**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:	**Houghton Mifflin Harcourt™ Publishing Company (HMH®)**
Assessment Provider Contact Information:	Phil Contompasis, Account Executive, at 518-313-9366 or via e-mail at phillip.contompasis@hmhco.com.
Name of Assessment:	*Gates-MacGinitie Reading Tests®* (*GMRT®*)
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: *GMRT* Growth Model
What are the grade(s) for which the assessment can be used to generate a 0-20 APPR score?	Grades K–12
What are the subject area(s) for which the assessment can be used to generate a 0-20 APPR score?	Reading
What are the technology requirements associated with the assessment?	The *GMRT* can be administered on paper at all levels; reports can be accessed online. Minimum technology requirements are: **Supported Windows OS/Browsers:** • OS: Windows 10, Windows Vista(If Service Pack 2 is installed), Windows 7, Windows 8.1 • Browsers: Microsoft Edge, IE 9.0 or later, Firefox 31.x(Extended Support), 37.0 or later **Supported Mac OS/Browser:** • OS: Mac 10.5.8(Intel) or later • Browsers: Safari 4.0 or more, Firefox 38.2 or later **Additional Requirements:** • Acrobat® Reader®: 8.0 (to view reports and

	documentation) • Adobe Flash 9.x • Cookies must be enable or allowed • Popup Blockers must be turned off • Modem or Internet connection • Mouse or compatible pointing device: For use in answering test questions • Super VGA (1024 x 768) or higher-resolution monitor with 256 colors • Browser supporting Java Script • Speakers or headphones and audio card if audio is needed **Minimum Network Requirements:** • Network Connections: 15 Megabits or higher • Firewall Configuration: Ports 80 and 81 • Content Filtering: Allow rpclearning.com URLs through content filters
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)**

NOTE TO REVIEWERS: Since this cell would not expand beyond one page, the three-page description follows.

Description of the *Gates-MacGinitie Reading Tests® (GMRT®)*

Houghton Mifflin Harcourt Publishing Company's™ (HMH®) *Gates-MacGinitie Reading Tests (GMRT)* are an individual- or group-administered norm-referenced reading survey battery that is administered on paper. These assessments are appropriate measures for students from Kindergarten through Grade 12, as well as adult learners.

GMRT measures overall reading achievement spanning each important stage along the reading continuum—from oral language to mature reading comprehension. The underlying foundation of *GMRT* is that it is useful for teachers and schools to know the general level of reading achievement of individual students throughout their entire school careers. The objective information obtained from the tests, complemented by teachers' evaluations and other sources of information, can be a basis for:

- Organizing students into appropriate instructional groups

- Selecting students for individual diagnosis and special instruction

- Identifying students who are ready for more advanced instruction

- Placing new students into proper instructional programs

- Evaluating the effectiveness of instructional programs

- Reporting student progress to parents, teachers, and the community

GMRT was constructed on the premise that reading is fundamental to developing an understanding of other subject matters. Thus, a student who struggles with reading will be likely to have difficulty completing a math story problem or comprehending a history lesson. Decisions on the content of the tests were based on current research in reading. The assessments offer:

- Vocabulary words that are appropriate to each grade level and chosen based on their usefulness to students

- Reading passages that are representative of students' actual reading in and outside of school

- Passages that are all taken from previously published books and periodicals

- Passage authors and content that fairly represent both genders and the ethnic diversity of the student population

- Passage content that includes fiction and non-fiction, science and social studies, and a wide variety of written styles

Each subtest in *GMRT* uses content that is age- and grade-appropriate and meaningful in the experience of students. Test results identify the strengths and needs of students throughout their development. Separate score information is provided for each subtest skill to allow for classroom decision-making and instructional planning.

Test Development and Design

GMRT was written by an exceptionally talented team of authors. **Walter H. MacGinitie, PhD**, is a former Professor of Psychology and Education at Teachers College, Columbia University, New York, and Lansdowne Scholar and Professor of Education at the University of Victoria, British Columbia. A past president of the International Reading Association and of the Reading Hall of Fame, he is an author of the first three editions of the *Gates-MacGinitie Reading Tests* and of numerous articles about reading.

Ruth K. MacGinitie has many years of experience as an author of educational materials and has worked with students who have reading comprehension difficulties. She played a major role in the development of the Second Edition of *GMRT* and is an author of the Third Edition of the test series.

Katherine Maria, PhD, is a former Professor in the Graduate Literacy Education Program at the College of New Rochelle, where she taught courses in children's literature, reading assessment, and comprehension instruction. Dr. Maria was an elementary school reading teacher for eight years. She has facilitated extensive, on-going staff development in reading in both urban and suburban school districts and has served as Professor in Residence in two elementary schools. She is the author of articles about reading and a text on comprehension instruction.

Lois G. Dreyer, PhD, is a former Professor of Reading and member of the Graduate Faculty at Southern Connecticut State University, where she was Director of the Reading Center and Coordinator of the Graduate Reading Program. She taught courses in diagnosis and remediation of reading difficulty and in reading research. Dr. Dreyer has been a classroom

teacher and school Reading and Language Arts Consultant in both urban and suburban settings. Her research and writing have focused on beginning reading, spelling, and the instructional needs of students experiencing reading difficulties.

Test Design

The validity of the *GMRT* is rooted in the overall design of the series, which measures the progression of students' skills in reading from Kindergarten through adulthood. Each level measures appropriate reading skills at the grade(s) for which it is intended, from background knowledge that is important for learning to read, to understanding appropriately sophisticated expository and narrative prose.

Reading vocabulary is assessed as it progresses from the use of important letter-sound correspondences (Levels PR and BR/Kindergarten–Grade 1), through reading sight words in context (Level BR/Kindergarten–Grade 1), to progressively more developed decoding and vocabulary (Levels 1 through 10/12/Grades 1–12). Reading comprehension is assessed as it progresses from understanding stories read aloud, to reading simple stories and expository text, to reading increasingly mature, age-appropriate text. The passages used throughout the *GMRT* are characteristic of texts students are learning to read in their respective grades and include a wide range of topics, representing a good balance of fiction, natural science, and social science as well as narrative and expository prose.

Scoring and Reporting

Paper/pencil responses to *GMRT* can be shipped to and scanned by the *HMH Scoring Service*, scanned and scored locally, or hand-scored on-site by teachers. When the tests are scored locally, teachers can make use of Booklet Scoring Keys that have been designed to be folded so that the answers for each test page can be placed next to the corresponding page in a test booklet as the teacher scores it. Class Summary Record forms are also available for recording individual scores and for calculating and recording group averages.

The score types available include national percentile ranks, national stanines, normal curve equivalents, grade equivalents, and extended scale scores for both the Vocabulary and Comprehension subtests as well as for Total scores. Local percentile ranks and local stanines are also available for Total scores when scoring is conducted through the *HMH Scoring Service*. Finally, Lexile measures are also available for all the levels of the Comprehension test. Norms are available based on the 2005–2006 school year renorming. All score reports are available online through HMH's *Interactive Results Manager* (*iRM*). Sample score reports are available in Appendix G of this submission.

Implementation and Technical Assistance

A key component of our approach to the management of an assessment program is a comprehensive system of customer support, which includes providing training in the administration of the test and in the use of the results. HMH's regional sales and support staff has extensive experience presenting and supporting training workshops for a wide variety of constituencies and groups. They will work with New York districts and Boards of Cooperative Educational Services (BOCES) that use the *Gates-MacGinitie Reading Tests* to determine the nature, purpose, and depth desired for pre- and/or post-test workshops.

Each New York Local Education Agency (LEA) will have the benefit of an experienced and knowledgeable Account Executive and the support of HMH's Customer Service and Technical Support departments.

Our Customer Service department serves as the primary point of contact for school personnel relative for support and assistance as they order, receive, and return testing materials, both those purchased from our catalog and those for large-scale assessment programs. In a recent year, HMH Customer Service handled over 100,000 inbound calls and processed nearly 60,000 orders for products and services. Our outstanding team of Customer Service Representatives is available by phone from 9 a.m. to 7 p.m. EST, Monday through Friday. Customers can also request assistance through fax, e-mail, or the HMH website. In addition to answering questions from schools and districts related to the procedures for the administration and handling of testing materials, Customer Service Representatives play a key role in ensuring test security by screening all requests for program-related assessment materials before they are shipped to individual schools.

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.
The *Gates-MacGinitie Reading Tests (GMRT)* provide appropriate metrics, such as scaled scores and normative data, to be used in multiple approaches for tracking and summarizing yearly growth in Reading. The Local Education Agency can use results from the initial administration to set an individual yearly Reading growth target for each student. Later in the year, scores from a second administration of the *GMRT* can be used to determine the percentage of students who met their individual growth targets for Reading. The LEA can convert those percentages to New York State's 0–20 Annual Professional Performance Review (APPR) Scale, which further segments these ratings into the four different reporting categories of educator effectiveness (Highly Effective, Effective, Developing, and Ineffective) in accordance with the existing rules for New York State's 0-20 metric.

New York State Next Generation Assessment Priorities Please provide detail on how the proposed supplemental assessment I 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 GMRT* is a battery of large-scale tests that assess students' skills in Reading. Characteristics of these assessments include content alignment and appropriateness, solid technical characteristics such as validity and reliability, and valuable information being reported to students and educators.
	The *GMRT* provides information that can improve instruction and influence student learning. Teachers can use test results to inform parents of an individual student's progress and to evaluate the progress of an entire class. Educators can monitor student growth by comparing results from multiple administrations. The *GMRT* reports student achievement, and it has been empirically validated for each of these purposes. The appropriate supporting documentation for each of these purposes can be found in the *GMRT Technical Report* in Appendix C*,* The *GMRT* is research-based and empirically validated and provides information in a fair, reliable and accurate manner. An integral component of *the GMRT* is the *GMRT* growth model.
	The *GMRT* growth model provides answers to important

	questions about student growth and changes to groups over time with a descriptive framework based on many years of research and development. Student growth information can be readily used for a variety of purposes in which the primary interpretation involves gain and improvement over time. Growth data based on the *GMRT* model are also amenable to various approaches for secondary analyses.
	The *GMRT* growth model uses an underlying vertical score scale, the Extended Scale Score (ESS), which permits several approaches to describing growth. It is a metric that ranges numerically from 74 to 714 and spans a developmental continuum from Kindergarten to Grade 12 in Reading.
	National research studies in the 1998–1999 school year were conducted to validate the reference points on the extended score scale representing the medians for each grade level and the model-based inferences about the amount of growth typical of students at different achievement levels. The primary interpretations supported by the SS scale have to do with: (1) how much a student is growing from one assessment occasion to the next compared to his or her assessment peers (a relative growth interpretation), and (2) how much growth would be expected for this student's assessment peers (a normative growth interpretation). This basic information about growth can be used for a variety of purposes in student and program evaluation such as individual and group decisions about instructional interventions, and responses to interventions that can be gauged by the amount of growth achieved.
	Another key feature of the *GMRT* growth model and its backbone, the ESS scale, is the ability to track student ESS values over time.
Assessments Woven Tightly Into the Curriculum:	Educators can use the results of *the GMRT* to improve instruction in a variety of ways. Teachers can use classroom-level results to gauge how well students comprehended particular content areas and adjust instruction accordingly. Growth results can be used for instructional planning and curriculum changes (e.g., in a Response to Intervention or RTI framework). This information can be used to set goals for the upcoming school year that will lead to student growth that exceeds expectations. Growth results can also help in determining professional development opportunities.
Performance Assessment:	Not applicable
Efficient Time-Saving Assessments:	The *GMRT* provides an effective and efficient assessment of reading. The *GMRT* balances efficient test administrations with rich reporting, including at the

	domain level. Paper-based with the availability of digital reporting provides flexibility in reporting options as well as faster turnaround times.
	Testing times for each *GMRT* level are listed in the Test Levels Times Chart, included in Appendix F of this submission. Note that for Levels PR and BR, there is no time limit. The teacher paces the students paces the students through the test at whatever rate is appropriate for the group. Testing times for these levels listed in the appendix are approximate for an average class; many classes will take a shorter or longer amount of time.
Technology:	*GMRT* offers flexible options for receipt of score reports. Customers may elect to receive digital reports through our *Interactive Results Manager*.
Degree to which the growth model must differentiate across New York State's four levels of teacher effectiveness (only applicable to supplemental assessments):	The essential pieces of this research work are already completed and validated on a national scale, including a technically sound and aligned test battery with a vertical scale.
	The process to determine an educator's effectiveness category using the *GMRT* is straightforward. First, student level growth scores from *the GMRT* are aggregated at the educator's level. These scores are then converted to New York State's 0–20 APPR scale in accordance with the existing rules for New York State's 0-20 metric.

STUDENT **A**SSESSMENTS FOR
TEACHER AND **P**RINCIPAL **E**VALUATION

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.	**X**
For supplemental assessments, this application contains a description of the growth model and how it is used in conjunction with the assessment.	**X**
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."	**X**
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.	**X**
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.	**X**
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.	**X**

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.*	
	Check all that apply:
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).	**X**
This application contains evidence of the *desired* criteria for reliability:	
• Standard errors provided for students growth scores.	**X**
• Student growth classifications have adequate decision consistency.	**X**
• Teacher effectiveness classifications demonstrate adequate consistency.	**X**
Examples include agreement statistics (e.g., kappa coefficients) based on simulation studies.	

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.*	
	Check all that apply:
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.*	**X**
This application contains evidence of the *desired* criteria for alignment validity:	
• 100% alignment between NYS Learning Standards and assessment.	☐

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.*	
	Check all that apply:
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).	☐
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.

	Check all that apply:
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).	**X**
This application contains evidence of the *desired* criteria for validity of internal structure: • Evidence students' scores are on an interval scale.	**X**
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.*	

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.

	Check all that apply:
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.	**X**
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.	**X**

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.	**X**
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).	**X**
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.	**X**
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.	**X**
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	**X**
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.)	**X**
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).	**X**
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.	**X**

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

Houghton Mifflin Harcourt Publishing Company 1. Name of Organization (PLEASE PRINT/TYPE)	 4. Signature of Authorized Representative (PLEASE USE **BLUE** INK)
Shawn Weirather 2. Name of Authorized Representative (PLEASE PRINT/TYPE)	August 24, 2016 5. Date Signed
Senior Director, Business Desk 3. Title of Authorized Representative (PLEASE PRINT/TYPE)	

N/A 1. Name of LEA (PLEASE PRINT/TYPE)	4. Signature of School Representative (PLEASE USE **BLUE** INK)
 2. School Representative's Name (PLEASE PRINT/TYPE)	5. Date Signed
 3. Title of School Representative (PLEASE PRINT/TYPE)	