single assessment, most state education agencies are mandating a uniform response for all schools in the state regardless of differing regional impacts of COVID-19, meaning each state or district will likely face only one of the possible scenarios for their primary assessment program. However, given that states have additional assessment programs (e.g. English Language Proficiency assessments) with differing administration windows, multiple scenarios may be relevant.

In Illinois, as in most other states, the most pressing concerns are with the state's large scale summative assessment given each spring and are twofold:

1. Can growth be calculated from Spring 2021 using Spring 2019 as the most immediate prior? This is often referred to as two-year growth.
2. If two-year growth can be calculated, can the quantities be used for accountability (e.g., school accountability) purposes?

The following report addresses these two questions using existing 2017, 2018, and 2019 SGP data from previous Illinois state assessment administrations. In general, the answer to the first question is straight forward: Yes, growth can be calculated from 2019 to 2021. Several states already use growth that spans two years due to a lack of testing in a given grade (e.g., growth from grade 8 to grade 10). Whether two-year growth can be utilized in place of one-year growth as part of accountability involves validating whether two year growth is comparable to one year as well as whether it is politically tenable to utilize two year growth in such a fashion.

Analyses

Illinois, as a member of what was formerly referred to as the PARCC consortium, has had SGPs calculated as part of the consortium since 2016. Illinois assesses students in grades 3 to 8 in English Language Arts (ELA) and Mathematics each spring. Due to the coronavirus pandemic of 2020 student assessment was called off so that no SGPs will be calculated in 2020. Because growth utilizes current and prior scores, the cancellation of testing in 2020 impacts the analysis of growth in 2021 and potentially later depending upon how many prior scores are used as part of the growth analysis.

To investigate the impact of the missing 2020 data on growth analyses in 2021, we used historical Illinois assessment data to calculate SGPs in two ways:

1. SGPs were calculated between 2017 and 2019 (two-year SGPs) using 2019 as the dependent variable and 2017 as the independent variables. SGPs of order 1 and 2 were calculated as part of these analyses.¹
2. SGPs were calculate between 2018 and 2019 (one-year SGPs) using 2019 as the dependent variable and 2018 and 2017 as the independent variable. SGPs of order 1 and 2 were calculated as part of these analyses.²

¹Order refers to the number of priors used in the analyses. Order 1 SGPs would utilize only 2017 as the prior score and order 2 SGPs would utilize both 2016 and 2017.

²Order refers to the number of priors used in the analyses. Order 1 SGPs would utilize only 2017 as the prior score and order 2 SGPs would utilize both 2016 and 2017.

Grade	ELA				Mathematics			
	SGP 2 YEAR	SGP 1 YEAR ORD 1	SGP 1 YEAR ORD 2	SGP 1 YEAR	SGP 2 YEAR	SGP 1 YEAR ORD 1	SGP 1 YEAR ORD 2	SGP 1 YEAR
All	541,363	689,217	533,925	689,217	541,236	688,203	532,786	688,203
4		135,833		135,833		135,167		135,167
5	135,827	139,260	134,307	139,260	136,019	139,286	134,329	139,286
6	137,429	140,613	135,752	140,613	137,494	140,559	135,665	140,559
7	134,271	137,744	132,338	137,744	134,271	137,674	132,304	137,674
8	133,836	136,567	131,528	136,567	133,452	135,517	130,488	135,517

Table 1: Frequencies of SGPs in 2019 by grade and content area and SGP type.

SGPs are calculated separately by grade and content area. One-year SGPs are calculated for grades 4, 5, 6, 7, and 8 in ELA and Mathematics. Two-year SGPs are calculated for grades 5, 6, 7, and 8.

Frequencies associated with SGP calculations are provided in Table 1. In general, there are between 135,000 and 140,000 students in each grade and content area with longitudinal data. Note that SGP 1 YEAR are the usual SGPs that are reported and represent the SGP calculated with the maximum number of priors given the students record. Given that grade 4 has only a single prior in the content area (i.e., grade 3), there are no order 2 SGPs or skip year SGPs.

Individual level results

Comparisons of individual level SGPs show, in general, high correlations depending on which version of 1 year SGPs one uses. Using the standard 1 year SGP which is the highest order SGP available for a student, the correlations (for each grade and content area) between one year SGP and two-year SGP are all between 0.85 and 0.9 with one exception at 0.81. Note, however, that when the 1st order, one-year SGP (i.e., 2018 to 2019 SGP using only 2018 as the prior) is correlated with the two-year SGP (2017 to 2019), the correlations drop to between 0.65 and 0.7.³ In general, the SGPs that include the 2nd prior yield higher correlations than those based solely upon the single, most proximal prior (2018). Almost all states use the SGP of highest order available for a given student.

High correlations do not always mean that the magnitude of differences between individual level SGPs is necessarily small. Table 2 provides .05/median/mean/.95 SGP SKIP YEAR - SGP differences. As the table shows, on “average” (median difference) there is no difference between the two-year and one-year SGPs. However, for 10 percent of the students the magnitude of difference exceeds 25 points. Even using broad categories to characterize individual SGPs (e.g., low, typical, high) will not ameliorate differences that large. Because differences are so large, at the individual level, we think it is unreasonable *at the individual level* to substitute a two-year SGP for a one-year SGP.

School level results

Illinois data provided by Pearson does not include demographic or school level information for students so we were not able to calculate school level summaries and compare school level

³Correlations between 1st order and 2nd order SGPs are approximately 0.95 for each grade and content area.

Grade	ELA			Mathematics		
	0.05	Median	0.95	0.05	Median	0.95
All	-25	0	26	-26	0	27
5	-24	0	25	-26	0	29
6	-25	0	26	-23	0	25
7	-27	0	27	-29	0	30
8	-25	0	26	-24	0	26

Table 2: Summary differences between two-year SGPs and one-year SGPs at the individual level.

Grade	ELA			Mathematics		
	0.05	Median	0.95	0.05	Median	0.95
All	-5	0	6	-6	0	5
Elementary	-4	0	5	-6	0	5
Middle	-5	0	4	-4	0	5

Table 3: Summary differences between two-year mean SGPs and one-year mean SGPs at the school level.

results based upon two-year SGPs versus school level results based upon one-year SGPs. We are currently working with another state on a similar set of analyses and can report preliminary data with regard to how two-year SGPs at the school level correlate and differ in magnitude there.

Correlations of mean two-year SGPs versus mean one-year SGPs exceed those at the individual level are 0.92 and 0.94 in ELA for elementary and middle schools, respectively, and 0.93 and 0.95 in mathematics. These correlations are very high and compare with correlations seen between different student growth models (e.g., VAM versus SGP) applied to the same data. Differences in magnitude between mean two-year SGPs and mean one-year SGPs are provided in Table 3

The observed differences between the different methods (i.e., two-year versus one-year) are less than observed year-to-year differences reported for accountability systems. One potential criteria to consider in the use of two-year SGPs in lieu of one-year SGPs is the extent to which they exceed year-to-year differences that are observed. That is, if mean SGPs are deemed accurate and reliable enough based upon year-to-year fluctuation, then they would be considered accurate and reliable enough based upon method-to-method fluctuation. We are currently investigating these comparisons more fully in the other state and hope to extend them to Illinois if we can merge data with school level indicators.

Summary

With the coronavirus interruption to education in the United States including the impact on student assessment, states are actively considering options for their school accountability plans in 2021. Forty-eight states in the United States currently utilize student academic growth as a part of their school accountability systems. Many states are investigating whether it is

possible to calculate growth and if so determine whether it is valid to use it as part of school accountability. For SGPs, preliminary results from Illinois and another state suggest that:

- At the individual level, two-year SGPs should not be substituted for one-year SGPs.
- At the school level, mean two-year SGPs demonstrate high correlation and minor differences with one-year SGP. This suggests it is feasible to use them as part of school accountability.

Determining whether to use mean two-year SGPs goes beyond just technical justification and extends to practical and political considerations. Two recommendations for further investigating this use case include:

1. Merge school level identifiers and demographic student identifiers with SGP data to create school level aggregates and calculate correlations and differences between mean two-year SGP and mean one-year SGPs.
2. Using historical accountability system data, investigate how replacing mean one-year SGPs with mean two-year SGPs changes accountability determinations for schools. Preliminary results indicate that changes will be extremely minor.

References

Betebenner, D. W. (2008, April). *Norm- and criterion-referenced student growth*. (Paper presented at the 2008 NCME Annual Conference, New York, NY. Available online at http://www.nciea.org/publications/normative_criterion_growth_DB08.pdf)