STUDENT ASSESSMENTS
AND ASSOCIATED GROWTH MODELS FOR
TEACHER AND PRINCIPAL EVALUATION

FORM C

PUBLICLY AVAILABLE SERVICES SUMMARY

Assessment Provider Information	
[illegible]	[illegible]
[illegible]	[illegible] 612-254-2534 [illegible]
[illegible]	[illegible]
[illegible]	☐ [illegible] [illegible] ☐ [illegible] [illegible] ☐ [illegible] ☒ [illegible] ☐ [illegible] ☐ [illegible] ☐ [illegible] :
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible]

Is the assessment available, either for free or through purchase, to other districts or BOCES in New York State?	☒ YES ☐ NO

Please provide an overview of the assessment for districts and BOCES. Please include:

- A description of the assessment;
- A description of how the assessment is administered;
- A description of how scores are reported (include links to sample reports as appropriate);
- A description of how the Assessment Provider supports implementation of the assessment, including any technical assistance. (3 pages max)

aMath (Adaptive Math) is a simple, efficient, fully automated computer adaptive measure of broad math and component skills.

Number & Operations – Fractions (NF)
The NF domain is a part of the standards for students in third through fifth grade. In third grade, fractions are introduced to students as a new set of numbers in addition to whole numbers. Students are expected to understand fractions as partitions and compare fractions by reasoning about their size. Students at this level use math models involving equal parts or partitions to develop their understanding of fractions. For students in fourth grade, the ability to compare fractions is required. Students may also need to convert between decimal numbers and fractions. In fourth and fifth grade, students continue to extend their knowledge on operations of fractions with whole numbers and of fractions with fractions. By the end of fifth grade, students are expected to solve real-world problems with operations including multiplication, division, addition and subtraction.

Measurement & Data (MD)
The MD domain extends from kindergarten through fifth grade and addresses conversion of units as well as the interpretation of data. At kindergarten, students are tested on classifying and comparing objects with measurable attributes. In first and second grades, students develop their ability to work with variables such as time, length, and volume. By fourth grade, students are required to convert various units in a given measurement system. Through all grade levels, students are expected to develop and understanding of data on diagrams. By the end of fifth grade, students may be asked to complete tasks such as creating a line plot of data or using different operations to calculate measurements.

Geometry (G from kindergarten tnntiGeomw8scext(SC /TT4 1 Tf 06mp MCIDstan75sng 821 Tudentrangating

needing improvement, and growth trends within and across school years. District Managers, School Managers, and Specialists within the FAST system may run grade-wide reports from the FAST Reports Manager. District Managers also have access to run reports for multiple schools in the district at once. In addition to the standard FAST reporting, FastBridge Learning offers additional ad-hoc and custom reporting capabilities via our “Off-Line Reporting” feature. These

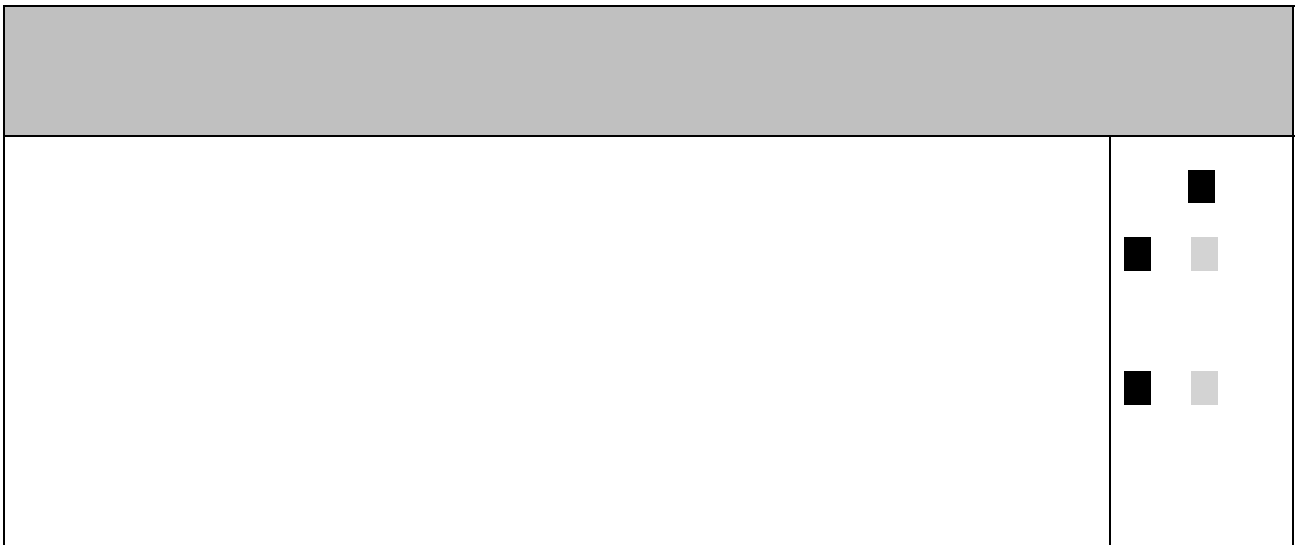
Educator effectiveness was estimated for evaluation purposes using medians of SGP, i.e., median growth percentiles (MGP), for those students associated with a given educator. MGP are expressed on the same metric as SGP, and, like SGP, range from 0.01 to 0.99. MGP can then be converted to an Annual Professional Performance Review score (APPR) using the crosswalk tables presented below for each assessment. APPR values are also linked to HEDI ratings (4 = highly effective, 3 = effective, 2 = developing, and 1 = ineffective). Note that these crosswalk tables are based on preliminary norming data for educators, and will be updated at the completion of the 2015/2016 and annual thereafter based on updated norming data. An aMath crosswalk table is provided on page 16 of the Growth Report in Appendix A-1. APPR scores were assigned to educator median growth percentiles (MGP) so that a HEDI rating of "Ineffective" corresponded to APPR scores from 0 to 12, "Developing" corresponds to APPR scores from 13 to 14, "Effective" to APPR scores from 15 to 17, and "Highly Effective" to scores from 18 to 20. Based on this crosswalk, MGP for the "Ineffective" category extend to 0.59, and MGP for "Developing" then extend from 0.60 to 0.74. "Effective" MGP range from 0.75 to 0.89, and "Highly Effective" MGP range from 0.90 to 0.99. The FAST online system handles the administration and scoring of assessments and reporting of results. Norming data collected during the 2015/2016 school year will be integrated into the online reporting functionality prior to the 2016/2017 school year. Student growth estimates over screening periods will be reported with standard errors, and SGP will be provided for any students enrolled for at least 70% of the school year having fall and spring assessment scores. Educators having SGP results from at least 15 students meeting these criteria will then be provided with MGP APPR scores, and HEDI ratings using updating crosswalk tables. For additional details, please reference Formative Assessment System for Teachers: Growth Modeling for Educator Evaluation submitted as part of Appendix A.

New York State Next Generation Assessment Priorities Please provide detail on how the proposed supplemental assessment l or assessment to be used with SLOs addresses each of the Next Generation Assessment Priorities below.	
Characteristics of Good ELA and Math Assessments (only applicable to ELA and math assessments):	The aMath assessment is consistent with best practices in measuring the New York State Learning Standards in mathematics. Reliability and validity evidence supports the use of aMath for the purpose of measuring student growth across the following domains, which are aligned with NYS standards in mathematics: Counting and Cardinality, Operations and Algebraic Thinking, Number and Operations in Base Ten, Number and Operations, Measurement and Data, and Geometry. aMath item development followed the process and standards presented by Schmeiser and Welch (2006) in the fourth edition of Educational Measurement (Brennan, 2006). Research assistants, teachers from each grade level (1st through 5th), and content experts in the area of math served as both item writers and reviewers. After items were written they were reviewed for feasibility, fairness, construct relevance, and content balance. A stratified procedure was used to recruit a diverse set of item writers from urban, suburban and rural areas. The

	item writers wrote, reviewed, and edited assessment materials. Item writing for aMath was a multi-year, collaborative, and iterative process. First the literature on item writing guidelines typically used in developing assessments was reviewed. Next, the literature on multiple-choice item writing was reviewed. Once the literature was reviewed, the guidelines were applied to aMath to examine relevance and utility. Extensive guidelines were provided to item writers and the process outlined above was followed. The aMath project uses a research-based skills hierarchy and unified construct of broad math achievement to establish an instructionally relevant assessment. The importance and emphasis on each component skill (domain) varies across children. Each assessment is individualized by the aMath software and built-in assessment algorithms. As a result, the information and precision of measurement is optimized regardless of whether a student functions at, above, or below grade level (i.e., same age and grade peers). The grade labels and content balancing that are proposed in the a-priori model derive from the recommendations of expert panels and are subject to empirical evaluation and refinement. Additional information about aMath item development is included in the Technical Manual submitted with Appendix A-2 starting on page 157.
Assessments Woven Tightly Into the Curriculum:	We believe the best assessments are those that are able to be seam v-2 s (mode0a ar) aMath pro Td [(Wnt .6 (t is in-1.14 Tc 0 Tw

Performance Assessment:	Reliability and validity evidence supports the use of aMath for the purpose of measuring student growth across the following domains, which are aligned with NYS standards in mathematics: Counting and Cardinality, Operations and Algebraic Thinking, Number and Operations in Base Ten, Number and Operations, Measurement and Data, and Geometry. The FAST assessments are evidence-based. Numerous studies were completed with diverse samples of students across many geographic locations and LEAs (e.g., NY, GA, MN, IA, and WI). Consistent with the definitions of "evidence-based," there are many large, multi-site studies with student samples from the populations and settings of interest (i.e., K–12 students). The samples size for almost all studies well-exceeded the requirement of 50 students per condition (e.g., assessment, grade, LEA, instructional condition). On aggregate, more than 15,000 students participated in well-controlled psychometric research. In addition, norms were developed from samples of approximately 8,000 students per grade (K to 8th) per assessment, which aggregates to 72,000 student participants. Consistent with the requirements for evidence, the psychometric qualities for reliability and validity were statistically significant, and the various assessments are meaningful and statistically robust indicators of relevant outcomes, such as state tests and future performance in school. FastBridge Learning uses standard setting processes to summarize student performance. Standards may be used to inform goal setting, identify instructional level, and evaluate the accuracy of student performance. The FastBridge Learning software provides various resources to assist administrators with test result interpretations. For example, a Visual Conventions drop down menu is available to facilitate interpretation of screening and progress monitoring group and individual reports. Percentiles are calculated for local school norms unless otherwise indicated. Local school norms compare individual student performances to their same grade and school peers. Methods of notation ar.153t result groual sch147 TD [(sch et isk(, wwer astwoe xclartationmmakus rfper)]TJ pgt al sti 0 -1.147 TD assessment63 dsbdenhmmaku, -base of FastBridge assessmentsmorsivolvrs

	monitoring. Additionally, online training modules include sections on administering the assessments, interpreting results, screen casts, and videos. Results should always be interpreted carefully considering reliability and validity of the score, which is influenced by the quality of standardized administration and scoring. It important to consider the intended purpose of the assessment, its content, the stability of performance over time, scoring procedures, testing situations, or the examinee. The FastBridge Learning system automates analysis, scoring, calculations, reporting and data aggregation. It also facilitates scaling and equating across screening and progress monitoring occasions.
Efficient Time-Saving Assessments:	Students typically complete the aMath assessments in 20-30 minutes, reducing testing time by up to 50-95% compared to traditional tests. Our extensiP16.830. which is infh25in-0.u



2.2(D)-i: Technical Documentation Related to Assessment and Student Growth Score Properties: RELIABILITY Both "minimum" and "desired" qualifications are listed. For the purposes of this RFQ, applications will only be rated against the "minimum" qualifications; however, NYSED's aspirational "desired" qualifications are also listed to identify possible future requirements for assessments and associated growth models.	
For supplemental assessments used in conjunction with growth models: This application contains evidence of the minimum criteria for reliability: x Student test scores have adequate levels of reliability (e.g., coefficient alpha > 0.75). This application contains evidence of the desired criteria for reliability: x Standard errors provided for students growth scores. x Student growth classifications have adequate decision consistency. x Teacher effectiveness classifications demonstrate adequate consistency. Examples include agreement statistics (e.g., kappa coefficients) based on simulation studies.	Check all that apply: ■ ☐ ☐ ☐
2.2(D)-ii: Technical Documentation Related to Assessment and Student Growth Score Properties: VALIDITY – ALIGNMENT Both "minimum" and "desired" qualifications are listed. For the purposes of this RFQ, applications will only be rated against the "minimum" qualifications; however, NYSED's aspirational "desired" qualifications are also listed to identify possible future requirements for assessments and associated growth models.	
For supplemental assessments used in conjunction with growth models: This application contains evidence of the minimum criteria for alignment validity: x Evidence that test content is sufficiently aligned with New York State Learning Standards and covers a range of measurable standards. Documentation that demonstrates that: (a) at least 80% of the test measures content aligned with NYS learning standards, (b) no more than 20% of test content is aligned with other learning standards or objectives, and (c) a range of content from the NYS learning standards is measured Note: Other relevant standards can be proposed if NYS Learning Standards do not apply to subject area. This application contains evidence of the desired criteria for alignment validity: x 100% alignment between NYS Learning Standards and assessment.	Check all that apply: ■ ☐
2.2(D)-iii: Technical Documentation Related to Assessment and Student Growth Score Properties: VALIDITY – RELATIONS TO OTHER VARIABLES Both "minimum" and "desired" qualifications are listed. For the purposes of this RFQ, applications will only be rated against the "minimum" qualifications; however, NYSED's aspirational "desired" qualifications are also listed to identify possible future requirements for assessments and associated growth models.	
For supplemental assessments used in conjunction with growth models: This application contains evidence of the minimum criteria for validity in relation to other variables: x Evidence students' growth scores are correlated with other measures of student progress (e.g., r > .5 with measures such as the number of objectives	Check all that apply: