

Student Growth Percentile in STAR™ Assessments

This document provides a technical summary of STAR Student Growth Percentile (SGP), specifically, the origins of SGP as an indicator of growth on summative state assessments, as well as how and why STAR SGP was developed, the sample upon which it is based, and how it is used by educators for various purposes.

Introduction

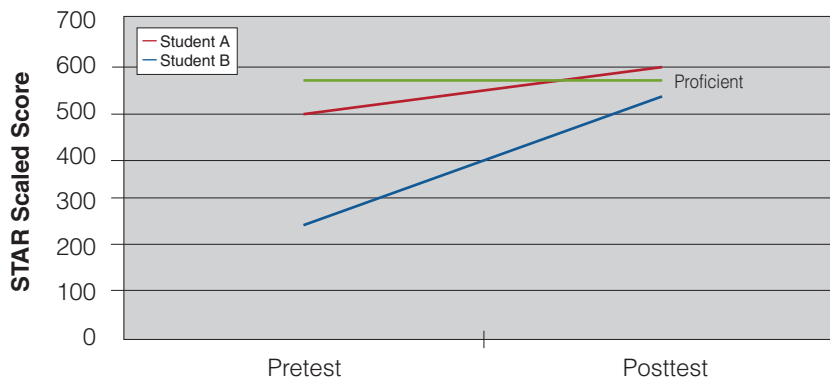
Considering growth in addition to achievement status greatly enriches an educator’s understanding of how well a student is performing (Betebenner, 2009). Student achievement is typically gleaned from a single score from a single point in time. While knowing a student’s level of *achievement* indicates whether the student is performing below, above, or on grade-level expectations, *growth* indicates what kind of progress the student is making. For example, a student may be performing at a low level, yet experiencing high rates of growth. Conversely, a high-performing student could be stagnating.

Growth, sometimes called slope or rate of improvement, has been of central importance in evidence-based instructional models such as Response to Intervention and Multi-Tiered Systems of Support for some time. When educators are able to capture and accurately interpret growth information, they can make informed, data-based decisions regarding the extent to which students are profiting from intervention or regular classroom instruction, or whether changes to instruction are warranted (Fox, Carta, Strain, Dunlap, & Hemmeter, 2009).

In recent years, federal and state educational policy has come to more formally accept and embrace growth in state accountability systems, reflecting broad agreement that such systems must go beyond the percentage of students obtaining proficiency status by the end of the school year (Domaleski & Perie, 2012).

Figure 1 highlights the importance of understanding growth by depicting the performance of two students, A and B, on a test administered twice during a school year. Student A ends the school year with a score of 600 which places him over the proficiency threshold (the green line). Student B does not reach the proficiency threshold by the end of the school year, but the student’s growth from pretest to posttest is greater than the growth shown by student A.

Figure 1: Illustration of Student Achievement Versus Growth



Measurement, however, is just the first step. Understanding that Student B grew more than A does not necessarily help an educator know if either grew as much as would be expected given their grade and starting points. To answer that question, normative data on student growth is needed, and that is where Student Growth Percentiles can be helpful.

Student Growth Percentiles

A number of 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, 2010). Adopted by a number of states for instructional and accountability purposes, Renaissance Learning's STAR assessments for Early Literacy, Reading, and Mathematics became the first interim tests to adopt the SGP approach. Below we provide a general description of SGPs and then explain how this approach, which had most commonly been used to characterize growth between annual state summative measures, was adapted for within-year purposes via the STAR measures.

SGPs are a norm-referenced quantification of individual student growth derived using quantile regression techniques (Betebenner, 2011). An SGP compares a student's growth to that of his or her academic peers.¹ SGPs provide a measure of how a student changes from one test to the next relative to other students with similar starting scores. SGPs range from 1–99 and interpretation is similar to that of percentile rank scores; lower numbers indicate lower relative growth and higher numbers show higher relative growth. For example, an SGP of 75 means that the student's growth from one test to another exceeds the growth of 75% of students with a similar beginning (pretest) score. All students, no matter their starting score, have an equal chance to demonstrate growth at any of the 99 percentiles. So, it is possible for a student who is scoring well above average at the beginning of the year to have an SGP that is relatively low. Take, for example, a student with a fall STAR percentile rank of 95 who receives an SGP of 19 at the end of the year. It may not seem fair that such a high performing student would receive a relatively low growth score, but this simply indicates that 81 percent of this student's academic peers from the same grade with the same starting score experienced more growth.

Renaissance Learning's STAR assessments for Early Literacy, Reading, and Mathematics became the first interim tests to adopt the SGP approach.

SGPs often are used to indicate whether a student's growth is more or less than can be expected. For example, without an SGP, a teacher would not know if a scaled score increase of 100 represents good, not-so-good, or average growth. This is because students of differing achievement levels in different grades grow at different rates relative to the test's scale. For example, a high-achieving second grader grows at a different rate than a low-achieving second grader. Similarly, a high-achieving second grader grows at a different rate than a high-achieving eighth grader. Scaled score growth of 100 points may be good for one student, but not so good for another student, depending on their grade level and pretest score.

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

Applying SGP to STAR™ assessments

STAR assessments are reliable, valid, and time-efficient assessments of early literacy skills (STAR Early Literacy), reading skills (STAR Reading), and mathematics skills (STAR Math). Quick and accurate results from the assessments provide teachers with specific benchmarking, screening, progress-monitoring, and diagnostic information to help tailor instruction, monitor growth, and improve achievement for all students. STAR is statistically linked to the research-based Core Progress learning progression, which provides a bridge between assessment and instruction to help educators ensure students are meeting college- and career-readiness standards, such as the Common Core State Standards. STAR assessments are also highly rated for screening and progress monitoring by the National Center on Response to Intervention (NCRTI) and the National Center on Intensive Intervention (NCII).

¹With STAR SGPs, academic peers are defined as students in the same grade with a similar scaled score on a STAR assessment at the beginning of the time period being examined.

Student Growth Percentiles represent the latest advancement in STAR assessments to help educators understand student growth. SGPs are available in STAR Reading and STAR Math for grades 1–12 and STAR Early Literacy for grades K–3. During the 2011–2012 school year, STAR assessments were the first interim tests to report SGPs.

To apply the SGP approach to STAR data, Renaissance Learning worked closely with the lead developer of SGP, Dr. Damian Betebenner, as well as technical advisor Dr. Daniel Bolt, an expert in quantitative methods and educational measurement from the University of Wisconsin–Madison. Because SGP was initially developed for measuring growth on state tests across years, applying the SGP approach to interim, within-year assessment data involved a number of technical challenges, primarily the differences regarding when STAR and state tests are administered.

Testing windows

State summative tests are typically administered once a year, at approximately the same time, to all students. Thus, score comparisons from one state test administration to another speak to growth across school years. On the other hand, STAR assessments are much more flexible, and may be administered to students as often as weekly, allowing for the examination of growth within a single school year. Decisions on when to administer STAR and to which students are left to local educators based on their purposes and needs for assessment. Most commonly, schools use STAR assessments as a screening and benchmarking test for all or nearly all students 2–4 times per year. Students requiring more frequent progress monitoring may take STAR assessments on a more regular basis to inform instructional decisions, such as whether the student is responding adequately to an intervention.

Because of this flexibility, not all students necessarily take STAR assessments at the same time; the number and dates of administration may vary from one student to the next. However, the majority of students test within at least two of the following time periods during the school year: fall (August 1–November 30), winter (December 1–March 31), and/or spring (April 1–July 31). We chose these date ranges when defining the data sets that would be used to determine Student Growth Percentiles. Therefore, we can provide SGPs for achievement that takes place between fall and winter STAR testing, winter and spring STAR testing, and/or fall and spring STAR testing. For example, to receive SGP scores for the fall-to-winter period, students must take at least one test between August 1 and November 30 and at least one test between December 1 and March 31. In addition, a sufficient amount of time must pass between tests. For half-year SGPs (fall to winter, winter to spring) to be calculated, students must take tests at least 60 calendar days apart. For full-year SGPs (fall to spring) to be calculated, students must take tests at least 180 calendar days apart.

Calculating SGPs

Growth models such as Student Growth Percentile require an enormous amount of data to generate reliable results (Castellano & Ho, 2013a). Fortunately, the widespread national use of STAR assessments provides a sufficient number of test data, enabling SGPs to be reported for nearly every student in every grade.²

SGPs often are used to indicate whether a student's growth is more or less than can be expected.

To calculate SGPs, Renaissance Learning collected millions of STAR test records from student data captured and stored on Renaissance servers. Because dates of STAR test administrations are not constrained to a specific time of year as is the case with most state summative tests, it was necessary to address the variability in the number of days between students' pre- and posttest dates. This was done by incorporating time into the SGP model. Taking this approach varied slightly from the typical SGP approach in that it considered a student's weekly rate of growth conditional on the pretest scaled score, instead of posttest scaled score conditional upon pretest scaled score. Quantile regression was applied to characterize the bivariate distribution of students' initial scores and weekly rates of growth. Students were grouped by grade and subject, and then quantile regression was used to associate every possible initial score and weekly growth rate combination with a percentile corresponding to the conditional distribution of weekly growth given the initial score.

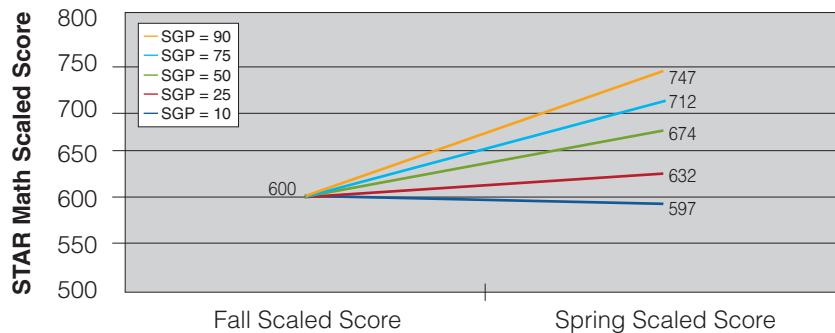
The result of these analyses was the creation of a look-up table in which initial STAR scores along with weekly growth rates were used as input to define SGPs for each grade, subject, and time period (e.g., fall to winter, winter to spring, fall to spring). The use of quantile regression techniques makes construction of such tables possible even though not all possible initial and

²Based on expert recommendation, Renaissance Learning does not report Student Growth Percentiles for extremely unusual pretest scores, such as for those students with extremely low or high pretest scores.

ending score combinations were observed in the student data. In general, the quantile regression approach can be viewed as a type of smoothing in which information from neighboring score values (initial scores and weekly rates of growth) can be used to inform percentiles for hypothetical score combinations not yet observed. As such, application of the methodology allows us to look up any score combination to obtain the percentile cut points for the weekly growth rate conditional achievement distribution associated with the given initial score. These cut points are the percentiles of the conditional distribution associated with the student's prior achievement. For example, using the quantile regression results of the sixth-grade STAR Math weekly growth rate on fall scores, we can calculate estimates for the 1st, 2nd, 3rd,...99th percentiles of growth from fall to spring. Using each of these cut points, SGPs can be reported for every subject, grade, and score combination.

Figure 2 illustrates how SGPs look in actual usage. Each line represents a different fifth-grade student starting out with a pre-test (fall) STAR Math scaled score of 600. While all students started at the same score at the beginning of the school year, each grew at a different rate throughout the year.

Figure 2: Depiction of Various SGPs Among Students With Same Pretest Score



- Student A (purple line) has a posttest score of 597, or a decline of 0.1 scaled score points per week, from fall to spring. This represents (extremely) low growth (SGP = 10).
- Student B (red line) has a posttest score of 632, or an increase of 1.0 scaled score points per week, from fall to spring. This represents relatively low growth (SGP = 25).
- Student C (green line) has a posttest score of 674, or an increase of 2.3 scaled score points per week, from fall to spring. This represents typical growth (SGP = 50).
- Student D (blue line) has a posttest score of 712, or an increase of 3.5 scaled score points per week, from fall to spring. This represents relatively high growth (SGP = 75).
- Student E (yellow line) has a posttest score of 747, or an increase of 4.6 scaled score points per week, from fall to spring. This represents extremely high growth (SGP = 90).

Description of Sample

Because STAR assessments are so widely used, Renaissance Learning has data for millions of testing events. Every summer, test records from recent school years are gathered and updated to recalculate Student Growth Percentiles.

While we do not report different SGPs for specific subgroups of students, all students—regardless of Special Education or English Learner status—are retained in the sample. However, we do limit the sample to STAR tests administered in typical school settings; tests administered by tutoring centers or virtual schools are excluded from the analysis.

Tables 1 through 3 describe the students in the SGP samples for STAR Early Literacy, STAR Reading, and STAR Math, respectively.

Table 1: Sample Characteristics, STAR Early Literacy™ SGP Study

		Sample %		
		Fall to Spring (n = 697,084)	Fall to Winter (n = 688,938)	Winter to Spring (n = 802,472)
Geographic Region	Midwest	20.8%	20.7%	22.3%
	Northeast	8.5%	9.3%	9.0%
	South	54.7%	53.7%	53.0%
	West	16.0%	16.3%	15.7%
	Response Rate	98.8%	98.7%	98.7%
School Type	Public	97.7%	97.6%	97.7%
	Private, Catholic	1.5%	1.6%	1.6%
	Private, Other	0.8%	0.8%	0.7%
	Response Rate	94.3%	94.1%	93.9%
School Enrollment	< 200	3.5%	3.6%	3.6%
	200-499	42.8%	43.2%	43.1%
	500-2,499	53.7%	53.2%	53.4%
	2,500 or more	0.0%	0.0%	0.0%
	Response Rate	96.3%	96.3%	95.8%
School Location	Urban	26.8%	25.9%	26.0%
	Suburban	25.1%	25.8%	26.4%
	Town	16.9%	17.1%	17.0%
	Rural	31.3%	31.1%	30.6%
	Response Rate	90.4%	90.6%	90.0%
Ethnic Group	Asian	2.6%	2.7%	2.7%
	Black	23.7%	22.8%	23.5%
	Hispanic	22.4%	21.8%	22.4%
	Native American	1.6%	1.8%	1.6%
	White	49.6%	50.9%	49.8%
	Response Rate	43.8%	42.7%	43.2%
Gender	Female	48.2%	48.0%	48.1%
	Male	51.8%	52.0%	51.9%
	Response Rate	82.0%	81.9%	81.7%

Table 2: Sample Characteristics, STAR Reading™ SGP Study

		Sample %		
		Fall to Spring (n = 3,528,829)	Fall to Winter (n = 4,019,291)	Winter to Spring (n = 4,114,791)
Geographic Region	Midwest	22.1%	21.0%	22.2%
	Northeast	9.7%	8.7%	9.7%
	South	47.9%	49.9%	48.8%
	West	20.3%	20.4%	19.3%
	Response Rate	97.8%	97.6%	97.7%
School Type	Public	96.2%	96.1%	96.4%
	Private, Catholic	2.4%	2.5%	2.3%
	Private, Other	1.4%	1.4%	1.3%
	Response Rate	93.6%	93.3%	93.3%
School Enrollment	< 200	3.4%	3.5%	3.5%
	200-499	36.5%	36.6%	36.9%
	500-2,499	59.8%	59.6%	59.4%
	2,500 or more	0.2%	0.3%	0.2%
	Response Rate	95.1%	94.8%	94.9%
School Location	Urban	28.2%	28.3%	27.9%
	Suburban	27.5%	27.1%	27.8%
	Town	16.1%	16.4%	16.1%
	Rural	28.2%	28.2%	28.1%
	Response Rate	89.1%	88.7%	89.0%
Ethnic Group	Asian	3.6%	3.5%	3.6%
	Black	25.1%	26.3%	24.8%
	Hispanic	18.3%	18.1%	18.7%
	Native American	1.7%	1.8%	1.9%
	White	51.4%	50.3%	51.1%
	Response Rate	45.6%	44.0%	44.5%
Gender	Female	49.1%	49.0%	49.0%
	Male	50.9%	51.0%	51.0%
	Response Rate	78.8%	77.6%	78.3%

Table 3: Sample Characteristics, STAR Math™ SGP Study

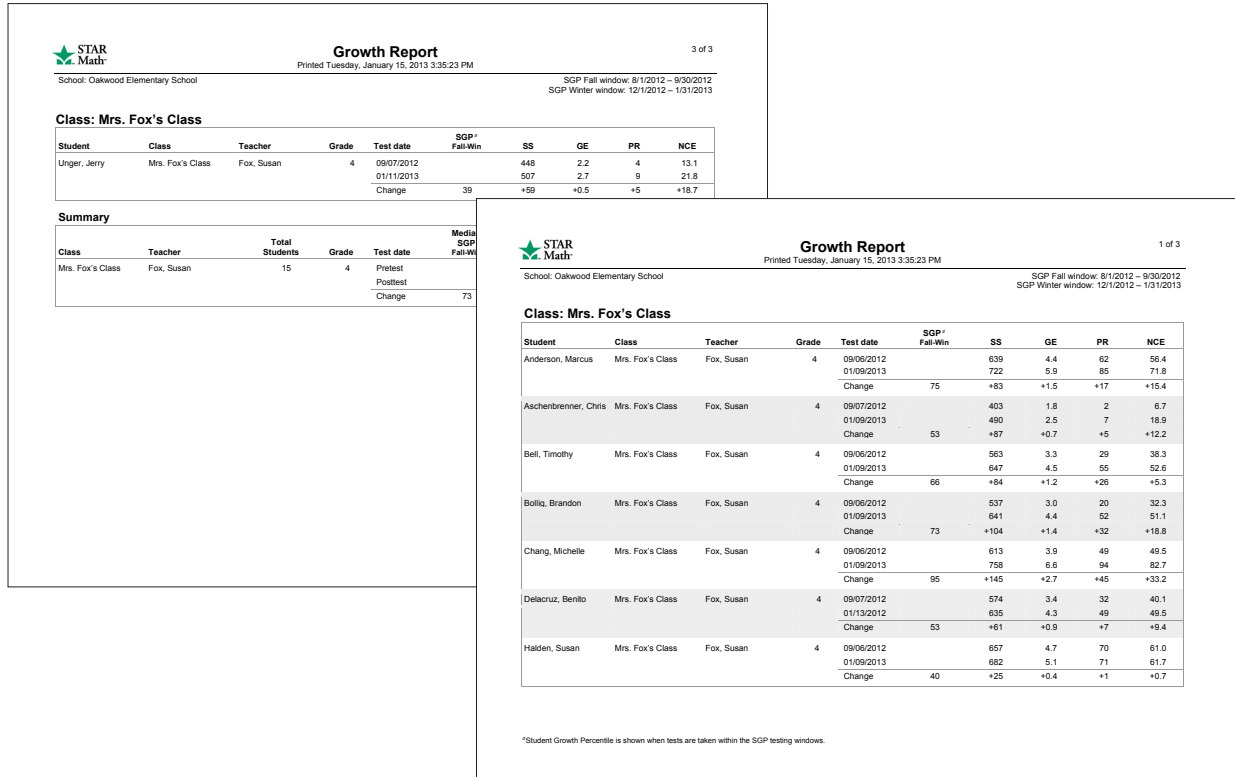
		Sample %		
		Fall to Spring (n = 2,290,143)	Fall to Winter (n = 2,607,713)	Winter to Spring (n = 2,507,238)
Geographic Region	Midwest	24.5%	23.5%	25.1%
	Northeast	11.6%	10.7%	12.6%
	South	50.6%	52.6%	49.2%
	West	13.3%	13.3%	13.1%
	Response Rate	98.3%	98.2%	98.3%
School Type	Public	97.6%	97.7%	97.7%
	Private, Catholic	1.6%	1.5%	1.5%
	Private, Other	0.8%	0.8%	0.8%
	Response Rate	93.3%	93.1%	93.0%
School Enrollment	< 200	3.3%	3.3%	3.4%
	200-499	37.3%	37.4%	37.6%
	500-2,499	59.3%	59.1%	58.8%
	2,500 or more	0.1%	0.2%	0.1%
	Response Rate	95.1%	94.8%	94.8%
School Location	Urban	22.7%	22.9%	22.6%
	Suburban	28.8%	27.9%	29.5%
	Town	16.8%	17.3%	16.8%
	Rural	31.7%	31.9%	31.1%
	Response Rate	88.8%	88.4%	88.6%
Ethnic Group	Asian	3.2%	3.1%	3.3%
	Black	27.1%	28.1%	25.8%
	Hispanic	13.1%	12.8%	14.6%
	Native American	2.0%	2.0%	2.0%
	White	54.6%	53.9%	54.2%
	Response Rate	49.9%	48.9%	49.3%
Gender	Female	49.0%	48.8%	48.8%
	Male	51.0%	51.2%	51.2%
	Response Rate	82.5%	81.7%	81.8%

Reporting SGPs in STAR™ Assessments

Student Growth Percentiles are shown in STAR assessments by viewing the Growth Report and the Growth Proficiency Chart.

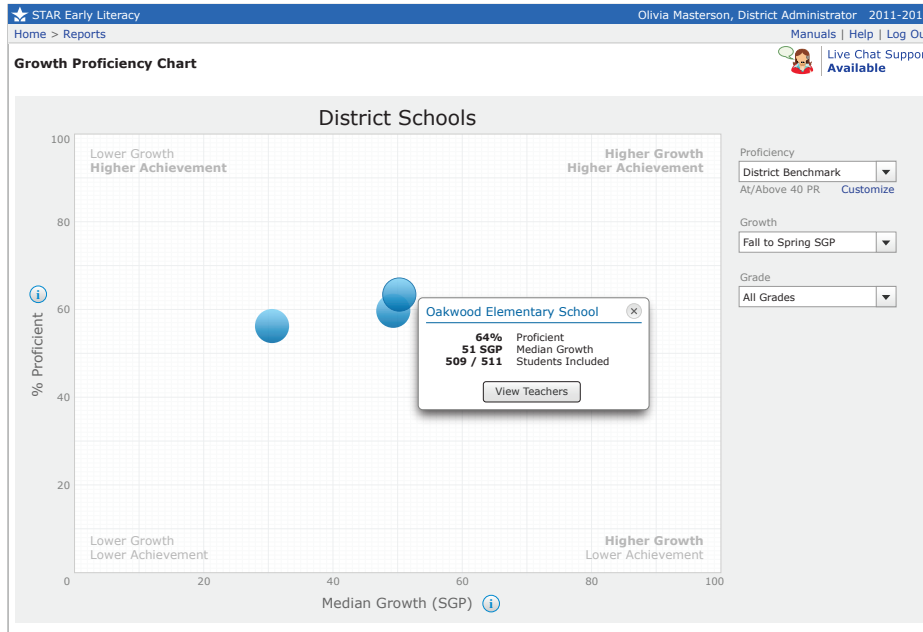
- Growth Report (see sample, Figure 3): The STAR Growth Report summarizes growth between two testing periods in the same school year. Teachers can run the report for a class or a specific group of students. Administrators can run it to see growth for each class or grade in their school. An SGP appears on the Growth Report as soon as a student has both a pretest and posttest score.

Figure 3: Sample STAR Math™ Growth Report



- Growth Proficiency Chart: The Growth Proficiency Chart is an interactive tool within the software. It visually displays data to show the relationship between estimated proficiency³ and growth. The chart shows at a glance which students, classes, or schools show the following: low proficiency and low growth, low proficiency and high growth, high proficiency and low growth, or high proficiency and high growth. Growth is expressed with SGPs. Figure 4 shows a district view.

Figure 4: Sample STAR Early Literacy™ Growth Proficiency Chart



Mean or Median?

In keeping with the vast majority of states who report SGPs on their state summative tests, Renaissance Learning reports median SGP. However, we recognize recent research on this topic (Castellano & Ho, 2013b) concluded that it can be appropriate to use either mean or median. Educators in states that report SGP on state summative tests may want to consult their state's position on this matter and use the statistic preferred by their state. All educators should exercise caution when aggregating SGP results for small classes/groups (fewer than 20 students) because both mean and median are subject to providing misleading estimates of central tendency, depending on the distribution of scores and the size of the group. For this reason, a number of states have chosen not to report SGP results for small groups.

³ STAR Reading and STAR Math are statistically linked to almost every state test to help answer one of the driving questions in education today, "Will students perform well on the state test?" In order to populate the Growth Proficiency Chart, we then combine data from the linking studies with expected weekly scaled score growth (from our decile-based growth norms, which take into account grade and starting observed score). STAR Early Literacy scores are not linked to state tests because most states do not test students until grade 3.

References

- Betebenner, D. W. (2009). Norm-and criterion-referenced student growth. *Educational Measurement: Issues and Practice*, 28(4), 42–51.
- Betebenner, D. W. (2010). *SGP: Student growth percentile and percentile growth projection/trajectory functions*. (R package version 0.0-6).
- Betebenner, D. W. (2011). *A technical overview of the student growth percentile methodology: Student growth percentiles and percentile growth projections/trajectories*. Dover, NH: The National Center for the Improvement of Educational Assessment.
- Castellano, K. E. & Ho, A. D. (2013a). *A practitioner's guide to growth models*. A paper commissioned by the Technical Issues in Large-Scale Assessment (TILSA) and Accountability Systems & Reporting (ASR) State Collaboratives on Assessment and Student Standards Council of Chief State School Officers.
- Castellano, K. E., & Ho, A. D. (2013b). Contrasting OLS and quantile regression approaches to student "growth" percentiles. *Journal of Educational and Behavioral Statistics*, 38(2), 190–215.
- Domaleski, C., & Perie, M. (2012). *Promoting equity in state education accountability systems*. Dover, NH: The National Center for the Improvement of Educational Assessment.
- Fox, L., Carta, J., Strain, P., Dunlap, G., & Hemmeter, M. L. (2009). *Response to intervention and the pyramid model*. Tampa, Florida: University of South Florida, Technical Assistance Center on Social Emotional Intervention for Young Children.

Independent technical reviews of STAR™ Assessments

- U.S. Department of Education: National Center on Intensive Intervention. (2012). Review of progress monitoring tools [Review of STAR Early Literacy]. Washington, DC: Author. Available online from <http://www.intensiveintervention.org/chart/progress-monitoring>
- U.S. Department of Education: National Center on Intensive Intervention. (2012). Review of progress monitoring tools [Review of STAR Math]. Washington, DC: Author. Available online from <http://www.intensiveintervention.org/chart/progress-monitoring>
- U.S. Department of Education: National Center on Intensive Intervention. (2012). Review of progress monitoring tools [Review of STAR Reading]. Washington, DC: Author. Available online from <http://www.intensiveintervention.org/chart/progress-monitoring>
- U.S. Department of Education: National Center on Response to Intervention. (2010). Review of progress-monitoring tools [Review of STAR Early Literacy]. Washington, DC: Author. Available online from <http://www.rti4success.org/ProgressMonitoringTools>
- U.S. Department of Education: National Center on Response to Intervention. (2010). Review of progress-monitoring tools [Review of STAR Math]. Washington, DC: Author. Available online from <http://www.rti4success.org/ProgressMonitoringTools>
- U.S. Department of Education: National Center on Response to Intervention. (2010). Review of progress-monitoring tools [Review of STAR Reading]. Washington, DC: Author. Available online from <http://www.rti4success.org/ProgressMonitoringTools>
- U.S. Department of Education: National Center on Response to Intervention. (2011). Review of screening tools [Review of STAR Early Literacy]. Washington, DC: Author. Available online from <http://www.rti4success.org/ScreeningTools>
- U.S. Department of Education: National Center on Response to Intervention. (2011). Review of screening tools [Review of STAR Math]. Washington, DC: Author. Available online from <http://www.rti4success.org/ScreeningTools>
- U.S. Department of Education: National Center on Response to Intervention. (2011). Review of screening tools [Review of STAR Reading]. Washington, DC: Author. Available online from <http://www.rti4success.org/ScreeningTools>



For more information, or for additional copies of this report, contact:

Educational Research Department
PO Box 8036 • Wisconsin Rapids, WI 54495-8036
(800) 338-4204 • www.renlearn.com

All logos, designs, and brand names for Renaissance Learning's products and services, including but not limited to Renaissance Learning, STAR, STAR Early Literacy, STAR Reading, and STAR Math, are trademarks of Renaissance Learning, Inc., and its subsidiaries, registered, common law, or pending registration in the United States and other countries.