# American Museum of Natural History MAT Residency Program TQP Grant: Year One External Evaluation Report

**Evelyn M. Gordon**
**P. Sean Smith**

**September 2020**

Submitted to: Rosamond Kinzler
American Museum of Natural History
Central Park West at 79th Street
New York, NY 10024

Submitted by: Horizon Research, Inc.
326 Cloister Court
Chapel Hill, NC 27514

# TABLE OF CONTENTS

***Page***

# Introduction

The Master of Arts in Teaching - Residency (MAT-R) program, based at the American Museum of Natural History (AMNH) in New York City, is a collaboration between educators and scientists at AMNH, and school districts in Yonkers, Queens, Brooklyn, and the Bronx. Horizon Research, Inc. (HRI) is conducting the external evaluation for AMNH's 2019–24 TQP grant to expand and innovate the residency program. The primary aims of the TQP project are to: (1) expand its Earth science student-teacher residency, mentoring, and induction program in high-need schools; (2) refine supports for culturally responsive teaching (CRT) in the teacher preparation curriculum; and (3) integrate Computational Thinking (CT) into the curriculum. The project has secured the services of Dr. Irene Lee as a consultant for leading the CT integration.

The Earth science teacher preparation program is based in AMNH's Richard Gilder Graduate School. The cohort-based, 14-month, 36-credit teacher residency program is designed to prepare and retain highly effective Earth science teachers to serve diverse student populations, including English language learners (ELLs) and students with special needs. Each cohort begins with a summer-long Museum Teaching Residency centered at AMNH. Program participants are then assigned two residencies in high-need public schools: one semester at a high school and one semester at a middle school. AMNH also trains and supports mentor teachers at the residency schools to provide a robust mentoring system for the residents. The program concludes with an AMNH-based Science Practicum Residency. After completing the residency program, MAT-R graduates are expected to begin their teaching careers at high-need schools and are supported by AMNH's two-year New Teacher Induction Program. The TQP grant will support residents in three new cohorts, as well as providing second-year induction activities for an earlier cohort.

Year One of the TQP project was notable for disruptions caused by the COVID-19 pandemic, including closures of the museum and schools. Despite the challenges these disruptions created, AMNH began the CT innovation process with an online faculty meeting focused on CT. At the meeting, Dr. Lee outlined the roots of CT in computer science and in computational science, its incorporation into K–12 education standards and teaching, and its use in the professional STEM workforce. Immediately following Dr. Lee's presentation, faculty members discussed the potential for using the Educational Global Climate Model[1] and OpenSpace[2] to develop residents' CT understanding. Plans for a second CT-focused faculty meeting were delayed so that faculty members could address urgent program adjustments due to the pandemic-related school and museum closures. In August and September 2020, AMNH engaged new and returning mentors in mentor training and support activities through an online adaptation of its Mentor Academy.

---

[1] http://edgcm.columbia.edu/

[2] https://www.openspaceproject.com/

# EVALUATION OVERVIEW

The external evaluation includes both formative and summative components. This section provides an overview of the formative and summative evaluations for the entire project, followed by a description of the Year One evaluation focus and activities.

## Formative Evaluation

The formative evaluation questions are aligned with the project's key activities, and Table 1 describes how the evaluation will address each question. As indicated by the data sources listed, the evaluation uses a multi-method, multi-source approach to addressing the questions.

**Table 1**
**Formative Evaluation Questions, Data Sources, and Timeline**

| Formative Evaluation Question | Data Sources | Years |
|---|---|---|
| 1. What are the nature, quality, and outcomes of the course revision process with respect to developing new CT components, refining CRT content, and developing additional guidance for supporting ELL students and students with special needs? | Document reviews, Meeting observations, Interviews with field test faculty and students | 1–2 |
| 2. To what extent is the project able to attract diverse, well-qualified applicants and select and enroll them as residents? | Demographic data for applicants and residents | 2–4 |
| 3. To what extent do AMNH and district superintendents function as partners to develop admissions goals and priorities, identify residency schools, and build partnerships with those schools? | Annual one-on-one interviews with leaders from AMNH and from partner districts and schools | 1–5 |
| 4. To what extent is the project able to attract, prepare, and support school-based mentors? | Teacher effectiveness data and subject area knowledge for mentor teachers, Observations of a sample of mentor program activities, Survey of all mentors, One-on-one interviews with project leaders from AMNH and a sample of residents, Focus group interviews with a sample of mentor teachers | 1–4 |
| 5. To what extent do clinical experiences focus on specific project objectives, including CRT and implementing CT activities? | Observations of a sample of clinical experiences, One-on-one interviews with a sample of residents and mentors, Survey of all residents | 3–5 |
| 6. To what extent do enacted course experiences align with project objectives and support residents' clinical experiences? | Observations of a sample of course meetings, Course evaluations, One-on-one interviews with a sample of residents, Survey of all residents | 3–5 |
| 7. To what extent does the induction program, including professional development opportunities, meet newly inducted teachers' needs? | One-on-one interviews with a sample of new teachers, Survey of all new teachers | 2, 4–5 |
| 8. In what ways and to what extent do residents and new teachers benefit from working with school-based and faculty mentors and coaching activities? | One-on-one interviews with a sample of residents and new teachers, Survey of all residents and new teachers | 3–5 |
| 9. In what ways and to what extent do residents and new teachers benefit from being part of a cohort? | One-on-one interviews with a sample of residents and new teachers, Survey of all residents and new teachers | 3–5 |

## Summative Evaluation

The summative evaluation will focus on project outcomes and impacts. MAT-R's goals include specific targets for persistence in the program, certification, high-need school placement, and teacher retention (Summative Question 2). The project also aims to positively impact graduates' preparedness as Earth science teachers, including their preparedness to use CRT practices and implement CT activities. HRI will collect data annually on each of these outcomes (see Table 2 for details).

**Table 2**
**Summative Evaluation Questions, Data Sources, and Timeline**

| Summative Evaluation Question | Data Sources | Years |
|---|---|---|
| 1. Did the project achieve its recruitment target of 72 residents, 24 of whom identify as Hispanic and/or non-white for the MAT-R program? | Demographic data for residents who enroll in the program | 3–5 |
| 2. Did the project achieve its preparation, certification, and high-need school hiring target rate of greater than 90 percent and its 3-year retention rate of greater than 80 percent? | Certification/licensure outcomes (Performance Measure 1), Program graduation results (Performance Measure 2), One-year persistence rate among any residents that do not graduate in 14 months (Performance Measure 3), Hiring rate (high-need LEA and overall), One-year employment retention (Performance Measure 4), Three-year employment retention (Performance Measure 5) | 3–5 |
| 3. What is the impact of the MAT-R program on residents' preparedness to (a) teach science effectively to high-need underserved students, including ELL students and special education students; (b) use CRT practices; and (c) implement CT activities? | One-on-one interviews with a sample of residents and new teachers, Survey of all residents, AMNH faculty observation scores, Mentor teacher observation scores | 3–5 |
| 4. What is the impact of the MAT-R program on graduates' preparedness to use CRT practices and implement CT activities, and to teach underserved students, including ELL students and special education students? | One-on-one interviews with a sample of new teachers, Survey of all new teachers, Hiring principal surveys | 4–5 |
| 5. What is the impact of MAT-R program graduates on high-need schools' performance in Earth science? | Results of quantitative comparisons of Earth Science Regents performance in schools with and without MAT-R program graduates, conducted by project partners at NYU (Performance Measure 6) | 4–5 |

## Year One Evaluation Activities

In Year One, HRI focused on becoming familiar with the current program and addressing formative evaluation question 1 (What are the nature, quality, and outcomes of the program tool and course revision processes with respect to developing new CT components, refining CRT content, and developing additional guidance for supporting ELL students and students with special needs?). HRI also collected initial evaluation data related to formative evaluation question 4 (To what extent is the project able to attract, prepare, and support school-based mentors?).

To address question 1, HRI met with project leaders periodically by videoconference, observed a sample of MAT-R faculty meetings by videoconference, conducted a survey focused on faculty members' CT instruction and perceptions of preparedness for CT instruction, and reviewed documents related to courses targeted for innovation. To address question 4, HRI observed two mentor preparation sessions by videoconference. In addition to this annual report, HRI provided formative evaluation feedback based on the faculty survey via a videoconference.

### *Faculty Computational Thinking Survey*

Following the May faculty meeting at which Irene Lee presented definitions and a brief history of CT, HRI administered a survey about CT instruction to faculty members (see Appendix A for the survey instrument). The survey was designed to encourage faculty members to reflect on how their instruction currently addresses CT and potential areas for increasing attention to CT. Open-ended questions addressed both of these areas, and faculty were also asked to identify challenges to increasing attention to CT in the program.

The online survey was sent to 34 MAT-R faculty members, administrators, and staff. Of these, 19 were course instructors, including 4 who also taught in a residency,[3] 8 taught only in a residency, and 7 did not teach a course or in a residency. HRI received completed surveys from 18 faculty members, and response rates by teaching roles are shown in Table 3

**Table 3**
**Faculty CT Survey Response Rates, by Teaching Role**

| | Teaches MAT course(s) | Teaches MAT residency† | Does not teach MAT course or residency | Total |
|---|---|---|---|---|
| Invited | 19 | 8 | 7 | 34 |
| Responded | 11 | 4 | 3 | 18 |
| Response rate | 58% | 50% | 43% | 53% |

† School-based residency, museum teaching residency, science research practicum, or induction; does not include faculty who teach MAT course(s).

HRI coded responses to the open-ended questions to identify common themes and any contradictory evidence to those themes across respondents. Responses to the scale items were tested to evaluate the items' reliability as a possible composite measure,[4] and descriptive

3 "A residency" means one or more of: school-based residency, museum teaching residency, science research practicum, or induction. Because fewer than five faculty members taught both course(s) and in a residency, their responses are not reported separately to maintain participant confidentiality.

4 A composite is more reliable than individual survey items. The composite is calculated by summing the responses to the items associated with the composite and dividing by the total points possible. To create a 100-point scale for the composite, the lowest response option is set to 0 with the other options adjusted accordingly, and the result of the division is multiplied by 100. The Faculty CT composite, with 7 items and a scale ranging from 1 to 4, was adjusted to have a scale of 0 to 3, making the total possible points 21. Using this approach, someone who marks the lowest point on every item in a composite receives a composite score of 0 (rather than some positive number). It also assures that 50 is the true midpoint.

statistics were calculated for the scale items. Results of these analyses are presented in the Findings section.

# FINDINGS

## *Faculty feel prepared to develop residents' understanding of CT, but less prepared to develop residents' readiness to teach CT or assess residents' CT understanding or teaching skills.*

Frequencies for faculty responses to the scale items are shown in Table 4. The majority of respondents agreed (somewhat or strongly) to statements related to their own understanding or use of CT and their preparedness to develop residents'[5] understanding or use of CT. However, the majority of respondents disagreed (somewhat or strongly) to statements related to their preparedness to develop residents' skills for teaching CT or to assess candidates' understanding of CT. This suggests that, as the program increases attention to CT, faculty may particularly benefit from assessment resources and professional learning about how to develop residents' skills for teaching CT.

**Table 4**
**Faculty Perceptions of CT Understanding and Preparedness**

| | Percent of Respondents (N = 17†) | | | |
|---|---|---|---|---|
| | Strongly Disagree | Somewhat Disagree | Somewhat Agree | Strongly Agree |
| I frequently use elements of computational thinking in my own research. | 18 | 18 | 24 | 41 |
| I find it easy to identify elements of computational thinking in scientific endeavors, whether my own or others. | 0 | 6 | 59 | 35 |
| I am prepared to engage candidates in using computational thinking as part of scientific inquiry. | 6 | 24 | 41 | 29 |
| I am prepared to develop candidates' skills for teaching computational thinking. | 6 | 47 | 24 | 24 |
| I am prepared to develop candidates' understanding of computational thinking. | 0 | 29 | 53 | 18 |
| I have strategies for assessing candidates' understanding of computational thinking. | 18 | 41 | 41 | 0 |
| I have strategies for assessing candidates' understanding of how to teach computational thinking. | 24 | 71 | 6 | 0 |

† One faculty member without teaching responsibilities did not respond to the scale items.

The seven scale items taken together form a reliable composite (Cronbach's Alpha = .840). The mean composite score across all respondents was 55.74 (out of 100 possible), with a minimum of 24, a maximum of 90, and a standard deviation of 19.54. The reliability of the composite and the

[5] This report uses AMNH's preferred term for students in the MAT-R program, which is "resident." However, the term "candidate" was used in the faculty CT survey and therefore appears in Table 4.

range of scores suggest that the composite may be useful for measuring changes over time in the faculty's readiness to address CT.

### *Faculty currently address CT-related content in their instruction and identified opportunities to increase attention to CT.*

The open-ended survey responses indicated several key ways that faculty currently address CT-related concepts. Most commonly, instructors engage residents in using CT (mentioned by 11 respondents) through some combination of using models (8 respondents), using online tools such as simulations or large datasets (7 respondents), and collecting or interpreting data (6 respondents). These strategies were also evident in HRI's review of assignments for the two courses targeted for CT innovation. The following survey response excerpts indicate the types of activities that many respondents described:

> *1. Students carry out data analysis exercises using tools like spreadsheets for data generated from physical experiments, observations (by themselves or others), and models. 2. Students use existing computational models to generate data (e.g. planetarium software that recreates sky positions over time). 3. Students retrieve data from existing databases of observations for analysis.* [Course Instructor]

> *In the field experiences, now virtual, we prioritize careful observation by candidates of natural phenomena . . . . We then ask them to question: Why is this so? Is it expected? These questions lead to additional observation to clarify inquiry. Observations lead to building a mental model.* [Residency Instructor]

In addition, 5 respondents described opportunities for residents to implement, discuss, or observe activities for engaging K–12 students in CT, and 4 respondents described connections between course activities and Next Generation Science and Engineering Practices related to CT. Although 4 respondents noted that they did not explicitly address CT in their instruction, 2 of these respondents described activities similar to those described as CT by other respondents, for example:

> *One of our course sessions focuses on the scientific practice of analyzing and interpreting data and teaching and learning opportunities with that practice. . . . This isn't currently explicitly connected to computational thinking, and it's a weak connection—but it's one assignment in [the course] where they are having to think about abstracting and creating a model.* [Course Instructor]

In general, respondents' ideas for increasing attention to CT in their instruction mirrored the current types of activities; one notable addition (mentioned by 2 respondents) was engaging residents with coding or examining code. Additional opportunities for engaging residents in using CT by using models, using online tools such as simulations or large datasets, and collecting or interpreting data were each described by 5 respondents. In addition, 3 respondents suggested that currently implicit CT connections could be made more explicit by pointing out connections to CT and/or deliberately using CT terminology.

### *Time for addressing CT, resident preparedness, and a coherent vision among faculty may present challenges to increasing attention to CT.*

Time was the most frequently mentioned challenge to increasing attention to CT to the MAT-R curriculum, specifically, creating time for residents to engage with additional content (8 respondents). The following response is representative of those raising concerns about time:

> *Time. Our program is already overloaded, and other emphases would have to be sacrificed to make this possible.*

Several additional challenges were raised by multiple respondents. Some respondents suggested that residents' readiness to engage with CT concepts may present a challenge because they will not perceive CT as important or relevant (4 respondents). Two of these respondents also raised concerns that not all residents have prerequisite mathematical knowledge, and another identified limited computer skills as a possible challenge. In addition, faculty members developing a coherent understanding of CT and common language around CT, including the need to express implicitly connected ideas more explicitly, was raised by 5 respondents, and two others shared a related concern that faculty members may need to develop their own understanding before they can effectively teach CT.

Dr. Lee's presentation to the faculty provided a foundation for addressing these last concerns about developing a shared and coherent understanding of CT among faculty members. At the same time, Dr. Lee identified various definitions and priorities related to CT historically and across related fields, and the Next Generation Science and Engineering Practices offer yet another perspective, illustrating the challenge the program will face in forging a shared vision.

## Summary and Recommendations

Despite unprecedented challenges to AMNH's and schools' functioning created by COVID-19, the MAT-R program was able to begin the process of developing a shared vision for CT innovations in Year One. A significant strength for the innovation process going forward is that current, pre-innovation instruction includes significant attention to CT-related concepts. However, not all CT-related content is explicitly connected to CT in current instruction, and faculty may not have a coherent, shared vision for making such connections. Furthermore, finding time to address CT in the curriculum may be challenging. Based on self-report, faculty appear better prepared to address residents' understanding of CT than their understanding of CT instruction.

- *Recommendation: Consider engaging a small team of faculty members to develop an initial vision for CT in the MAT-R to share and refine with the full faculty.*
  Having a small team develop an initial vision for CT in the MAT-R and refining the vision with the full faculty is likely to be more efficient than engaging the full faculty at the beginning of the development process. In addition to project leaders and guidance from Dr. Lee, it may be useful to include a few motivated and knowledgeable faculty members on the smaller team. Elements of the initial vision might include:

  - A framework or definition of CT for MAT-R use, including connections to the Next Generation Science and Engineering Practices;
  - Outcome goals for developing residents' understanding of CT content; and
  - Outcome goals for developing residents' skills for teaching CT.

- *Recommendation: To target innovation priorities, consider mapping existing CT-related activities across the curriculum and comparing them to outcome goals for developing residents' understanding of CT content and skills for teaching CT.*
  The MAT-R faculty periodically review residents' assignments across courses and residencies to manage the residents' workload, and they have previously compared and coordinated attention to specific content across courses, residencies, and workshops. These established strategies may be useful for identifying current CT-related instruction to better understand the extent to which outcome goals are already being addressed and opportunities to refine or add instructional activities to address goals. Tying revised or new CT-related activities to existing activities may also help to relieve faculty concerns about the time needed to address additional learning goals in the program.

# APPENDIX:
# FACULTY COMPUTATIONAL THINKING SURVEY

**MAT-R Faculty Computational Thinking Survey**

Horizon Research, Inc. is the external evaluator for AMNH's 2019-2024 TQP grant. As part of the evaluation, we are asking all MAT-R program faculty members to complete a brief survey about your thinking about how the role of computational thinking in the program is evolving.

This survey is intended as a tool for thinking and reflecting upon computational thinking in the program in preparation for discussion at a future faculty meeting. It provides each individual with an opportunity to reflect on computational thinking and to share your current thoughts about how computational thinking might fit into your work with candidates. It also provides a starting point for conversations and thinking about CT as a faculty. Collectively, the reflections will provide a snapshot of how computational thinking is addressed in the program at this moment. In addition, de-identified results from the survey will be used to provide formative evaluation feedback, which will be shared with MAT program leaders and at a future faculty meeting.

As computational thinking can have multiple definitions, we provide descriptions both by Dr. Irene Lee (from her presentation at the May faculty meeting) and of Next Generation Science Standards (NGSS) practices that we ask you to consider. The next page includes brief descriptions of NGSS science and engineering practices related to computational thinking and key points from Dr. Lee's talk. Please refer to them as needed.

[new page]

Computational thinking is addressed in the Next Generation Science Standards as part of Science and Engineering Practices 2 and 5:

> Practice 2 Developing and Using Models: A practice of both science and engineering is to use and construct models as helpful tools for representing ideas and explanations. These tools include diagrams, drawings, physical replicas, mathematical representations, analogies, and computer simulations.
>
> Practice 5 Using Mathematics and Computational Thinking: In both science and engineering, mathematics and computation are fundamental tools for representing physical variables and their relationships. They are used for a range of tasks such as constructing simulations; statistically analyzing data; and recognizing, expressing, and applying quantitative relationships.

Dr. Lee shared the following definition of computational thinking at the May 7 faculty meeting:

> "Computational Thinking is the thought processes involved in formulating problems and their solutions so that the solutions are represented in a form that can be effectively carried out by an information-processing agent." (Cuny, Snyder, & Wing, 2010, cited in Wing, 2011; Aho, 2011)

Dr. Lee described abstraction, automation, and analysis as three pillars of computational thinking. She also identified several elements that bridge computational thinking in K-12 education and in the work of STEM professionals, including: understanding complex systems, innovating with representations, modeling and simulation, collective sensemaking and crowd sourcing data, and understanding consequences and making predictions.

Click here if you would like to open a pdf version of this information.

[new page]

1. How, if at all, do you incorporate computational thinking in your course(s) or other work with candidates? (Describe relevant discussions, activities, resources, assignments, etc., and how each addresses computational thinking in a few sentences or enter "none." Please indicate if any of the assignments are expected to be carried out in the school residency.)

| |
|---|
| |

2. In any course(s) you teach and your work with candidates in other settings, what opportunities, if any, do you see for adjusting course goals or for including or adding learning opportunities related to computational thinking? (Describe potential changes that could increase attention to computational thinking. This does not commit you to make these changes!)

| |
|---|
| |

[new page]

3. Considering the MAT-R program and our preparation of candidates more generally, what challenges do you anticipate for increasing attention to computational thinking?

| |
|---|
| |

[new page]

4. Please rate the degree to which you agree or disagree with each statement.

| | **Strongly Disagree** | **Somewhat Disagree** | **Somewhat Agree** | **Strongly Agree** |
|---|---|---|---|---|
| a. I find it easy to identify elements of computational thinking in scientific endeavors, whether my own or others'. | | | | |
| b. I frequently use elements of computational thinking in my own research. | | | | |
| c. I am prepared to engage candidates in using computational thinking as part of scientific inquiry. | | | | |
| d. I am prepared to develop candidates' understanding of computational thinking. | | | | |
| e. I am prepared to develop candidates' skills for teaching computational thinking. | | | | |
| f. I have strategies for assessing candidates' understanding of computational thinking. | | | | |
| g. I have strategies for assessing candidates' understanding of how to teach computational thinking. | | | | |