Student Growth Percentile in Star Assessments
Contents

4    Introduction
4    Growth
6    Student growth percentiles
7    About Renaissance Star Assessments
7    Applying SGP to Star Assessments
11   Reporting SGPs
12   Frequently asked questions
20   References

Figures
5    Figure 1. Growth is better understood when performance history and peer group are considered
10   Figure 2. Current SGP score: Decision rules for score selection in the SGP model
11   Figure 3. Sample Star Math Growth Report
Introduction

Student achievement typically is gleaned from one score at a single point in time. When you consider growth in addition to achievement, it can greatly enrich your understanding of how well a student is performing (Betebenner, 2009; Thurlow et al., 2010). While achievement indicates whether performance is below, above, or on par with grade-level expectations, growth explains the type of progress the student is making over time. For example, one of your students may be performing at a low level, yet experiencing high rates of growth. Conversely, a high-performing student's growth could be stagnating. It is important to know how a student is performing, but this information must have context—such as, how remarkable is this growth given a student's achievement history?

Many state accountability systems incorporate a plan for measuring growth over time, reflecting broad agreement that such systems must go beyond reporting the percentage of students obtaining proficiency status by the end of the school year (Domalesski & Perie, 2012). Let's examine the student growth percentile (SGP) score—a widely accepted indicator of student progress used by states for purposes including instructional decisions and accountability reports—that is used in the Renaissance Star Reading®, Renaissance Star Math®, Renaissance Star Early Literacy®, Renaissance Star Reading Spanish®, Renaissance Star Math Spanish®, and Renaissance Star Early Literacy Spanish® assessments.

Growth

Growth over time, which is sometimes called slope or rate of improvement, is of central importance in evidence-based instructional models such as Response to Intervention and Multi-Tiered Systems of Support. When you can capture and accurately interpret growth information, you are able to make informed, data-based decisions regarding the extent to which students are benefiting from intervention or regular classroom instruction, or whether changes are warranted (Fox et al., 2009).

Figure 1 illustrates why we need context to fully understand differing rates of growth by depicting the performance of two high jumpers.
Figure 1. Growth is better understood when performance history and peer group are considered

Over a four-month period, Athlete A increased her high jump by 4 inches, while Athlete B increased his jump by 1 inch. How should we interpret these gains? Whose growth was more impressive? At first glance, it may seem Athlete A has made greater improvement; however, to determine the significance of the athletes’ achievements, we must also consider their performance history and peer groups.

Athlete A is a novice who had room for improvement, while Athlete B is an Olympian who, even while performing at his peak, was able to improve. Having background information helps us know that the growth achieved by the expert Olympian was more impressive than the novice’s improvement. Without data about the growth that would be expected for each type of athlete, it is difficult to draw these conclusions.

Likewise, in education, knowing absolute change in achievement—in scaled score, for example—is not helpful for making meaning from data. Without context, we do not know if the growth was expected, below what was expected, or extraordinary. The amount each student grows can vary by test/subject, grade, and prior achievement, so simply knowing that a student’s scores increased is only half the story.

Many statistical models have been designed to measure student growth; Castellano and Ho (2013a) provide an overview of seven such models. One of the most widely used is student growth percentile, which was developed by Dr. Damian Betebenner of the National Center for the Improvement of Educational Assessment and piloted in partnership with various state departments of education (Betebenner et al., 2016). SGPs have now been adopted by several states for instructional and accountability purposes.

Star Early Literacy, Star Reading, and Star Math were the first interim tests to report SGPs. Growth models like SGP require an enormous amount of data to generate reliable results (Castellano & Ho, 2013a). Fortunately, widespread national use of Renaissance assessments provides a wealth of data,
allowing SGPs to be reported for nearly all students in every grade, no matter how high or low their initial achievement level.

Student growth percentiles

SGPs are a norm-referenced quantification of individual student growth derived using quantile regression techniques (Betebenner, 2011). The SGP score compares a student’s growth from one period to another with that of his or her academic peers nationwide—defined as students in the same grade with a similar scaled score history. SGPs range from 1–99 and interpretation is similar to percentile rank (PR) scores: lower numbers indicate lower relative growth and higher numbers indicate higher relative growth. For example, an SGP of 75 means the student’s growth exceeds the growth of 75 percent of students with a similar score history.

SGPs help us understand, given where a student started, to what extent the growth achieved was as expected. Without an SGP, a teacher may not know if a scaled score increase of 100 is good, not-so-good, or average, because what is considered expected growth for one student may not be for another. An SGP of 50 can be thought of as typical growth for a student, given his/her grade and prior score history; however, state and local policy makers may define typical growth as a less precise range, such as 35 to 65 or 40 to 60 SGP.

SGPs can be aggregated to describe growth for groups of students—such as for a whole class, grade, or school—by calculating the group’s mean or median (middle) growth percentile. No matter how SGPs are aggregated, the statistic and its interpretation remain the same. For example, a median SGP of 62 for a class means the middle student in that group achieved higher growth than 62 percent of his or her academic peers.

Sample characteristics

Each year, approximately 60 to 70 million Star tests are taken, and nearly all these scores are included in the sample used to report SGPs. The SGP data used in the model is refreshed periodically, and while we do not report SGPs for specific subgroups of students, all students—regardless of special education or English learner status—are retained in the sample.

Distribution of scores

A common misunderstanding regarding SGP scores is that their statistical distribution is normal, like a bell curve. This would mean that there are more SGPs reported in the middle (near 50) than at the tails, near 1 and 99. This is not true. While it is possible for SGP scores at local (e.g., class) levels to have any type of distribution, nationally the distribution is approximately flat for all grades and subjects. Thus, within any subject/grade, the number of scores reported at every point between 1 and 99 will be about the same (each score is reported for about 1 percent of students). There will be approximately the same

---

1 Scores for Star Early Literacy are included in the model through third grade. SGPs for high school are not yet available for Star Math Spanish and Star Reading Spanish.

2 In rare instances, for some grade and testing window combinations in the Spanish assessments, data may have been pooled across nearby grades in order to increase sample sizes.
number of students with an SGP of 50 and 6 and 92 and 37, and so on. Because of this uniform distribution, all students, regardless of score history, have as good a chance of demonstrating high growth as low growth (i.e., scoring at any of the 99 SGPs). See page 18 for FAQs covering goal setting and what to expect when examining distributions of SGPs at student, class, and school levels.

It is important to remember that no matter how high, low, or middle-of-the-road a student’s PR score, the student has an equal chance of receiving any SGP score from 1–99. Take, for example, a student with a fall percentile rank of 95 who receives an SGP of 19 at the end of the year. It may not seem reasonable that such a high-performing student would receive a relatively low growth score, but this just means that 81 percent of the student’s academic peers from the same grade with a similar score history experienced more growth. SGP compares each student’s performance to that of a group of unique academic peers, which is recalculated each time the student takes an assessment. No assumptions can or should be made about a student’s SGP based on PR performance. (Note, although we reference PR scores to illustrate points about achievement and growth, PRs are not used in the SGP calculation.)

**About Renaissance Star Assessments**

Renaissance Star Assessments® are reliable, valid, and time-efficient assessments of early literacy (Star Early Literacy and Star Early Literacy Spanish), reading (Star Reading and Star Reading Spanish), and mathematics (Star Math and Star Math Spanish) skills. Quick and accurate results from these assessments provide specific benchmarking, screening, progress-monitoring, and diagnostic information to help you tailor instruction, monitor growth, and improve achievement for all students. Renaissance Star Reading and Renaissance Star Math are highly rated for academic screening and academic progress monitoring by the National Center on Intensive Intervention (NCII, 2018a, 2018b, 2019a, 2019b). For more information on the reliability, validity, and other technical aspects of Star Assessments, see the Star technical manuals, available by request to research@renaissance.com.

**Applying SGP to Star Assessments**

In the 2011–2012 school year, we first reported SGPs in Star Reading and Star Math for grades 1–12 and in Star Early Literacy for grades K–3. Updates to the model now allow students who are transitioning to independent reading to receive an SGP to show growth between the Star Early Literacy and Star Reading assessments (for example, if a student began the year taking Star Early Literacy, but ended the year taking Star Reading). As of the 2020–2021 school year, SGPs are reported in Star Reading Spanish and Star Math Spanish for grades 1–8 and in Star Early Literacy Spanish for grades K–3.

From the very beginning, Renaissance researchers have worked closely with SGP creator Dr. Betebenner to apply the SGP approach to Star Assessment data.

**Testing windows**

Because SGP was initially developed for measuring growth on state tests across school years, applying the SGP approach to interim assessment data involved several technical challenges, primarily regarding differences in the timing of administering Star tests versus state tests.
State summative tests are typically given once per year, at approximately the same time, to nearly all students. Thus, score comparisons from one state test administration to another speak to growth across school years. Consequently, the original SGP model developed by Dr. Betebenner for state use assumed fairly constrained administration parameters with approximately the same amount of time between tests. In stark contrast, Star Assessments can be considered “on-demand” tests with far more flexibility. Administration decisions (when and to which students) are left to local educators based on purposes and needs for assessment. Most commonly, schools choose to use Star Assessments as screening or benchmarking tests for all, or nearly all, students 2–4 times per year. Students requiring progress monitoring may take the assessments more frequently to inform instructional decisions, such as whether a student is responding adequately to an intervention. Teachers who administer the Star Early Literacy assessment to beginning readers at the start of the school year and then have students test in Star Reading as they transition to reading on their own can track students’ progress with SGP scores derived using scores from both assessments.

Given that not all students take Star Assessments at the same time, and that the number and dates of test administrations may vary from one student to the next, it was necessary to make two adaptations for Star SGP: (1) identify testing windows and, (2) adjust for variable time between tests. Our analysis of Star data revealed a clear pattern for the majority of tests taken during the school year, which corresponded closely with the timing of district screening or benchmarking: Fall (August 1–November 30), Winter (December 1–March 31), and Spring (April 1–July 31).

While defining the data sets used to determine SGPs, these common date ranges became our testing windows. Establishing testing windows allowed Star SGPs to be reported within-year in a manner consistent with most district testing calendars.

**Calculating SGPs**

Quantile regression is a statistical process used in SGP models to estimate the conditional distribution of an outcome variable (a test score) given prior information (a student’s prior scores). An SGP reflects the likelihood of a specific outcome (an amount of growth over a period of time) given a student’s prior score history, while using data available from all students from recent years that characterize how different students grow. In general, this method can be viewed as a type of smoothing, in which information from neighboring score values can be used to inform percentiles for hypothetical score combinations not yet observed (Betebenner, 2016).

The SGP model prioritizes available data points to make the best use of information across time, by factoring a student’s current test score (the posttest) and up to two prior test scores (the pretest and, if available, an additional prior test may be used) into the calculation:

- **Posttest**: A score from the most recent test taken within the last 18 months.
- **Pretest**: A score from a test in an SGP window prior to the window the posttest falls within.
- **Additional prior test**: A score, if available, from a window in the previous school year may be used. English SGPs always use additional priors when they are available, but Spanish SGPs may not use an additional prior score in all subject/grade/season instances, even if it is available. This difference is due to the relative newness of the Spanish measures. Empirical evidence shows that using a student’s prior-year score, when available, ensures the most accurate representation of growth within an academic year (Betebenner, 2016).
The purpose of the additional prior score

Using an additional prior score, when available, in the SGP calculation serves to stabilize a student’s pretest score, minimize the impact of measurement error, and provide valuable baseline information to ensure the most accurate picture of the student’s optimal academic peer group.

For example, suppose two students have very similar posttest (current) and pretest scores. You might expect their SGP scores to also be similar. Their SGPs may turn out to be the same or close, but only looking at similar growth between a posttest and pretest does not provide a complete picture of the students’ growth. Adding an additional prior score to the calculation provides needed context, such as that the timing of their prior test events may have differed, giving each student varying levels of exposure to skills and learning time. Or, one student’s additional prior score might have been higher than his/her pretest score, while the other student’s prior score might have been much lower, meaning their academic peer groups were different. Using an additional data point provides a more accurate picture of each students’ individual score histories.

Score selection in the SGP model

Each time a student takes a Star Assessment, he/she receives a current SGP score. The score is reported based on the available Star test score history for that student. Figure 2 shows the decision rules that guide how an official SGP score is reported. The type of SGP score a student receives (i.e., the period of growth it covers) is prioritized from top to bottom in the table, depending on available data. As shown, when more than one test has been taken in an SGP window, the model selects the following prior scores: the first test taken in fall, the test taken closest to January 15 in winter, and the last test taken in spring. See the Frequently asked questions for details about Window Specific SGP.

---

3 Standard error of measurement (SEM) is unavoidable and is present to some degree in all assessments. Assessment developers can seek only to minimize the impact of SEM. Tests with sound technical characteristics, such as Star Assessments, should reliably generate consistent and accurate estimates of student achievement.
Adjusting for time

To provide the best possible indication of how a student is growing, we conduct ongoing research to determine what adjustments to the model will ensure the most fair and accurate comparisons of the data. Thus, based on a study that validated these adjustments, the Star SGP model has evolved to use time in two ways: 4

1. The amount of days between the posttest and the pretest. The testing windows alone do not address that students in the same window may have spans of time between tests that vary greatly—and, consequently, different opportunities to learn and grow. For instance, a student with tests taken the first day of the fall window and the last day of the spring window would have 364 days between test events, while another student testing the last day of the fall window and the first day of the spring window would have 122 days between tests. The more days between two testing events, the more growth that can be expected.

2. At what point in the window a student took the current test (which indicates how close or far the student is from the start of the testing window). Students at the end of the testing window have had more exposure to content and, thus, their scaled scores are likely higher.

4 For more on the time-sensitive calculation implemented in the SGP model, see the technical paper by Betebenner (2016).
**Reporting SGPs**

The SGP model provides an SGP for every student at the start the school year, after the first fall test is taken (as long as data exists from the previous school year). For example, if a student has a Star score from the prior spring and another test score from fall of the current school year (from the same assessment—e.g., Star Math), an SGP will generate indicating spring to fall growth. Having SGP scores available in fall allows educators to begin the year understanding students’ recent growth history, which can provide immediate insight and assist with initial instructional decisions. As the year progresses and additional assessments are taken, Star Assessments report students’ current (most recent) SGPs.

The Star Growth Report (see Figure 3) summarizes growth between two testing periods in the same school year as soon as a student has both a pretest and posttest score. Teachers can run the report for a class or specific groups of students by SGP window, and administrators can see growth for each class or grade in their schools.

**Figure 3. Sample Star Math Growth Report**

![Sample Star Math Growth Report](image)
Frequently asked questions

What is a student growth percentile (SGP)?
A student growth percentile, or SGP, compares a student’s growth to that of his or her academic peers nationwide. Academic peers are students in the same grade with similar achievement history on Star Assessments. SGP is reported on a 1–99 scale, with lower numbers indicating lower relative growth and higher numbers indicating higher relative growth. For example, an SGP score of 90 means the student has shown more growth than 90 percent of his/her academic peers.

How are SGPs determined?
SGPs are based upon the best available score information using a statistical model of growth and achievement called quantile regression. The model is designed to prioritize data points to make the best use of data across time. The calculation uses test scores from at least two SGP windows, and a third SGP window may be used when available:

- **Posttest**: A score from the most recent test taken within the last 18 months.
- **Pretest**: A score from a test in an SGP window prior to the window the posttest falls within.
- **Additional prior test**: A score, if available, from a window in the previous school year may be used. English SGPs always use additional priors when they are available, but Spanish SGPs may not use an additional prior score in all subject/grade/season instances, even if it is available. This difference is due to the relative newness of the Spanish measures, and the fact that Renaissance did not always have sufficient historical data to use additional prior scores. We do intend to incorporate them in the future. Empirical evidence (Betebenner, 2016) shows that using a student’s prior-year score, when available, helps to pinpoint a student’s optimal academic peer group and ensure the most accurate representation of growth within an academic year.

- **SGP windows**
  - Fall (August 1–November 30)
  - Winter (December 1–March 31)
  - Spring (April 1–July 31)
    - When more than one test has been taken in an SGP window, the model selects the following prior scores:
      - Fall: first test taken
      - Winter: test closest to January 15
      - Spring: last test taken

Can I get a SGP for the Star Spanish Assessments?
Yes, SGPs are offered for Spanish versions of Star Assessments in grades 1–8 in Star Math Spanish and Star Reading Spanish and grades K–3 in Star Early Literacy Spanish. SGPs are not yet available in grades 9–12 for Star Math Spanish and Star Reading Spanish.

Can SGP be used with English learners or students receiving special education services?
Yes, the SGP norming sample includes students categorized as English learners and as participating in special education. However, much remains to be learned about how these students grow and whether it is reasonable to expect their growth to be comparable to other students.

How do I know which test scores were used to determine the SGP score?
The Growth Report, see p. 11, displays the scores used to determine a student’s SGP score.
I tested a student in the fall, winter, and spring. Why is the previous year’s test score used to determine the SGP score?
SGP creator Damian Betebenner’s ongoing research (2016) has shown convincing evidence that by improving the association of a student’s scores with those of his/her peers (by considering an additional prior score for the student from the previous school year), the SGP model provides as complete a picture of individual student growth as possible. Including this additional data point helps to better pinpoint a student’s optimal academic peer group, which results in a more accurate measurement of growth between fall and spring.

How do I get an SGP for an earlier time in the year or a prior school year?
When running a Growth Report, you can select the testing window to view your desired SGP score.

Can I obtain an SGP for students who start the school year taking Star Early Literacy and transition to Star Reading during the same school year?
Yes! As long as a student has taken a Star Early Literacy assessment in a different SGP testing window than a Star Reading assessment, AND he/she is not already receiving an SGP based solely on Star Early Literacy OR Star Reading tests.

Can SGP scores be compared from year to year?
Because of the important role SGP scores play in instructional and accountability decisions, both Renaissance and SGP creator Dr. Betebenner are committed to a continuous improvement cycle. Since first reporting SGPs during the 2011–2012 school year, yearly refinements have improved the model’s functionality and accuracy. These changes impact neither the interpretation nor the general distribution of the SGP score. All students, whether low, average, or high performing, always have an equal opportunity to achieve any of the 99 SGP scores.

Thus, although adjustments are made to the SGP model each year, yes, scores can be compared over time. In 2016, Renaissance researchers studied the comparability of the scores by examining SGP scores from a large set of Star student records from the 2014–2015 school year. The scores were run through two SGP models, the model from the 2014–2015 school year and the enhanced model used in 2015–2016.

Results revealed high average correlations in the mid .9s, with a range of coefficients from .82 to .99 when looking at specific grade/subject combinations. Overall, the analysis showed that although recent changes provided meaningful improvement in the accuracy of the SGP score, both model calculations sort students consistently and provide reliable estimates of student growth. (Results did not perfectly correlate, which would call into question the efficacy of model enhancements producing precisely the same results.)

Why might SGPs for the same student vary across time or between different assessments?
You may expect to see very consistent SGPs for a given student within or across school years, but this is highly unlikely for several reasons: changes in instruction, the school environment, and the student’s aptitude, as well as the impact of measurement error (common in all educational tests), all can explain why a student does not receive the same SGP repeatedly.

In the case of varying SGPs on different assessments, consider expert recommendations (e.g., Hamilton et al., 2009) regarding how to use multiple sources of data to inform instructional decisions. Although Star SGP is a robust growth measure on its own, it should be used in combination with other reliable and valid sources of information about student achievement and growth.

Does whether a test is adaptive versus fixed form make a difference in the interpretation of growth with the SGP score?
SGP scores have the same basic interpretation regardless of test type. The SGP calculation for Star Assessments (which are adaptive) was developed by Dr. Damian Betebenner, who also worked to report SGPs with numerous
state summative assessment programs (which are mostly fixed form). Interpretation of SGP is the same across grade, subject, and assessment, assuming Dr. Betebenner’s model is followed.

One consideration when interpreting SGPs is that the timeframe of growth characterized by the SGP score may differ between Star Assessments and summative assessments. Star Assessments can be given throughout the year, whereas state summative tests are usually administered once per year, often in the spring. Thus, Star SGPs may reflect growth from any number of time periods (e.g., Fall to Spring, Spring to Fall, Fall to Winter), while state summative SGPs generally reflect Spring to Spring growth.

I have heard it can be difficult to differentiate growth in older (e.g., high school) student populations. Does that make SGP less useful or valid for secondary students?

No, the process to calculate SGP and the interpretation of its scores is consistent across all grades. It is well documented that growth over time is not perfectly linear—rather, rate of achievement growth tends to slow as students get older. This phenomenon, however, does not impact the interpretation of SGP nor threaten its validity. Although high school students tend to grow less than elementary students, in absolute scaled score change from beginning to end of year, there is enough variability in growth in the upper grades for a model such as SGP to work as well as it does in earlier grades. The SGP calculation takes into consideration grade level, meaning students are compared only with academic peers within their current grade. SGPs for high school are not yet available for Star Math Spanish and Star Reading Spanish.

Where are SGPs reported?

Star Assessments report a student’s SGP score in the Growth Report.

What does the dash mean on the Growth Report?

There are certain circumstances when a dash (–) will appear in lieu of an SGP on the Growth Report:

- A kindergarten student taking only Star Reading or Star Math assessments will not receive SGP scores, as these tests were designed for grades 1–12. (SGPs are reported for kindergarten students who take Star Early Literacy assessments alone or in combination with a Star Reading assessment.)

- For a student to receive an SGP based on prior-year spring to current-year fall growth, advancement of one grade is necessary between test administrations. (In addition, mid-year promotion or demotion of grades may result in a SGP score not being reported.) This also applies to other cross-year SGPs for Fall-Fall and Spring-Spring time periods.

- Student data populates with a minimum of one assessment in at least two consecutive SGP windows (from the same assessment, e.g., two test events in Star Reading—except in the case of transitioning readers who may begin the year testing in Star Early Literacy but then the year testing in Star Reading) must be taken in order to calculate an SGP score. Thus, if the report is being viewed the same day the test was taken, an SGP score may not be visible.

- Window Specific SGP score data is reported based on the decision rules displayed in the table (below), with the type of score received prioritized from top to bottom, depending on available data. Score history that deviates from what is shown here may result in the program being unable to report an SGP score.

- Window Specific SGPs are determined (during report customization) by selecting a current or prior school year as well as the SGP window time frames that you want to view growth for. The tests used to determine Window Specific SGPs are based on the following rules:
  - Fall Window: First test taken
  - Winter Window: Test taken closest to 1/15
Spring Window: Last test taken

In the Star Growth Report, Window Specific SGPs are calculated when you select a prior or current school year during the report customization (in the **Timeframe** drop-down list).

**Window Specific SGP score: Rules for score selection in the SGP model**

<table>
<thead>
<tr>
<th>Type of Window Specific SGP Calculated</th>
<th>Test Windows in Previous School Year*</th>
<th>Test Windows in Selected School Year*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fall 8/1-11/30</td>
<td>Fall 8/1-11/30</td>
</tr>
<tr>
<td></td>
<td>Winter 12/1-3/31</td>
<td>Winter 12/1-3/31</td>
</tr>
<tr>
<td></td>
<td>Spring 4/1-7/31</td>
<td>Spring 4/1-7/31</td>
</tr>
<tr>
<td>Fall–Fall</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spring–Fall</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fall–Winter</td>
<td><img src="image.png" alt="Diagram" /></td>
<td></td>
</tr>
<tr>
<td>Spring–Spring</td>
<td><img src="image.png" alt="Diagram" /></td>
<td></td>
</tr>
<tr>
<td>Fall–Spring</td>
<td><img src="image.png" alt="Diagram" /></td>
<td></td>
</tr>
<tr>
<td>Winter–Spring</td>
<td><img src="image.png" alt="Diagram" /></td>
<td></td>
</tr>
</tbody>
</table>

* Test window dates are fixed, and may not correspond to the beginning/ending dates of your school year. Students will only have Window Specific SGPs calculated if they have taken at least two tests, and the date of the most recent test has to be within the past 18 months.

The Window Specific SGP is calculated using at least two tests: the most-recently taken test in the past 18 months, and one test from a prior test window. If more than one test was taken in a test window, the test used for the calculations will be the following:

- ![Diagram](image.png) Two tests used to calculate SGP
- ![Diagram](image.png) Test in window, but skipped when calculating SGP
- ![Diagram](image.png) Third test used to calculate SGP (if available)

<table>
<thead>
<tr>
<th>Test Window is</th>
<th>The test used for Window Specific SGP Calculation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall Window</td>
<td>First test taken in the window</td>
</tr>
<tr>
<td>Winter Window</td>
<td>Test taken closest to 1/15 (red line)</td>
</tr>
<tr>
<td>Spring Window</td>
<td>Last test taken in the window</td>
</tr>
</tbody>
</table>
Why do I see an SGP in the fall?
Improvements to the SGP model allow teachers to start the school year with an SGP for all students (who have data from the previous school year) after they take a test in the current Fall window. For example, if the student has a Star score from the prior spring and takes their first test in Fall, an SGP will generate indicating spring-to-fall growth. This feature provides teachers with immediate information about students’ growth history to provide insight and assist with initial instructional decisions.

How do Star Assessments taken in the summer months affect a student’s fall SGP score?
If you test students with Star Assessments during summer school, and that testing takes place on or after August 1, SGP projections and calculations for the new school year may be based on that summer test, rather than the first test of the new school year. To remedy this, we recommend the following options:

Option 1: Do not use Star Assessments after August 1, until the first screening window of the new school year begins. This will ensure that (1) the first test of the year is linked to the student in their new grade, (2) SGP calculations will anchor to this test, and (3) SGP growth projections are used correctly in multiple reports/functions.

Or if you prefer, Option 2 is to: Start your new school year in Star on August 1st and promote students at that point. Keep in mind that the first test of this new school year will be used as the Fall test for both SGP calculation and growth projections on Star reports.

How can my student have a high Star Percentile Rank and a low Star SGP? Is SGP accurate for high-achieving kids?
With SGP, all students, no matter their score history, have an equal chance to demonstrate growth at any of the 99 percentiles. High-achieving students are compared against a national sample of other high-achieving students with similar achievement history (i.e., their academic peers). Thus, it is possible for a student who is scoring well-above average to have an SGP that is relatively low, typical, or relatively high.

For example, a student with a fall PR of 95 receives an SGP of 19 at the end of the year. It may not seem reasonable that such a high-performing student would receive a relatively low growth score, but this SGP indicates that 81 percent of the student’s academic peers (from the same grade with a similar scaled score (SS) history) experienced more growth. No matter how high, low, or middle-of-the-road a PR score, each student has the same chance of receiving any SGP score from 1–99. (Although we reference PR scores to illustrate points about achievement and growth, PRs are not used in the SGP calculation. No assumptions can be made about a student’s SGP based on PR performance.)

What is the difference between PR and SGP?
Percentile rank (PR) and student growth percentile (SGP) are very different metrics (see table, next page). PR is an achievement (i.e., performance) score that describes a single point in time, while SGP is a growth measure that explains student growth between points in time. Both measures are norm-referenced but have different norming groups—for PR, this group includes all students in a particular grade level; for SGP, each student has a unique academic peer group.

Because both scales span 1–99, score interpretation is similar between PR and SGP: lower numbers indicate lower relative achievement/growth and higher numbers indicate higher relative achievement/growth (e.g., an SGP of 75 means the student’s growth exceeds that of 75 percent of students, just as a PR of 75 means a student’s achievement exceeds that of 75 percent of students).

Although these scores can be interpreted similarly, that does not mean a student with a high PR score will receive a high SGP score. A high PR score means a student is performing well at a certain point in time, but no matter how high or low a student’s PR, there is an equal chance for that same student to receive any SGP score from
Because SGP compares a student’s performance to students in the same grade with a similar scaled score history, this unique group of academic peers is recalculated each time an assessment is taken. No assumptions can be made about a student’s SGP based on PR performance, and PRs are not used in the SGP calculation.

### Percentile rank (PR) vs. Student growth percentile (SGP)

<table>
<thead>
<tr>
<th>Percentile rank (PR)</th>
<th>Student growth percentile (SGP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Based on a 1–99 scale</td>
<td>Based on a 1–99 scale</td>
</tr>
<tr>
<td>Performance score</td>
<td>Growth score</td>
</tr>
<tr>
<td>PR reported after one test</td>
<td>At least two tests are needed to report an SGP</td>
</tr>
<tr>
<td>Describes a student’s achievement at a single point in time</td>
<td>Measures a student’s growth</td>
</tr>
<tr>
<td>Norm-referenced—compares students in the same grade</td>
<td>Norm-referenced—compares students in the same grade with similar achievement history</td>
</tr>
<tr>
<td>Scaled score is compared to national norm group of grade-level peers</td>
<td>Scaled scores are compared to national norm group of grade-level academic peers</td>
</tr>
</tbody>
</table>

**What does it mean when a student has a high PR and a low SGP?**

It is important to look at both achievement (percentile rank/scaled score (SS)) and growth (SGP). Achievement scores, like PR and/or SS, tell us how students are performing at a single point in time, but this data is only a piece of the puzzle. It is also important to know how students fare over time in relation to their peers, a question comparative growth data from SGP can help answer. For example, consider a student who begins and ends the year at the 50th PR with an SGP of 30. This student has consistently performed better than 50 percent of students in the same grade nationwide; however, the SGP score shows that the student has grown more than only 30 percent of students in the same grade with a similar score history. For more information, see FAQ: *What is the difference between PR and SGP?* above.

**How are individual student SGP scores distributed nationally?**

While it is possible for SGP scores at local (e.g., class) levels to have any type of distribution, nationally the distribution is approximately flat for all grades and subjects. Within any subject/grade, the number of reported scores at every point between 1 and 99 will be about the same (each score is reported for about 1 percent of students), so that there will be approximately the same number of students with an SGP of 50 as 6 as 92 as 37, and so on.

It is a common misunderstanding that the statistical distribution of SGP scores is normal, like a bell curve, meaning there are more SGPs reported in the middle (near 50) than at the tails, near 1 and near 99. *This is not true.* See the next question for more information about typical distributions of SGP scores at student, class, and school levels, which has implications for setting SGP goals at those levels.

**How should I aggregate SGP scores? Mean or median?**

In keeping with the many states that report SGPs on summative tests, Renaissance also reports median SGP. We recognize, however, recent research on this topic (Castellano & Ho, 2013b) concludes it may be appropriate to use mean or median. You may want to consult your state’s position and use the preferred statistic. All educators should exercise caution when aggregating SGP results for small classes/groups (fewer than 20 students), as both
mean and median may provide misleading estimates of central tendency, depending on distribution of scores and group size. It is for this reason that some states do not report SGP results for small groups.

**What should I consider when interpreting SGP results or setting SGP goals?**

Here are some considerations for setting goals or otherwise evaluating student growth either at individual or group levels.

- No matter the subject, grade, time of year, or whether you are considering an individual student or a group of students, SGP 50 always indicates typical or expected growth.

- That said, educators may wish to consider ranges of SGP values around 50 as indicating typical growth, rather than 50 by itself, given measurement error. In other words, the relative growth demonstrated by a student with SGP 49 may not be appreciably different than a student with SGP 50 or 51. That is why many districts and states who use SGP encourage the use of ranges around 50 to define typical growth. The most common range indicating typical growth at a student level is 35–65. Therefore, 1–34 would indicate below typical growth and 66–99 would indicate above typical growth. Learn more about how states use data from this metric here: [http://doc.renlearn.com/KMNet/R00585975038A824.pdf](http://doc.renlearn.com/KMNet/R00585975038A824.pdf).

When considering distributions of individual student SGPs, because the range 35–65 encompasses 31 points on the SGP scale, and because the scale is uniformly (not normally) distributed, we would expect roughly 31 percent of students nationally to have an SGP between 35–65 at any subject, grade, or time frame. Likewise, about 34 percent would be between 1–34 and about 34 percent would be between 66 and 99.

However, when SGP scores are aggregated by taking median or mean of SGPs within a class or school, those expectations no longer hold. The more student SGPs that are included in the group average or median, the more likely it is that those will bunch up around 50. In other words, it is not true that about 34% of class median SGPs will be below 35.

The following table summarizes approximate national distributions of scores at student, class, and school levels.

<table>
<thead>
<tr>
<th></th>
<th>Approximate Star SGP Score Ranges</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bottom third</td>
</tr>
<tr>
<td>Individual Student</td>
<td>1–34</td>
</tr>
<tr>
<td>Class Median</td>
<td>1–39</td>
</tr>
<tr>
<td>School Median</td>
<td>1–44</td>
</tr>
</tbody>
</table>

Following are notes to assist with interpretation of the table:

- Nationally, roughly 1/3 of individual students will have SGPs between 1–34, about 1/3 will be between 35–65, and the other 1/3 will be above 66.

- Nationally, approximately 1/3 of class median SGPs will fall below 40. If your class median Star SGP is 61, your class growth was better than about 2/3 of all the classes in the US.

- Nationally, approximately 1/3 of school median SGPs will fall below 45. So, if your school median Star SGP was 44, then roughly 2/3s of the schools nationally had stronger growth.

When reviewing data and considering SGP goals for students or groups of students, educators may wish to focus on supporting those students in bottom third, those with SGPs below 1 and 34. Of particular concern are those students with a history of relatively low growth; consecutive periods of low growth can result in their falling farther
behind their peers. Providing additional supports to those students should boost the proportion of students in the middle or upper third and the overall class or school medians.

Was any student data from the COVID-19 pandemic era used to create SGPs?
No. At the time of this update (June 2021), no data from spring 2020 or the 2020–2021 school year has been used to create Star SGP growth norms. In other words, the SGPs generated in Star reports, dashboards, and extracts describe growth expectations prior to the pandemic. To help educators and other stakeholders understand what performance and growth looks like during this timeframe, Renaissance has released a series of reports called How Kids Are Performing, available here: https://www.renaissance.com/how-kids-are-performing/

How do I get SGP scores based only on students in my state?
Growth models like SGP require an enormous amount of data to generate reliable results (Castellano & Ho, 2013a). We examine data for students nationwide so that our sample size is adequate to calculate reliable and valid SGPs that compare students accurately to their academic peers throughout the U.S.

In states that have adopted SGP, how will a student’s SGP from the state test compare to a Star SGP?
A student’s SGP on any assessment can vary from a Star SGP for many reasons: differences in test content, blueprint, and delivery, as well as the amount of time between test administration and the norming groups used. Consider expert recommendations (e.g., Hamilton et al., 2009) regarding how to use multiple sources of data to inform instructional decisions. Although Star SGP is a robust growth measure on its own, it should be used in combination with other reliable and valid sources of information about student achievement and growth.

Are there other ways besides SGP to understand student growth?
Yes, there are many ways to understand student growth. Castellano and Ho (2013a) provide an exhaustive list of methods. One approach is to calculate the change in a normative score such as a normal curve equivalent (NCE). NCEs provide a way of representing PR scores so they can be accurately averaged and compared. Because NCEs are derived from percentiles, they measure growth in comparison to national norms. Positive NCE change means student achievement grew at a faster rate than the national average (an NCE gain of zero). Another widely used model is value-added. Scores from Star Assessments can be used in such models to illustrate growth.
References


Independent technical reviews of Star Assessments


Adviser

Dr. Damian Betebenner is senior associate at the National Center for the Improvement of Educational Assessment in Dover, New Hampshire. He specializes in applied statistics, and his current research focuses on longitudinal data analysis—specifically with regard to state and federal performance mandates.