STUDENT ASSESSMENTS
AND ASSOCIATED GROWTH MODELS FOR
TEACHER AND PRINCIPAL EVALUATION

FORM *C*

PUBLICLY AVAILABLE SERVICES SUMMARY

This form will be posted on the New York State Education Department's Web site and distributed through other means for all applications that are approved in conjunction with this RFQ to allow districts and BOCES to understand proposed offerings in advance of directly contacting Assessment Providers regarding potential further procurements.

Assessment Provider Information	
Name of Assessment Provider:	Data Recognition Corporation
Assessment Provider Contact Information:	Genevieve Olvera
Name of Assessment:	*pre*LAS Assessment System (LAS Links family of products) Early Childhood Assessment of English and Spanish Language Proficiency
Nature of Assessment:	☐ ASSESSMENT FOR USE WITH STUDENT LEARNING OBJECTIVES WITH A TARGET SETTING MODEL; OR ☒ SUPPLEMENTAL ASSESSMENT WITH AN ASSOCIATED GROWTH MODEL: ☒ GAIN SCORE MODEL ☐ GROWTH-TO-PROFICIENCY MODEL ☐ STUDENT GROWTH PERCENTILES ☐ PROJECTION MODELS ☐ VALUE-ADDED MODELS ☐ OTHER:
What are the grade(s) for which the assessment can be used to generate a 0-20 APPR score?	Prekindergarten – 1st Grade (3 – 6 year olds)
What are the subject area(s) for which the assessment can be used to generate a 0-20 APPR score?	English Language Proficiency, Spanish Language Proficiency, Pre-Literacy Skills
What are the technology requirements associated with the assessment?	*pre*LAS is an observational assessment delivered by paper-based format. DRC provides scoring options which include handscoring and local scanning.
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)**

DESCRIPTION OF ASSESSMENT

The *pre*LAS Assessment System (*pre*LAS) is designed to measure English and Spanish language proficiency and pre-literacy skills of learners in early childhood—students aged 3 to 6 years old. This assessment can be used for pre-kindergarten programs to first grade. *pre*LAS assesses children through authentic, naturalistic observations and provides a standardized way to screen and place students by evaluating listening and speaking skills. *pre*LAS can be used for program entry and exit, instructional placement, identifying academic strengths, and monitoring progress and growth. The assessment is aligned with recommendations by the National Association for Education of Young Children (NAEYC) and is strongly aligned to the New York State Learning Standards (See Appendix C). An overview of the assessment and details on how the test is administered is described below.

- *pre*LAS is a paper-based assessment that can be administered in 20 minutes or less in a one-on-one setting.
- The assessment is available with two forms in English and one form in Spanish.
- *pre*LAS comprises of five sub-tests that measure receptive and expressive language.
- The test was normed using a national sample of students (see Appendix D).
- The assessment has five performance levels ranging from levels 1 (non-speaker) to level 5 (fluent speaker). Each level is further broken down into total scores used to measure progress over time.
- The *pre*LAS Oral component is designed for children ages 3 to 6. The assessment will provide school administrators and staff a proficiency level and total score for each student based on language skills.
- The test also includes a Pre-Literacy component appropriate for children ages 5 and 6. The Pre-Literacy score is a separate score designed to provide school administrators and staff with additional information on student pre-literacy skills in the following categories: numbers, letters, shapes, colors, reading and writing.
- *pre*LAS is a test designed to measure young children's expressive and receptive abilities in three linguistic components of oral language: morphology, syntax, and semantics.

The format of the test has been carefully designed to include colorful illustrations and a game-like appearance to keep the child's attention. The assessment is designed to be administered one-on-one and takes approximately 10-20 minutes.

The six subscales used in *pre*LAS are:

1. The child's ability to follow instructions.
2. The child's ability to understand simple oral instructions as well as language used to talk about relationships, likenesses and differences. The child points to one of two pictures, which best represents, the oral stimulus sentence.
3. The child's ability to provide labels for common household objects—articles of clothing, eating utensils, and furniture. The administrator points to various items and asks the

student to identify the item.

4. The child's expressive ability with morphological and syntactical features through the repetition of oral stimulus sentences.
5. The child's ability to provide an appropriate clause to complete a compound or complex sentence.
6. The child's ability to listen to a short story and then retell it. Sample probe questions are provided for use when the child is shy or reluctant to talk.

The Oral component of *pre*LAS is made up of five subtests that measure receptive and expressive language. Listed below are the five subtest descriptions:

Simon Says: This section tests receptive language, listening comprehension, following directions, and total physical response (TPR). This section utilizes simple directives typically encountered in early kindergarten classrooms. The vocabulary words refer to parts of the body and to items commonly encountered in household and pre-school environments such as pencil, floor, paper, and door.

Art Show: This section assesses expressive language and utilizes graphic stimuli to elicit labels for a number of concrete nouns without inflectional markers. The lexical items assessed in this subtest include concrete nouns, single-word responses, words that are commonly used in a household environment, etc. This section of the test shows a student's ability to produce oral vocabulary and verb phrases at appropriate levels of development.

Say What You Hear: This session assesses the student's receptive and expressive abilities with morphological and syntactical features and focuses on sentence repetition and grammatical features.

The Human Body: This section of the test asks students to name 10 parts of the human body and focuses on lexicon commonly acquired by native-speaking children.

Let's Tell Stories: This session assesses expressive language and represents an integrative approach to language testing. Students transform "storytelling and retelling" into their own words, reflecting a level of syntactic development. Students demonstrate their ability to produce complete sentences, retell simple narration with picture cues.

DESCRIPTION OF HOW THE ASSESSMENT IS USED

*pre*LAS comes in two forms in English; C and D. Form C is typically used for placement and identification or base lining a student's language proficiency. Form D is used as the follow-up test to measure student progress.

The test can be administered by any teacher or appropriate school personnel 1) who are qualified to work with four-to-six year-olds, 2) who are proficient speakers of standard English, 3) who are completely familiar with all administrative aspects of the test, either through a workshop or through self-study, and 4) who are able to distinguish correct from incorrect responses. CDs with test items and stories are provided as a means of standardizing test stimuli. The CDs are optional to the extent that they may be distracting to young children.

Details for establishing the reliability of test administration and scoring are described in the *pre*LAS Exam Manual located in Appendix E. Sample protocols are provided in this manual.

*pre*LAS is individually administered. Approximately 20 minutes should be allocated for the testing of each student. Testing should be done in a quiet area, and the student should not

have heard the test items before the test is administered to him or her.

Picture stimuli are presented in an easel-style book and the test administrator records correct responses on an individual score sheet as the test progresses. An optional CD may be used for "Simon Says", "Choose a Picture", and "Say What You Hear" to ensure that standard English is modeled for the examinee. Test instructions may be given in any language that helps the examinee understand what is expected, but all test items are administered in English. The Scoring and Interpretation Manual provides guidelines for identifying appropriate responses and resolving scoring ambiguities.

Pre-literacy skills are assessed with a fun board game that is designed to capture receptive and expressive language skills in action. As the teacher helps the student move around the board, these skills are assessed:

- Upper and lowercase letter recognition
- Number recognition and concepts
- Color recognition
- Shapes and spatial relationships such as “in front of,” “under”
- Reading two– and three–letter sight words, such as “and,” “is,” and “up”
- Writing name, age, and two– and three–letter sight words

DESCRIPTION OF HOW SCORES ARE REPORTED

Scoring is completed locally by a trained teacher. Raw scores correspond to one of three categories: Non-English Speakers, Limited English Speakers, and Fluent (Proficient) English Speakers, which are reported in one of five proficiency levels. Scoring is completed on the front cover of the student answer document. *pre*LAS provides the following set of reports:

- Teacher: Group list Report
- Student: Performance Profile (Student Proficiency Report)

Sample reports are provided in Appendix F of this submission.

THE IMPLEMENTATION OF THE ASSESSMENT AND TECHNICAL ASSISTANCE

Our professional development team enables educators to gain new skills to improve the instruction and assessment of students and enhance teachers’ ability to understand and use assessment measures and results and link them to instructional strategies.

The *pre*LAS Administrators Guide describes how to interpret *pre*LAS test results and help educators understand test content, standards assessed, research basis of the tests, and scoring tables.

Customized on-site professional development, available upon request, provides hands-on training in test administration and interpretation of *pre*LAS results. All New York LEA’s will have access to our Customer Care team with members from Customer Service, Scoring, Research, Technology, Test Development, and Product Management. Working together, this team ensures that every customer receives personal, helpful, and timely responses regardless of whether service is requested though the website, email, or through a phone call. For general support including scoring services representatives are available from 9:00 AM to 7:00 PM Eastern Time.

Please provide an overview of the student-level growth model or target setting model for SLOs for districts and BOCES, along with how student-level growth scores are aggregated to the create teacher-level scores, and how those teacher-level scores are converted to New York State's 0-20 metric.

*pre*LAS provides two types of normative scores: percentile rank (PR) and normal curve equivalent (NCE). The NCEs have many characteristics in common with percentile ranks, but have the additional advantage of being based on an equal-interval scale. The use of NCEs allows meaningful comparisons between different assessment series and between different tests within the same assessment series.

For each component being measured (Oral or Pre-Literacy), the obtained NCE score from the pre-test can be subtracted from that from the post-test to derive the growth score for each student. We propose using the median of student-level growth scores per grade class per teacher as the raw teacher-level score at each grade level, given that classes may have varying sizes and some of them may also show extreme cases. The median has been well known statistically being more stable and robust to outliers than the mean, particularly when the sample size is small (N<30). In addition, the median is easy to understand to educators, which supports interpretability and usability of the teacher-level growth scores.

When a teacher has multiple class types and a single score needs to be calculated for that teacher, we recommend using a weighted average of the teacher-level scores across class types as the single teacher-level score. DRC will assist districts in determining the optimal weight values to use based on the instructional settings and curriculum emphasis in each district and observed empirical statistical distributions.

To account for learner difference and compare teachers to their peers of similar student groups, we recommend that a Z score (or its transformation to a T score, depending on the use and district preference) of the raw teacher-level growth score be used for formal reporting and decision making. DRC will actively partner with districts to collect student-level and teacher-level growth data and the teacher's APPR and HEDI data over time to support a well-maintained, on-going validated crosswalk between the teacher-level growth scores and the APPR scale.

To support the crosswalk, teacher growth scores will be calibrated in consideration of teachers with similar student groups regarding student prior academic history, social-economic status, disability, and English language learner status using linear regression, in which the raw teacher-level growth score is predicted with student-level demographic statistics. After obtaining the regression coefficients, the predicted teacher-level growth score, as well as the associated standard deviation, for the teacher of interest will be calculated based on the particular student cohort of that teacher.

The predicted score and standard deviation can be used as the sample mean and standard deviation to convert the raw teacher-level growth score to a Z score. The Z score will then be mapped to the 0-20 APPR and the HEDI rating categories. The decision rules for mapping are proposed to be similar to those for the state HEDI category assignment to support consistency and interpretability. The proposed decision rules as listed below.

Highly Effective: $Z \geq 1.5$; **Effective**: $1.5 > Z \geq -1$; **Developing**: $-1 > Z \geq -1.5$; **Ineffective**: $Z < -1.5$

The obtained ratings will be correlated with those for HEDI empirically, with the correlation results monitored over time to determine if any finer adjustment to the rules, such as introducing the confidence range into the rules, is desirable to support better consistency with the state HEID category assignment. The use of linear regression for Z score calculation makes it possible to compare the raw teacher-level growth score to those from teachers of similar student groups in an empirical, consistent, and principled way.

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):	N/A
Assessments Woven Tightly Into the Curriculum:	The *pre*LAS proficiency assessment represents a convergent approach to language assessment. It is developmentally, linguistically, and psychometrically appropriate for children 3 to 6 years old. The definition and refinement of content specifications were continuous during the development of *pre*LAS. The tests fully meet the strict criteria for reliability and validity of the American Psychological Association. After items were matched to the initial theoretical rationale, test content was verified as part of the procedures for development of items, analysis of pilot data, analysis of national and international standardization data, and selection of final test items. The procedures were designed to ensure that stimulus materials and items meet the content criteria established for the testes, were well constructed, and were written in language appropriate for this level of testing. The final set of items for each part of the test was selected from a pool of items that reflected the identified objectives by a committee of experts. Criteria for item selection included level of difficulty, level of discriminate validity, inter-item and inter-subscale reliability, age and grade level appropriateness, and culture bias.
Performance Assessment:	The *pre*LAS assessment reports in proficiency levels and total scores. This data provides information indicating student performance and growth over time.
Efficient Time-Saving Assessments:	*pre*LAS is designed to take approximately 15 minutes to administer, per student, and provides the flexibility of on-site scoring and tabulation for immediate results.
Technology:	*pre*LAS is delivered via paper-based format.
Degree to which the growth model must differentiate across New York State's four levels of teacher effectiveness (only applicable to supplemental assessments):	The use of Z score (or its transformation to a T score) for reporting of the teacher-level growth score as proposed in the growth model supports meaningful differentiation of obtained teacher scores across New York State's four levels of teacher effectiveness. The distribution of the observed teacher scores across the four levels, along with their correlations, will be monitored on an ongoing basis to support validity and any desired enhancement of the growth model to keep up-to-date with the particular educational context, needs, and emphasis in districts.

STUDENT ASSESSMENTS FOR TEACHER AND PRINCIPAL EVALUATION

FORM *G*

ATTESTATION OF TECHNICAL CRITERIA – SUPPLEMENTAL ASSESSMENTS WITH CORRESPONDING GROWTH MODELS

Please read each of the items below and check the corresponding box to ensure the fulfillment of the technical criteria outlined in the Technical Application on "FORM B-2".

PLEASE SUBMIT ONE "FORM G" FOR EACH APPLICANT. CO-APPLICANTS SHOULD SUBMIT SEPARATE FORMS.

COMPLETE THIS SECTION:

2.2(A) Narrative Overview of Proposed Supplemental Assessment and Associated Growth Model	
This application contains a short overview of the assessment being proposed, including the intended purpose of the assessment, and how the assessment is administered.	☑
For supplemental assessments, this application contains a description of the growth model and how it is used in conjunction with the assessment.	☑ ☐ N/A
For K-2 assessments, this application contains evidence that the proposed assessment is consistent with this RFQ's requirement that the assessment not be a "Traditional Standardized Assessment" as defined above in the section "Definitions of Key Terms Used in this RFQ."	☑ ☐ N/A
2.2(B) Evidence of Capability	
This application provides an overview of services provided by the Assessment Provider, including a description of the range of support / technical assistance that the Assessment Provider would provide to an LEA if selected by an LEA for this service.	☑
This application contains information as to whether the Applicant or Assessment Provider has been denied approval as a provider of assessment services in another state(s) and the reason(s) for such denial. If denied within New York State, the location and reason are indicated.	☑ ☐ N/A
2.2(C): Evidence of Copyright Owner/Assessment Representative History of Assessment Development	
This application contains evidence that the Copyright Owner/Assessment Representative has a history of developing assessments of student learning (achievement or growth) for the purpose of making defensible judgments about educator effectiveness.	☑ ☐ N/A

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: • Student test scores have adequate levels of reliability (e.g., coefficient alpha > 0.75). This application contains evidence of the *desired* criteria for reliability: • Standard errors provided for students growth scores. • Student growth classifications have adequate decision consistency. • 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: • 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: • 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: • Evidence students' growth scores are correlated with other measures of student progress (e.g., r > .5 with measures such as the number of objectives mastered by a student over the course of the year, teachers' ratings of students' progress, or scores from other assessments).	**Check all that apply:** ☑

This application contains evidence of the *desired* criteria for validity in relation to other variables: • Evidence teacher effectiveness ratings are positively correlated (e.g., r > .5) with other measures of teaching effectiveness.	☐
2.2(D)-iv: Technical Documentation Related to Assessment and Student Growth Score Properties: VALIDITY – INTERNAL STRUCTURE ***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 of internal structure: • Scale properties appropriate for growth model used (*see notes*). Total scores and subscores on student assessments should be supported by dimensionality analyses (e.g., IRT residual analyses, factor analyses). This application contains evidence of the *desired* criteria for validity of internal structure: • Evidence students' scores are on an interval scale. **Notes: If **gain score model** is used, evidence is needed that students' pretest and posttest scores are on the same scale. If **student growth percentile** model used, justification for the number of years included in the model should be provided. If **growth-to-proficiency**, projection, or **value-added models** are used, evidence is needed that the model explains a significant amount of variability in student achievement. Also, models should demonstrate robustness to missing data.*	**Check all that apply:** ☑ ☑
2.2(D)-v: Technical Documentation Related to Assessment and Student Growth Score Properties: UTILITY AND COMPREHENSIBILITY ***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 utility and comprehensibility: • Technical documentation that describes how student growth and educator effectiveness are calculated. This application contains evidence of the *desired* criteria for utility and comprehensibility: • Student growth reports support instructional improvement. Resources and supporting materials available to the field.	**Check all that apply:** ☑ ☐
2.2(E)-i: Technical Documentation Related to Aggregating Student-Level Growth Scores to Teacher-Level Scores: CREATION OF TEACHER LEVEL SCORES	
For supplemental assessments used in conjunction with growth models: This application includes a narrative description of how student-level scores are aggregated to create a single teacher-level score for each teacher.	☑ ☐ N/A

2.2(E)-ii: Technical Documentation Related to Aggregating Student-Level Growth Scores to Teacher-Level Scores: EXCLUSION RULES	
This application includes a description of any exclusion rules that remove students associated with a given teacher from the teacher's teacher-level score (either through a growth model or in conjunction with an SLO).	☑ ☐ N/A
2.2(F): Technical Documentation Related to Converting Teacher-Level Growth Score to New York State's 0-20 APPR Scale	
This application includes a crosswalk that maps scores on the assessment's aggregated teacher-level growth score to the required New York State teacher and principal evaluation metric, which ranges from 0-20.	☑
This application includes procedures for converting teacher-level growth scores to the 0-20 APPR scale comply with the New York Standards for each evaluation rating category, which are based on the following definitions.	☑
For supplemental assessments used in conjunction with growth models: This application includes an explanation of the assignment of HEDI rating categories based on the following ranges: • Highly Effective: results are well-above State average* for similar students • Effective: results meet State average* for similar students • Developing: results are below State average* for similar students • Ineffective: Results are well-below State average* for similar students	☑ ☐ N/A
2.2(G)-i: Technical Documentation Related to Fairness: TEST TAKERS Consistent with the new Testing Standards (2014), there is an increased focus in the industry on fairness of assessments and their uses. Please provide evidence of fairness for both the proposed assessment and, if applicable, the proposed growth model.	
This application includes evidence that the proposed assessments are fair to all test takers (e.g., Differential Item Functioning [DIF] / bias information, fairness evaluation / sensitivity review plan.)	☑
2.2(G)-ii: Technical Documentation Related to Fairness: TEACHER GROWTH SCORES	
This application includes evidence of fairness of the proposed aggregated teacher growth scores (e.g., lack of correlation between aggregated teacher growth scores and student demographics).	☑
The evidence of fairness of the proposed aggregated teacher growth scores includes an explanation of how the growth model incorporates (a) prior academic history, (b) poverty, (c) students with disabilities, and (d) English language learners.	☑ ☐ N/A

To be completed by the Copyright Owner/Assessment Representative of the assessment being proposed and, where necessary, the co-applicant LEA:

1. Name of Organization (PLEASE PRINT/TYPE) Data Recognition Corporation	4. Signature of Authorized Representative (PLEASE USE BLUE INK) 
2. Name of Authorized Representative (PLEASE PRINT/TYPE) Susan S. Engeleiter	5. Date Signed February 2, 2016
3. Title of Authorized Representative (PLEASE PRINT/TYPE) Chief Executive Officer and President	

1. Name of LEA (please print/type) N/A	4. Signature of School Representative (Please use Blue ink) N/A
2. School Representative's Name (please print/type) N/A	5. Date Signed N/A
3. Title of School Representative (please print/type) N/A	