How Youth Experience Work With Data in Summer STEM Programs: Findings From An Experience Sampling Method Approach

|  |
| --- |
| Joshua M. Rosenberg1, Jennifer A. Schmidt2, and Matthew J. Koehler2 |
| 1University of Tennessee, Knoxville |
| 2Michigan State University |
|  |

# Author note

Joshua M. Rosenberg is an Assistant Professor in the Department of Theory and Practice in Teacher Education, College of Education, Health, and Human Sciences, University of Tennessee, Knoxville. Jennifer A. Schmidt is an Associate Professor in the Department of Counseling, Educational Psychology, and Special Education, College of Education, Michigan State University. Matthew J. Koehler is a Professor in the Department of Counseling, Educational Psychology, and Special Education, College of Education, Michigan State University.

This study was completed as part of National Science Foundation Grant No. DRL-1421198. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not reflect the views of the National Science Foundation. Correspondence concerning this article should be addressed to Joshua M. Rosenberg, 1122 Volunteer Blvd., Knoxville, TN 37919. E-mail: [jmrosenberg@utk.edu](mailto:jmrosenberg@utk.edu)

Abstract

Working with data – forming questions, constructing measures and attending to precision, and creating models and representations of data, for example –can serve as a context for learning across STEM content areas. Furthermore, becoming proficient in work with data can be personally empowering because of the parts of our lives–from paying energy bills to interpreting news articles–that use data. While there has been research on work with data in mathematics content areas, less has focused on such practices in science and engineering educational settings. The present study explores learners’ engagement in work with data in the context of nine summer STEM programs. Data from measures of learners’ engagement was collected through the Experience Sampling Method (ESM) and was collected from more than 200 youth in nine summer STEM programs over four weeks in the Northeastern United States. Findings show that aspects of work with data were fairly common in the programs overall. Six profiles of youth engagement were identified, representing distinct configurations of the five indicators of engagement. Relations between the profiles of engagement and each of the aspects of work with data were small: Notable exceptions were the generating data and data modeling were significantly associated with full engagement. Youth with higher pre-program interest in STEM were more likely to be engaged and competent but not challenged, though other youth characteristics were not highly related to the profiles. Implications of the findings and the implications for practice with respect to work with data in general and to engagement in informal learning environments, such as summer STEM programs, in both cases with an emphasis on how work with data can serve as a promising context for learning in STEM subject areas.

*Keywords:* Work with data, data science education, experience sampling method, engagement, out-of-school STEM programs

Understanding Work With Data in Summer STEM Programs Through An Experience Sampling Method Approach

In modern society, data are all around us, shaping how we work, how we socialize, how we teach, and how we learn. These sources of data are created by us, for us, and about us. Through work with data learners transform from consumers of knowledge to creators of knowledge (Hancock, Kaput, & Goldsmith, 1992; Lehrer & Schauble, 2015; Lee & Wilkerson, 2018; Finzer, 2013). Work with data empowers learners to ask meaningful questions and to answer them with arguments and explanations that draw from data as evidence (McNeill & Krajcik, 2007). This work, then supports learners to create new knowledge in learning environments and classrooms = an aim of recent reform efforts that cast a vision of learning that emphasizes participation in the practices of STEM disciplines (NGSS Lead States, 2013; National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010).

Work with data is more than just crunching numbers or interpreting a figure created by someone else. It refers to a number of broad processes aimed at making sense of phenomena or solving problems in the world. Work with data includes collecting, creating, and modeling data, as well as generating questions that can be answered with data. This focus on phenomena is particularly relevant to those designing and enacting learning opportunities focused on work with data (Lee & Wilkerson, 2018; Singer, Hilton, & Schweingruber, 2006; Wild & Pfannkuch, 1999).

Work with data provides a capability that can be used across content areas, particularly in advanced coursework. Aspects of work with data are recognized as core competencies across recent curricular documents for STEM subject area learning. They are found, for example, in the *Next Generation Science Standards* and the *Common Core State Standards* for mathematics. These standards highlight the role of authentic work with data as part of engaging in scientific and engineering and mathematical practices, respectively. These capabilities may be particularly useful in STEM domains because advanced coursework in these domains often involves demanding and abstract work with data, work that may be more accessible to more learners when they encounter it earlier in their education.

Past research on work with data has mostly been set in mathematics contexts and has focused on mathematical practices, like generating measures of phenomena and creating data models (English, 2012; Lehrer & Romberg, 1996; Lesh, Middleton, Caylor, & Gupta, 2008). It has often focused on specific cognitive outcomes (e.g., Gelman & Markman, 1987), strategies to support work with data (Petrosino, Lehrer, & Schauble, 2003), and some opportunities and challenges facing both teachers and learners when working with data (e.g., Konold & Pollatsek, 2002; Finzer, 2013). There has been some research about work with data in science settings, too. However, scholarship has pointed out that what it means to work with data can vary greatly in actual classrooms and other learning environments (McNeill & Berland, 2017). Even so, this past research broadly suggests that engaging in work with data is powerful concerning learning both about and how to do mathematics and science (Lee & Wilkerson, 2018; Lehrer & Schauble, 2015). Lehrer and Schauble (2015), summarizing past research on the use of mathematical practices in science contexts, note that work with data “has an exceptionally high payoff in terms of students’ scientific reasoning” (p. 696).

To date, past research shows that using a framework from contemporary engagement theory to characterize students’ experiences has been informative both in research and to practicing educators. Work with data is similar to hands-on, laboratory work which research has shown to be engaging to students (Schmidt, Rosenberg, & Beymer, 2018). In addition, work with data is demanding and requires sustained effort and focus (Lehrer & Schauble, 2015; National Research Council, 2015), and past work has shown that when learners are more challenged (and competent), they are more likely to be engaged (Schneider et al., 2016; Shernoff et al., 2016). Knowing more about how youth engage in work with data is valuable as engagement is a meaningful outcome for STEM learners in its own right (Sinatra, Heddy, & Lombardi, 2015). It may also be an antecedent of changes in other outcomes, such as their well-being, achievement, and the pursuit of an area of study or career (Wang & Eccles, 2012). However, research has not examined how students engage in work with data. Because engaging in work with data seems to be so potentially beneficial to learners, better understanding the nature of work with data and learners’ engagement in such practices is needed.

The purpose of this study, then, is to examine youth engagement in a variety of learning activities that involve work with data. We explore youths’ engagement in the context of outside-of-school STEM enrichment programs carried out during the summer. Such settings are an especially useful context for exploring work with data because they can be designed around youths’ interests (Lauer, Akiba, Wilkerson, Apthorp, Snow, & Martin-Glenn, 2006). One promise of work with data in outside-of-school settings is that relevant sources of data can be inherently interesting to learners. Such sources of data can be used as a context for learning about the world, allowing youth to ask and answer personally and socially meaningful questions, whereas many outside-of-school programs are focused around commercial aims, such as developing mobile device applications. Knowing more about how youth engage can also provide a foundation for subsequent work to explore how particular curricula and engaging experiences for youth spark their interest in work with data, including hobbies and occupations related to data science, but also in STEM domains in general.

# Defining Work with Data

Work with data has been conceived in different ways (i.e., Hancock et al., 1992; Lehrer & Romberg, 1996; Wild & Pfannkuch, 1999). For instance, Wild and Pfannkuch (1999) consider the process in terms of identifying a problem, generating a measurement system and sampling plan, collecting and cleaning the data, exploring the data and carrying out planned analyses, and interpreting the findings from the analysis. Such a process is common in STEM content areas and is instantiated in standards for some (especially mathematics) curricula. Franklin et al.’s guidelines focus on the Framework for statistical problem solving: formulating questions, collecting data, analyzing data, and interpreting results (2007). The goals of this framework and its components are similar to Hancock et al.’s (1992) description of data modeling, the process of “using data to solve real problems and to answer authentic questions” (p. 337). Hancock et al. (1992) focus in on two goals, data creation and analysis, arguing that the former (data creation) is under emphasized in classroom contexts. Scholars have subsequently expanded Hancock et al.’s definition of data modeling to include six components: asking questions, generating measures, collecting data, structuring data, visualizing data, and making inferences in light of variability (see Lehrer & Schauble, 2004, for use of this conceptualization of data modeling applied to the task of understanding plant growth). The last of these components is crucial across all of the visions of data modeling reviewed here and distinguishes these processes from other aspects of data analysis: Accounting for variability (or uncertainty) is central to solving real-world problems with data and the process of data modeling.

Because there is not an agreed-upon definition of work with data–particularly across subject area domains (i.e., across all of the STEM content areas)–WE focus on the core aspects that scholars have most often included in their conceptualizations of work with data. These core components, synthesized from definitions across studies, are better for understanding work with data across STEM content areas–as in the present study–than the components from specific examples, which were developed for use in only one domain. The aspects of work with data that have been articulated in prior studies are distilled into five key aspects for use in this study. They are:

* *Asking questions*: Generating questions that can be answered with empirical evidence
* *Making observations*: Watching phenomena and noticing what is happening concerning the phenomena or problem being investigated
* *Generating data*: The process of figuring out how or why to inscribe an observation as data about phenomena, as well as generating tools for measuring or categorizing
* *Data modeling*: Activities involving the use of simple statistics, such as the mean and standard deviation, as well as more complicated models, such as linear models and extensions of the linear model
* *Interpreting and communicating findings*: Activities related to identifying a driving question regarding the phenomena that the question is about

These aspects of work with data are not stand-alone practices but are a part of an iterative cycle. For example, interpreting findings leads to new questions and subsequent engagement in work with data. Also, scholars have pointed out some key features of how work with data is carried out that impact their effectiveness as a pedagogical approach. These key features include an emphasis on making sense of real-world phenomena and iterative cycles of engaging in work with data and collaboration and dialogue, through which ideas and findings are critiqued and subject to critique, and revised over time (McNeill & Berland, 2017; Lee & Wilkerson, 2018).

# What is Known About How Youth Work with Data

There is a good amount of past research on cognitive capabilities as outcomes from working with data. Much of this—laboratory-based—research has focused on how children develop the capability to inductively reason from observations (Gelman & Markman, 1987). Other research has focused on the development of causal, or mechanistic, reasoning, among young children (Gopnik & Sobel, 2000; Gopnik, Sobel, Schulz, & Glymour, 2001). A key outcome of engaging in work with data has to do with how learners account for variability (Lehrer, Kim, & Schauble, 2007; Petrosino et al., 2003; Lesh, Middleton, Caylor, & Gupta, 2008; Lee, Angotti, & Tarr, 2010), arguably the main goal of engaging in work with data (Konold & Pollatsek, 2002). From this research, we know that learners can develop the capacity to reason about variability.

Past research has also shown that there are strategies that can support work with data. These include the design of technological tools and the development of curricula. From this research, we know about specific strategies and learning progressions for learners to develop this capability. For example, past research has illustrated the role of measurement in exposing learners in a direct way to sources of variability (Petrosino et al., 2003) or the place of relevant phenomena, such as manufacturing processes, such as the size of metallic bolts, which can help learners to focus on “tracking a process by looking at its output” (Konold & Pollatsek, 2002, p. 282).

Finally, past research has shown that different aspects of work with data pose unique opportunities and challenges. Asking empirical questions requires experience and ample time to ask a question that is both able to be answered with data and which is sustaining and worth investigating (Bielik & Yarden, 2016; Hasson & Yarden, 2012). Making observations and generating data, such as of the height of the school’s flagpole, requires negotiation not only of what to measure, but how and how many times to measure it (Lehrer, Kim, & Schauble, 2007). Regarding modeling, not only teaching students about models, such as that of the mean, but also asking them to create them, are valuable and practical (Lehrer & Schauble, 2004; Lehrer, Kim, & Jones, 2011), but also time-intensive. Interpreting findings, especially in light of variability through models, and communicating answers to questions, means not only identifying error but understanding its sources, and can be supported through exploring models that deliberately represent the data poorly, but can be instructive for probing the benefits and weaknesses of models (Konold & Pollatsek, 2002; Lee & Hollebrands, 2008; Lehrer, Kim, & Schauble, 2007).

Though very valuable past research has been carried out, valuable insight into how learners and youth participate in different aspects of work with data through the lens of engagement has not been explored. This work can compliment past research by showing, for instance, how certain strategies of work with data or how enacting aspects of work with data in particular ways engage learners (two foci of past research). Consider the practice of modeling data, commonly described as a-or *the*-key part of many data analyses (Konold, Finzer, & Kreetong, 2017). When modeling data, learners may use data they generated and structured in a data set on their own or may model already-processed, or use already-plotted, data (McNeill & Berland, 2017). How challenging do students perceive the different enactments of these activities to be and how do learners perceive their competence regarding them? Importantly, how hard are learners working? How much do they feel they are learning? Knowing more about these beliefs, characteristics, and processes could help us to develop informed recommendations for teachers and designers intending to bring about opportunities for learners to engage in work with data in a better-supported way that is sustained over time.

**Engagement in General and in STEM Content Areas**

Engagement is defined in this study as active involvement, or investment, in activities (Fredricks, Blumenfeld, & Paris, 2004). Explaining how learners are involved in activities and tasks is especially important if we want to know about what aspects of work with data are most engaging (and in what ways), and therefore can serve as examples for others advancing work with data as well as those calling for greater support for engagement. Apart from being focused on involvement, engagement is often thought of as a meta-construct, that is, one that is made up of other constructs (Skinner & Pitzer, 2012; Skinner, Kindermann, & Furrer, 2009). By defining engagement as a meta-construct, scholars characterize it in terms of cognitive, behavioral, and affective dimensions that are distinct yet interrelated (Fredricks, 2016).

We know from past research that the cognitive, behavioral, and affective dimensions of engagement can be distinguished (Wang & Eccles, 2012; Wang & Holcombe, 2010) and that while there are long-standing concerns about the conceptual breadth of engagement (Fredricks et al., 2016), careful justification and thoughtful use of multidimensional engagement constructs and measures is warranted. Engagement is also considered to be changing in response to individual, situation or moment contextual factors, Skinner and Pitzer's (2012) model of motivational dynamics, highlighting the community, school, classroom, and even learning activity, shows the context-dependent nature of engagement on the basis of the impacts of these factors on learners' engagement.

Engagement in STEM settings shares characteristics with engagement across disciplines, yet there are some distinct aspects to it (Greene, 2015). While one type of engagement—behavioral—is associated with achievement-related outcomes, many STEM practices call for engagement in service of other outcomes, especially around epistemic and agency-related dimensions (Sinatra et al., 2015,). For example, many scholars have defined scientific and engineering practices as cognitive practices, which involve applying \*epistemic considerations\* around sources of evidence and the nature of explanatory processes (see Berland et al. 2016, Stroupe, 2014).

The emphasis on developing new knowledge and capabilities by engaging in STEM practices must be reflected in how the cognitive dimension of engagement is measured. Because of the importance of constructing knowledge to engagement in STEM practices, then, I define cognitive engagement in terms of learning something new or getting better at something. While sometimes defined in terms of extra-curricular involvement or following directions, I define behavioral engagement in this study as working hard on learning-related activities (Fredricks et al., 2004; Singh, Granville, & Dika, 2002). Finally, I define affective engagement as emotional responses to activities, such as being excited, angry, or relaxed (Pekrun & Linnenbrink-Garcia, 2012).

Finally, some critical conditions facilitate engagement. Emergent Motivation Theory (EMT; Csikszentmihalyi, 1990), provides a useful lens for understanding these conditions. From EMT, a critical condition for engagement that can change dynamically, from moment to moment, is how difficult individuals perceive an activity to be, or its \*perceived challenge\*. Another critical condition is how good at an activity an individual perceives themselves to be, or their \*perceived competence\*. What is most important--and necessary concerning being engaged--is being both challenged by and good at a particular activity.

Past research has supported this conjecture (Csikszentmihalyi, 1990). As one empirical example, Shernoff et al. (2016) demonstrated that the interaction of challenge and competence was associated with positive forms of engagement. These findings suggest that learners’ perceptions of the challenge of the activity, and their perceptions of how skillful they are, are important conditions that co-occur with learners' engagement. Conceptualizing perceptions of challenge and competence as conditions, rather than factors that influence engagement, is in recognition of their co-occurrence within individuals, in that youth experience engagement and their perceptions of the activity (perceived challenge) and of themselves (perceive competence) together and at the same time. Thus, these two conditions (challenge and competence) are considered together with engagement in this study, as described in the section below on analyzing multidimensional data on engagement.

## Need for the Present Study

While many scholars have argued that work with data can be understood in terms of the capabilities learners develop and the outcome learners achieve, there is a need to understand learners’ experiences working with data. The present study does this in terms of contemporary engagement theory (Fredricks, Blumenfeld, & Paris, 2014). Doing this can help us to understand work with data in terms of learner’s engagement, which we know from past research impacts what and how students learn (Sinatra et al., 2015). Knowing more about students’ engagement can help us to design activities and interventions focused around work with data. In particular, an Experience Sampling Method (ESM; Hektner, Schmidt, & Csikszentmihalyi, 2007) approach may be useful for measuring youths’ engagement not before or after their experiences, but rather during the specific activities that involve work with data. In addition to this need to study learners’ experience working work with data through the lens of engagement theory, no research has yet examined work with data in the context of summer STEM programs, though such settings are potentially rich with opportunities for highly engaged youth to analyze authentic data sources.

## Conceptual Framework and Research Questions

To summarize in Figure 1, engagement in work with data is associated with different profiles of engagement. The theoretical framework for the profile approach suggests that engagement is a multi-dimensional construct consisting of cognitive, behavioral, and affective dimensions of engagement and perceptions of challenge and competence. The five aspects of work with data and youths’ pre-program interest, gender, and URM status are predictor variables at the youth level.

The four research questions, then, are:

1. How frequent is work with data in summer STEM programs?
2. What profiles of engagement emerge from data collected via ESM in the programs?
3. How does work with data relate to the profiles of engagement?
4. How do youth characteristics relate to profiles of engagement?

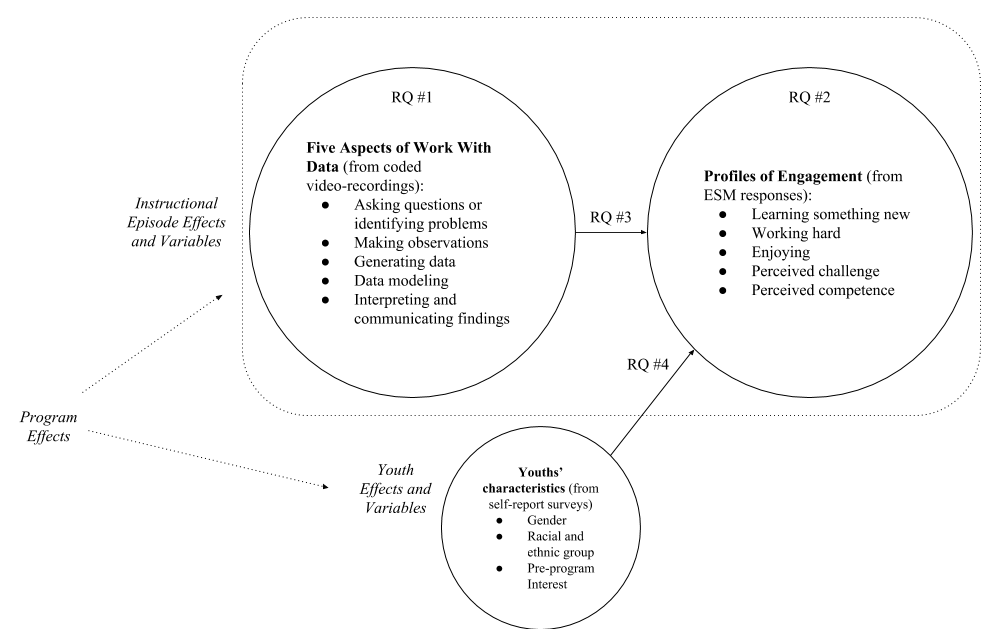


Figure 1. A conceptual framework for this study and research questions.

# Methods

## Context

The setting for the present study was nine out-of-school STEM programs during 2015 in the Northeast United States. Descriptions of the programs are provided in Appendix A. Two *intermediary organizations* which were contracted by the local school districts to administer the summer programs. The two intermediaries were responsible for soliciting and enrolling youth; establishing guidelines for the design of the programs, and the goals of the programs; and providing training and professional development for the staff, or youth activity leaders.

Many of the programs aim to involve youth in work with data in their aims These learning environments bring together youth activity leaders, educators, and those with technical expertise in STEM domains. Youth spent around three hours per day for four days per week for the approximately four-week programs, which were taught by youth activity leaders and scientists, engineers, and other community members with technical expertise.

## Participants

Participants consisted of 203 youth. Participants were from diverse racial and ethnic backgrounds (see Table 1). The mean age of participants was around 13 years old (from youth whose age was available: *M* = 12.71, *SD* = 1.70, *min.* = 10.75, *max.* = 16.36).

Table 1. Demographic characteristics for participating youth.

|  |  |
| --- | --- |
| Youth | Percentage |
| Sex | NA |
| Male | 50 |
| Female | 50 |
| Race/Ethnicity | NA |
| Hispanic | 48 |
| White | 6 |
| Black | 36 |
| Multi-racial | 3 |
| Asian/Pacific Islander | 7 |
| Parent Education | NA |
| High School or Below | 79 |
| Graduated from College (B.A. or B.S.) | 21 |

## Procedure

Before the start of the programs, youth completed a pre-survey that included questions about their experience in STEM, intention to pursue a STEM major or career, and other motivation and engagement-related measures; items about youths’ interest in STEM were the only items used from this survey in this study.

At the beginning of the programs, youth were introduced to the study and the phones used for data collection related to the ESM (Hektner, Schmidt, & Csikszentmihalyi, 2007). ESM is a method of data collection that involves asking youth to respond to short questions on phones (that were provided as part of the study). In particular, youth were signaled at random times (within intervals, so that the signals were not too near or far apart) in order to obtain a sample of their experience throughout the program. ESM data were collected two days each week, for three weeks (weeks 2-4 of the program). In all of the programs, about equal video-recording time was dedicated to classroom and field experiences. The programs associated with one of the intermediaries rotated between classroom and field experience days, while the other used the first half of each day for one and the second for the other. Each day, youth were signaled four times. These signals were at the same time for all of the youth within their program, but at different times between programs and between days within programs (with the constraint that no two signals could occur less than ten minutes apart).

The programs were video-recorded by research team members on the days during which ESM data were collected. So that the measures relating the video-recording and ESM data can be matched, the videos included a signal from the person doing the video-recording that identified the ESM signal to which youth were signaled to respond.

## Data Sources and Measures

Data sources consist of ESM measures of engagement and youths’ perceptions of themselves and the activity, pre-survey measures of youths’ interest, youths’ demographic information, and the video-recordings of programs.

### ESM measures of engagement for the profiles

. Measures for engagement were created from five ESM questions, as described in Table 2. All five items were ultimately used to construct the profiles of engagement. Each of the ESM items consisted of the item text and the following four item response options, of which youth were directed to select one: Not at all (associated with the number 1 on the survey; 1), A little (2), Somewhat (3), and Very Much (4).

Table 2. Items for constructs measured through ESM

|  |  |
| --- | --- |
| Construct | Item |
| Cognitive engagement | As you were signaled, were you learning anything or getting better at something? |
| Behavioral engagement | As you were signaled, how hard were you working? |
| Affective engagement | As you were signaled, did you enjoy what you are doing? |
| Perceived challenge | As you were signaled, how challenging was the main activity? |
| Perceived competence | As you were signaled, were you good at the main activity? |

### Measures from video for work with data

### Measures from video for work with data. Different aspects of work with data were identified from video-recordings. Specifically, codes for work with data were generated on the basis of the activity that the youth activity leaders were facilitating. The activity youth activity leaders were facilitating were from the STEM-Program Quality Assessment (STEM-PQA; Forum for Youth Investment, 2012), an assessment of quality programming in after-school programs. We then identified the specific activities that corresponded to the five aspects of work with data, as defined in Appendix A. While we chose to match the five aspects of work with data to the STEM-PQA code(s) that we interpreted as aligning most closely (in the cases of *generating data* and *interpreting and communicating findings*, choosing to use two STEM-PQA items for the codes), there are other possible ways that these could be matched, a point discussed later.

We then used these codes as part of the following coding frame, with the code names, possible values, code description, and examples from this study, as presented in Table 3. Note that this coding frame was not developed to assess work with data but rather was adapted for this purpose based on aligning dimensions of the STEM-PQA with the categories of the coding frame for work with data in this table.

Table 3. Codes for the aspects of work with data.

|  |  |  |  |
| --- | --- | --- | --- |
| Code Name | Values | Description | Example |
| Asking questions | 1: Present; 0: Not Present | Discussing and exploring topics to investigate and pose questions. | Youth generated questions they investigated related to tide ponds in an estuary ecosystem. |
| Making observations | 1: Present; 0: Not Present | Watching and noticing what is happening with respect to the phenomena or problem being investigated. | Youth observed the projectile motion of an object launched with a catapult. |
| Generating data | 1: Present; 0: Not Present | Figuring out how or why to inscribe an observation as data and generating coding frames or measurement tools. | Youth wrote in a table the number of pieces of recyclables they collected in parts of local waterways. |
| Data modeling | 1: Present; 0: Not Present | Understanding and explaining phenomena using models of the data that account for variability or uncertainty. | Youth calculated the average number of plant species found across a number of sites in the field. |
| Interpreting and communicating findings | 1: Present; 0: Not Present | Discussing and sharing findings. | Youth presented the outcomes of an investigation or engineered design in light of a research question or problem. |

Raters contracted by American Institute of Research (AIR) were trained in the use of the Program Quality Assessment tool (PQA), the broader assessment tool for which the STEM-PQA is a supplement. For the STEM-PQA, three of the same raters contracted by AIR to code the overall PQA measure used the STEM-PQA supplement to score one video segment, for which there were no disagreements on scoring for any of the items. The programs were divided up among all of the raters, so raters coded some of the videos for all of the programs. When the raters encountered a situation that was difficult to score, they would discuss the issue by telephone or more often by email after viewing the video in question and reach a consensus on how to score the specific item. Out of the 248 instructional episodes, 236 were code-able for work with data; for the 12 that were not codeable, issues with the video-recordings were the primary source of the missing data. These 236 responses were used for all of the analyses.

### Survey measures of pre-interest in STEM

. Measures of youths’ pre-interest were used as youth-level characteristics that predict the profiles of engagement. In particular, three items adapted from Vandell, Hall, O’Cadiz, and Karsh (2012) were used, with directions for youth to rate their agreement with the items’ text using the same scale as the ESM items. Reliability and validity information on this scale is presented in Vandell et al. (2008).

This measure was constructed by taking the maximum value for the scales for the different content areas (science, mathematics, and engineering), so that the value for a youth whose response for the science scale was 2.5 and for the mathematics scale was 2.75 would be 2.75. See Beymer, Rosenberg, and Schmidt (2018) for more details on this (use of the maximum value) measurement approach. The three items were: 1) I am interested in science / mathematics / engineering; 2) At school, science / mathematics / engineering is fun; 3) I have always been fascinated by science / mathematics / engineering).

Overall levels of this measure were high (*M* = 3.044 (*SD* = 0.901). The individual interest measure represented the mean of interest items across all relevant domains. Thus for some students, the mean was based on three items, while for others it was based on as many as nine items representing all three domains (with Cronbach values ranging from .77 - .86 for each domain-specific interest scale).

### Other youth characteristics

. In addition to the measures described in this section, demographic information for youths’ gender, and their racial and ethnic group are used to construct demographic variables for gender and membership in an under-represented (in STEM) group; membership in an under-represented group was identified on the basis of youths’ racial and ethnic group being Hispanic, African American, Asian or Pacific Islanders, or native American.

## Data Analysis

**Preliminary analyses**. Correlations (first-order Pearson) and the frequency, range, mean (*M*), and standard deviation (*SD*) are first presented for all variables. In addition, the frequency of the codes for aspects of work with data and the numbers of responses by youth, program, and instructional episode are presented.

### Analysis for Research Question #1 (the frequency of work with data)

.

The codes for the aspects of work with data (described above in the section on the measures) were counted up and presented as a proportion of the number of code-able instructional episodes. As the signals represent a sample of youths’ experiences in the programs, results from this analysis provide insight into how often each of the aspects took place during the programs.

### Analysis for Research Question #2 (what profiles of engagement emerge)

. Latent Profile Analysis (LPA; Harring & Hodis, 2016; Muthen, 2004) was used to identify profiles of engagement. LPA allows for capturing the multidimensional nature of engagement through profiles in terms of discovering groups of the ways in which youth experience engagement together and at once. A key benefit of the use of LPA, in addition to likelihood estimation-based fit indices, is probabilities of an observation being a member of a cluster (unlike in cluster analysis). These profiles make it possible to analyze the multivariate data collected on engagement in a way that balances the parsimony of a single model.

For these analyses, five variables were included: the three indicators for the experience of engagement (cognitive, behavioral, and affective) and the two necessary conditions for it (perceptions of challenge and competence). In addition, solutions with between two and ten profiles were considered. As part of LPA, the model type selection-where the type refers to which parameters are estimated–is a crucial topic. For the present study, six model types were considered. See Rosenberg et al. (2018) for a description of the models. The tidyLPA package (Rosenberg et al., 2018) and the MPlus software (Muthen & Muthen, 1998-2017) was used to carry out LPA through open-source statistical software we developed and published to the Comprehensive R Archive Network.

To select a solution in terms of the model type and the number of profiles to be interpreted and used in subsequent analyses, a number of fit statistics and other considerations were taken into account. These include a range of information criteria (AIC, BIC, and sample adjusted BIC [SABIC]), statistics about the quality of the profile assignments (entropy, which represents the mean posterior probability), a statistical test (the bootstrapped LRT [BLRT]), and concerns of interpretability and parsimony. On the basis of these criteria, a particular solution was selected and used as part of subsequent analyses; as the model selection process is an integral part of providing an answer to this question, the model and number of profiles selected are described in the section for the results for this research question.

### Analysis for Research Question #3 (how work with data relates to engagement)

. This question is focused on how work with data relates to the profiles of engagement. For the primary results for this question, mixed effects models that account for the cross-classification of the instructional episode (because of the dependencies of the responses associated with each of the 248 distinct ESM signals) and youth are used and for the “nesting” of both within each of the nine programs are used. The *lme4* R package (Bates, Martin, Bolker, & Walker, 2015) was used. All of the models for this and the subsequent research question use random effects for youth, instructional episode, and program effects. Youth and the instructional episode can be considered to be crossed with both nested within the program. The probability of a response belonging to the profile was the dependent variable, and the aspects of work with data are the independent variables. There are six models, for each of the six profiles. Because the outcome from LPA is not a classification (i.e., an observation is in a profile—or not) but a probability, the dependent variable is treated as a continuous variable.

First, models with only the random parts (i.e., random youth, instructional episode, and program effects) were specified. Then, the five aspects of work with data were added as predictors to the model. The results will be interpreted on the basis of which of the statistical significance and the magnitude and direction of the coefficients: For example, if the coefficient for the effect of the asking questions aspect of work with data upon one of the profiles was 0.10, and is determined to be statistically significant, then this would indicate that when youth are engaged in this aspect of work with data, then they are ten percentage points more likely to report a response in that particular profile.

For this question, models with the aspects of work with data both separate from and together with the youth characteristics were fit. The models with both sets together were also used as part of research question #4, though they are presented here (and interpreted in the sections for both results). Because the results were found to be identical when the aspects of work with data and the youth characteristics were considered in separate and in the same model, the results from the two sets of variables being in the same model were used both to provide answers to both this and the next research question. Specifically, pre-program interest in STEM, gender (entered s a dummy code with the value of “1” indicating female), and URM status (also entered as a dummy code, with “1” indicating a youth from a URM group) were added as predictors along with the aspects of work with data. Note that a composite for work with data (made as the sum of the individual aspects of work with data) was considered, but as it did only yield one (small) statistically significant result, the results for this analysis are not presented in the results.

# Results

## Descriptive statistics for the engagement measures

First, descriptive statistics for the five engagement variables that were used to estimate the profiles are presented in Table 4. These descriptive statistics show high overall levels of cognitive (*M* = 2.768, *SD* = 1.063), behavioral (*M* = 2.863, *SD* = 1.044) and affective (*M* = 2.831, *SD* = 1.051) engagement.

Table 4. Descriptive statistics for the variables used to create the profiles.

|  |  |  |  |
| --- | --- | --- | --- |
|  | n | Mean | SD |
| Cog. eng. | 2969 | 2.768272 | 1.0634473 |
| Beh. eng. | 2959 | 2.863467 | 1.0438073 |
| Aff. eng. | 2970 | 2.831313 | 1.0514690 |
| Challenge | 2970 | 2.270034 | 1.1168559 |
| Competence | 2970 | 3.000000 | 0.9518555 |

These statistics also show high perceptions of competence (*M* = 3.000 (*SD* = 0.952)) and moderate perceptions of challenge (*M* = 2.270 (*SD* = 1.117)). There was a similar degree of (moderate) variability across the engagement measures (see the *SD*s).

## Correlations among the study variables

Correlations between the variables that were used to create the profiles of engagement and the one other variable which was continuous (rather than a code for groups, in particular youths’ gender and URM status), pre-program interest in STEM (Table 5), were specified. In addition, relations between these variables and those for the five aspects of work with data were identified.

Table 5. Bivariate correlations among the study variables

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Pre-interest | Cog. eng. | Beh. eng. | Aff. eng. | Chall. | Comp. | Ask. | Obs. | Gen. | Mod. | Com. |
| Pre-interest |  |  |  |  |  |  |  |  |  |  |  |
| Cog. eng. | .14 |  |  |  |  |  |  |  |  |  |  |
| Beh. eng. | .13 | .60 |  |  |  |  |  |  |  |  |  |
| Aff. eng. | .12 | .59 | .57 |  |  |  |  |  |  |  |  |
| Chall. | .15 | .30 | .27 | .27 |  |  |  |  |  |  |  |
| Comp. | .06 | .40 | .41 | .47 | .08 |  |  |  |  |  |  |
| Ask. | -.18 | .02 | .01 | .01 | -.01 | -.01 |  |  |  |  |  |
| Obs. | .11 | .01 | .03 | -.01 | -.02 | -.00 | .38 |  |  |  |  |
| Gen. | -.08 | .02 | .02 | -.03 | -.01 | -.05 | .31 | .30 |  |  |  |
| Mod. | -.03 | .02 | .01 | .01 | .03 | -.00 | .42 | .19 | .35 |  |  |
| Com. | -.10 | .00 | -.02 | -.05 | -.06 | -.03 | .42 | .20 | .38 | .50 |  |

## Results for Research Question #1

### Frequency of the aspects of work with data. Of the 236 instructional episodes used in the analysis, 170 (72%) were coded as involving one or more of the five aspects of work with data. As a reminder, the instructional episode refers to the ten-minute block of time immediately preceding an ESM signal. As presented in Table 6, the five aspects of work with data occurred regularly. Making observations was found to be the least frequent of the five aspects, occurring in 24% of instructional episodes. Data modeling was the next most frequent aspect, occurring in 29% of the episodes, followed by asking questions (38%), generating data (43%), and communicating findings (again 43%).

As suggested by the proportions reported in Table 6, the different aspects of work with data often co-occurred within a single instructional episode. On average, there were 1.86 (*SD* = 1.61) aspects of work with data present during each instructional episode. This value indicates that, on average, youth were engaged in around two of aspects of the work with data during each instructional episode. There was a considerable amount of variation in the extent to which these types of work with data were supported in each program.

Table 6. Proportion of signals for which each of the aspects of work with data was present

|  |  |  |
| --- | --- | --- |
| Aspect of Work with Data | Proportion of Instructional Episodes | N |
| Asking Questions | 0.381 | 90 |
| Making Observations | 0.242 | 57 |
| Generating Data | 0.432 | 102 |
| Data Modeling | 0.288 | 68 |
| Communicating Findings | 0.436 | 103 |

## Results for Research Question #2: What profiles of youth engagement emerge from experiential data collected in the programs?

On the basis of fit statistics and statistical tests and concerns of interpretability and parsimony, a solution with six profiles of engagement was selected. Note that only models associated with the model 1 type and the model 2 type converged, and so only solutions associated with these two model specifications were considered.

Taking the information criteria, likelihood ratio tests, and concerns of interpretability and parsimony into consideration when reviewing the models that converged, either a model one type, six profile solution or a model one type, seven profile solution was found to be most reasonable. The results from analyses using *both* sets of profiles were found to be comparable and so the six profile solution was chosen on the basis of parsimony and its greater interpretability.

The result of this model selection process was the estimation of *six distinct profiles* identified from the data, as presented in Figures 2 and 3. Figure 2 shows the profiles with variables that were centered to have a *mean* of 0 and a *standard deviation* of 1. Thus, the *y*-axis for this plot is labeled “Z-score”). Figure 3 shows the profiles with the raw data (not transformed). Thus, the *y*-axis for this plot is labeled “Value.” This solution represents the profiles of engagement identified to answer this research question and for use in subsequent analyses. The two plots are presented because they provide a different view into the composition of the profiles.

Figure 2. The six profiles of engagement (with variable values standardized)



Figure 3. The six profiles of engagement (with raw variable values)

The six profiles are characterized by both varying levels on both the indicators of engagement (cognitive, behavioral, and affective) and perceptions of challenge and competence. Also, the number of observations across the profiles is relatively balanced (with no profiles associated with a very large or small number of observations). The universally low profile was associated with the most substantial number of observations (*n* = 667), followed by the all moderate profile (*n* = 638); each of the other four profiles was associated with 300 to 400 observations. The results for research questions 3-5 use this solution and the six profiles in subsequent analyses.

A MANOVA was carried out to determine whether the values of variables differ across the profiles, with multiple ANOVAs used to determine which variables (and for which profiles) there were differences. The MANOVA was statistically significant (*Pillai-Bartlett* = 0.633, *p* < .001). The *F*-test associated with each ANOVA was also statistically significant.

Descriptions of each the profiles taking account of their size (in terms of the number of responses for which the profile was most likely), their variable values, and what the profiles suggest about youth engagement follow.

* A *universally low* profile consisted of a substantial proportion of responses (22.55%) was identified. This profile was characterized by low levels of working hard, learning something new, and enjoying the activity, and perceptions challenge and competence. This profile reflects very low levels of youth engagement during the specific instructional episodes during which youth were signaled to respond.
* An *only behaviorally engaged* profile with a small proportion of responses (12.51%) was identified. This profile was characterized by moderate levels of working hard, very low enjoyment of the activity, and moderate levels of learning something new and challenge and competence. These responses suggest that youth perceive themselves to be working hard, but to not be enjoying what they were doing and to not report learning something new, nor to be particularly challenged or good at what they were doing when signaled.
* An *only affectively engaged* profile with a small proportion of responses (11.66%) was identified. This profile was characterized by moderate levels of enjoyment, low levels of hard work, and moderate levels of learning something new, challenge, and competence. When youth report this response, they enjoy what they were doing, but were not working hard or learning something new, nor do youth report being challenged by or good at the activity they were doing.
* An *all moderate* profile with a large proportion of responses (21.57%) was identified. This profile was characterized by moderate levels of the three indicators of working hard, learning something new, enjoying the activity, challenge, and competence. For youth reporting *all moderate* engagement were engaged, but may have the potential to be more highly engaged (and challenged by and good at the activity).
* An *engaged and competent but not challenged* profile with a modest proportion of responses was identified (15.21%). This profile was characterized by high levels of working hard, learning something new, enjoying the activity, and competence, but low levels of challenge. This profile suggests youth can be highly engaged, while not being challenged by the activity they were involved in at the time they were signaled.
* A *full* profile with a modest proportion of responses (16.50%) was identified. This profile was characterized by high levels of working hard, learning something new, enjoying the activity, challenge, and competence. These responses reflect a very high level of engagement, both relative to the other profiles and in absolute terms. When youth report engagement in ways that were associated with this profile, they report being challenged and good at what they were doing and, on the basis of these variables and the indicators of engagement, youth very highly engaged.

For all six profiles, the *ICC*s represent the variability (the proportion of variance explained) associated with each of the levels for each profile. These models can show how much variability in the profiles was systematic at these different levels and was potentially attributable to youth, instructional episode, and program. The systematic variability at the youth level, for example, could be .10 for the *Full* profile and .025 for the *Universally Low* profile. At the program level, the *ICC*s were found to be small, with values ranging from 0.00 to 0.023, suggesting that little variability can be explained by the program. For the instructional episode level, the *ICC*s were also small, ranging from 0.004 to 0.01. Finally, at the youth level, the *ICC*s ranged from .093 to .432. These findings show that there was substantial variability in the profiles present at the youth level, with less variability was explained by either the program youth were in or the nature of the particular instructional episode present when youth were signaled. These results set the stage for those for the next two research questions, on the relations between the aspects of work with data (for research question #3) and the youth characteristics (for research question #4) and the profiles of engagement.

## Results for Research Questions #3 and #4: Aspects of work with data, youth characteristics, and engagement

To understand how aspects of work with data were related to engagement, six models were specified – one for each engagement profile. The five aspects of work with data and the three youth characteristics were the predictor variables. Various aspects of work with data tended to co-occur and so simultaneously entering indicators for all five aspects served to isolate the association for any single aspect while controlling on the presence of the others. All models also include some youth characteristics which will be used to answer research question five below.

In Table 7, each column represents the output from one of the six different models. As an example, the first column reports the coefficients for the associations between the predictor variables and the *Only behavioral* profile. Because the outcome was in the form of a probability (ranging from 0.00 to 1.00), it can be interpreted as the change in the probability of a response being associated with each profile. Note that the *p*-values were calculated using the most conservative and recommended by recent research Kenward-Rogers approximation (Halekoh & Hojsgaard, 2014).

Table 7. Findings for relations between the aspects of work with data and youth characteristics and the profiles of engagement.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Profile** | **Universally low B (SE)** | **Only behavioral B (SE)** | **Only affective B (SE)** | **Eng. and comp., not chall. B (SE)** | **All moderate B (SE)** | **Full B (SE)** |
| Pre-interest | -0.047 (0.022) | -0.013 (0.012) | -0.012 (0.019) | 0.039 (0.016)\* | 0.007 (0.01) | 0.018 (0.021) |
| Gender-Female | 0.06 (0.037)+ | 0.019 (0.019) | -0.038 (0.033) | 0.025 (0.028) | -0.02 (0.018) | -0.035 (0.037) |
| URM status | -0.01 (0.052) | 0.031 (0.026) | -0.076 (0.046) | -0.012 (0.04) | 0.018 (0.025) | 0.043 (0.053) |
| Asking | -0.015 (0.018) | 0.015 (0.015) | 0.023 (0.017)+ | -0.011 (0.015) | 0.004 (0.014) | -0.019 (0.016) |
| Observing | 0.003 (0.018) | 0.013 (0.015) | 0.007 (0.017) | 0.009 (0.015) | -0.017 (0.014) | -0.025 (0.016) |
| Generating | -0.014 (0.017) | 0.014 (0.014) | 0.012 (0.016) | -0.014 (0.014) | -0.02 (0.013) | 0.027 (0.015)\* |
| Modeling | 0.004 (0.019) | -0.023 (0.016) | -0.004 (0.018) | 0 (0.015) | -0.012 (0.015) | 0.034 (0.017)\* |
| Communicating | 0.002 (0.018) | 0.018 (0.015) | -0.011 (0.017) | 0.004 (0.015) | 0.016 (0.014) | -0.027 (0.016) |

*Note.* + p < .10; \* p < .05

The only engagement profile that was significantly associated with any aspects of work with data was the Full profile (see the column with the column name Full for these results). When program activities involved modeling data, youth were around 3% more likely to be fully engaged ( = 0.034 (0.017), *p* = .020; partial = .002). In other words, when program activities included modeling data, youth were more likely to report working harder, learning more, enjoying themselves more, and feeling more competent and challenged.

Youth were also more likely to be in the Full engagement profile when program activities included generating data ( = 0.027 (0.015), *p* = .033; partial = .002). These particular program activities increased the probability of full engagement by around 3%. To sum up these two findings, modeling data and generating data were associated with a (very) positive form of engagement, that exhibited by the Full profile. However, the effect sizes indicate quite small effects in substantive terms. Note that interactions between the individual aspects of work with data and youth characteristics were also specified. However, none of these interactions were found to be statistically significant.

Sensitivity analysis was carried out for the statistically significant two effects was carried out to determine just how robust they were. This follow-up analysis revealed that the effect of modeling data on *Full* engagement much more robust than that for generating data: 9.835% of this effect (of data modeling) would have to be due to bias to invalidate the inference about its effect. For generating data, only 1.884% of the effect of generating data would need to be due to bias to invalidate the inference about its effect. These values were not minuscule but were also not very large given what we known past research (Frank, 2003).

Associations between youth characteristics and the six profiles are reported in the top half of Table 7. Youth who enter the program with higher levels of interest (in STEM) were more likely to report being in the engaged and competent but not challenged profile ( = 0.039, *p* = .009; partial = .001). In other words, youth who were more interested at the outset of the program report working harder, learning more, enjoying themselves more, and feeling more competent when they were involved in program activities, though they also report lower levels of challenge. For this effect, 17.879% would be needed to invalidate the inference, suggesting a moderately robust effect.

In terms of youths’ pre-program interest, these analyses show that youth who enter the program with higher levels of interest (in STEM) were more likely to report being in the *Engaged and competent but not challenged* profile ( = 0.039, *p* = .009; *partial*  = .001). For each one-unit increase in pre-program interest in STEM, youth were around 4% more likely to report this profile. In other words, youth who were more interested at the outset of the program report working harder, learning more, enjoying themselves more, and feeling more competent when they were involved in a program’s activities, though they also report lower levels of challenge. For this effect, 17.879% would be needed to invalidate the inference, a slightly larger value for the follow-up sensitivity analysis than those found for the (statistically significant) relations involving the aspects of work with data, suggesting a moderately robust effect.

There were not any statistically significant effects of youths’ URM status. This lack of relations between URM status and youth engagement may be a function of the large proportion of youth from under-represented (in STEM) racial and ethnic groups. Hispanic (48%), African American or Black (36%), and youth who identify as being from multiple racial and ethnic groups (3%) made up 87% of the youth in the programs, so there were not many youth *not* from under-represented groups in the sample, suggesting that the absence of findings may be due to this small sample (and low statistical power). Nevertheless, no relations between URM status and youths’ engagement were found, indicating that there is at least no evidence that youth from such backgrounds do engage in different ways.

# Discussion

Each of the disciplines that contribute to STEM learning - science, technology and computer science, engineering, and mathematics - involve work with data. In this study, engagement was used as a lens to understand the experience of youth working with data during summer STEM programs. In particular, five aspects of work with data, a) asking questions, b) observing phenomena, c) constructing measures and generating data, d) data modeling, and e) interpreting and communicating findings, were occurred regularly in the programs. There were some examples of ambitious activities centered on working with real-world data as well as some that highlight substantial heterogeneity in how work with data was enacted.

We identified six profiles of engagement using LPA. These profiles represented different configurations of how youth were working hard, learning, enjoying themselves, and feeling challenged and competent at the time they were signaled as part of the ESM approach. Relations of the five aspects of work with data and youth characteristics (pre-program interest in STEM and youths’ gender and status in terms of being a member of under-represented groups in STEM) were, overall, not strongly related with the profiles of engagement, though some key findings were identified. Generating and modeling data were both related to the most potentially beneficial profile (full engagement), one characterized by high levels of all five of the engagement variables.

This study suggests that work with data and contemporary engagement theory as interpreted in this study can serve as a frame to understand what youth do in summer STEM programs. These findings also show the value of an innovative method, ESM, and an analytic approach designed to identify engagement holistically, LPA, that together to provide some access to youths’ experience in-the-moment of the activities they were involved in during the program. Data, and how youth and students in K-12 settings can themselves work with data, is an important, yet perhaps under-emphasized part of STEM learning. In the remainder of this section, We discuss key findings with respect to a) work with data, b) youths’ engagement, and c) what relates to youths’ engagement. Also, some limitations and recommendations for future research as well as implications for practice are identified and described.

## Key findings related to work with data in summer STEM programs

Results showed that work with data was common in the summer STEM programs. There was variability in *which* aspects of work with data was present: Making observations, in some form, occurred during 24% of the program’s time, for example, while generating data and communicating findings both occurred more frequently, during 43% of the instructional episodes. These findings, broadly, suggest that work with data occurred enough that we might expect to see differences in youths’ engagement. They align with what may be expected given past research: Such programs are designed to engage youth in the practices, including and as we argued earlier *especially* those relating to work with data, of STEM domains (Dabney et al., 2012; Elam et al., 2012). Even still, these are the first results of this kind (in terms of the proportion of the time spent in the programs). Using video-recording data and a sampling strategy that can provide insight into the amount of overall time spent was an important component of achieving these findings. While there are no other results of this particular kind, a related, an area of related work concerns other studies that have used the PQA measure.

## Key findings related to engagement

Six profiles of engagement were found using a rigorous model selection approach. In terms of comparing the make-up of the specific profiles to other, past research, little work has examined profiles of engagement. Schmidt et al. (2018) did examine profiles of engagement, which were constructed from indicators cognitive, behavioral, and affective engagement (but not perceptions of challenge and competence, as in this study). Schmidt et al. (2018) found six profiles, some of which partially overlap with those found in the present study. In particular, on the basis of the items shared between the studies, a *Universally low*, *All moderate*, and *Full* profile were found in both studies. However, as these profiles are characterized by the (uniform) level across all of the variables, this is only limited evidence for the presence of these profiles in the larger population of youth engaged in science and STEM-related learning activities. The number of profiles found is broadly similar to that found in past researcg: Schmidt et al. (2018) found six profiles of engagement.

The six profiles lend insight into how youth engage during summer STEM programs. In particular, the *Only behavioral*, *Only affective*, and *Engaged and competent but not challenged* profiles were found in the present study, but not in Schmidt et al.’s (2018) study. Youth were highly engaged (as may be anticipated given the goals and design of such programs), but perceive a misalignment between their (high) competence and how (not very) challenged they were. According to past theory (e.g., Csikszentmihalyi, 1997) and some research (e.g., Shernoff et al., 2016), such a profile would be unexpected, as high levels of engagement are expected to be associated with high levels of *both* challenge and competence. In this study, a profile characterized by high competence but (very) low challenge was associated with very high engagement. This profile, *Engaged and competent but not challenged*, then, seems to suggest a type of engagement that may be unique and common to summer STEM programs. Perhaps such a profile may be expected given the lower stakes (compared to formal educational settings) of summer STEM programs (and other informal learning environments) and the degree of competence that youth–many of whom have chosen to attend the particular program (Beymer et al., 2018)–perceive during them.

In addition to suggesting a profile of engagement that is distinct to summer STEM program, this profile and the other two not found in past research have some implications for youth activity leaders. In particular, they suggest that lower challenge may *not*, as would be anticipated given theory and past research, be associated with lower engagement. Because of this, it may be recommended that activities that are not challenging but have other possible benefits to youth (i.e., benefits from activities designed to support youths’ social skills), can be integrated into programs, along with other, more challenging activities that are also highly engaging to youth.

## Key findings related to work with data and youth characteristics and their relations to engagement

Why might these relations be so minimal? First, and foremost, the little variability at the instructional episode level was noteworthy because it means that few relations between variables at the instructional episode were expected. In particular, there were small ICCs at the instructional episode level for all six profiles. This suggests that there was very little systematic variability at the particular level that a variable for work with data was at. Additionally, the ICC values found in this study were smaller than those found in the one other past study that employed the same analytic approach (Strati et al., 2017). The relative absence of variability at the instructional episode level may be due to the summer STEM setting: Perhaps youth are less likely to engage differently from instructional episode to instructional episode (compared to in K-12 educational settings) because there is less variability in what took place across the episodes or because youth perceive there to be lower stakes for the programs’ activities and therefore do not perceive the changes in the instructional episode as a salient factor in terms of their engagement. This consideration is described in greater detail in the limitations section.

Even so, there were *some* noteworthy findings that could be anticipated on the basis of the importance of the two aspects of work with data that were found to relate positively to youths’ engagement. In particular, both generating and modeling data were found to be positively (and statistically significantly) related to the *Full* profile, suggesting that when youth were involved in these practices, then they were more likely to be highly engaged. In particular, given the makeup of this profile, this suggests that when youth were involved in these aspects of work with data, they were more likely to report high levels of cognitive, behavioral, and affective engagement, and high perceptions of competence and challenge. Generating and modeling data may have such relations because they were particularly important aspects of work with data. As Lehrer and Schauble (2006) explain, *inscriptions serve commitments*: Choosing to record an observation or an idea as data involves the process of identifying something that is worth recording and then recording the parts that are of interest. Thus, generating data may be fully engaging to youth because it is, generally, demanding and important with respect to work with data. Modeling, too, is an important practice. It has been described as *the* central scientific and engineering practice (Lehrer & Schauble, 2015; Weisberg, 2012), and its relations with full engagement provides some actionable evidence for its importance in the context of summer STEM programs. Modeling may be especially engaging to youth because such work positions learners as the creators of new information, in addition to using models created by others to learn about authoritative sources of information. This is one of the affordances of modeling in teaching and learning contexts (Berland et al., 2016; Schwarz et al., 2009). Moreover, when learners create new knowledge through activities such as modeling, they can begin to shape the process of constructing new knowledge in a domain, a challenge in science education contexts (Stroupe, 2014) and likely in other STEM content areas, we well.

The null findings for the relations of asking questions, making observations, and interpreting and communicating findings were noteworthy, too. They suggest that their effects were not large enough to be detected. They may not be able to be detected for a number of reasons: simply because they were very small given all of the other factors that impact youths’ engagement, because the aspects of work with data were enacted in a myriad of ways which may be more or less engaging, and because they simply were not as engaging. Nevertheless, future research may seek to understand why these aspects did not relate to engagement.

As there is no research on how work with data relates to youths’ engagement, the findings associated with this research question provide some initial evidence for how some aspects of work with data relate to youths’ engagement. These findings suggest that these activities may not be more engaging *per se*. Instead, it may be the way that youth engage in them that matters, in alignment with past research (Berland et al., 2017). While the findings for this question were somewhat minimal, there are key findings from both the important relationships that were found to be statistically significant (between generating data and data modeling and *Full* engagement) and from those that were not. Other samples, other enactments of work with data, and, possibly, other analytic approaches can build on this work to further substantiate what is known about how work with data engages youth and other learners.

There was substantial variability in the profiles of engagement at the youth level, but there were not many relations in terms of youths’ gender, URM status, or pre-program interest. Given past theory and research have suggested that learners’ gender, URM status, and individual or pre-program interest can predict engagement (Bystydzienski, Eisenhart, & Bruning, 2015; Hidi & Renninger, 2006; Shernoff & Schmidt, 2008). Despite these surprising findings, youth with higher pre-program interest were found to be more likely to be *Engaged and competent but not challenged*. This suggests that youth with a higher interest in STEM were inclined to be highly engaged and good at what they were doing, but were not challenged by the activities they experience. This finding is in line with past research suggesting a relationship (direct or as a moderator) between youth characteristics (including interest) and their engagement (Shernoff et al., 2003; Shernoff et al., 2016; Strati et al., 2017). More specifically, this finding suggests that for youth who were particularly interested (and those who choose to attend) summer STEM programs, what they were involved in may not challenge them very highly. This finding has implications for past research that shows youth who choose to attend summer STEM programs were more engaged (but that does not speak to their degree of challenge; Beymer, Rosenberg, Schmidt, & Naftzger, 2018).

## Limitations to the present study and recommendations for research

First, the programs participating in this study were not designed especially to support youth in work with data. Instead, the programs were designed around best practices for summer STEM programs to support youth to engage in a wide variety of STEM-related practices–and in other activities, such as those intended to build a sense of camaraderie among the youth in the programs. In this study, aspects of work with data were identified and were found to be common, but some of the heterogeneity in the nature of working with data may be due to this reason: Planning and instruction for the programs did not aim to foster rich work with data any more than the other activities (STEM and otherwise) that made up their programming.

In addition to the varied ways in which youth worked with data, some of the relations of the variables for the five aspects of work with data to youths’ engagement may be due to the ways that the variables for work with data indicated, in fact, many different ways of working with data. Some of these aspects of working with data, particularly those that were highly-specific with respect to how the data was involved and to how focused and sustained the work with data-related activity was, may be more engaging to youth than the others, such as those that were more general, instructor-focused, or brief. These two types of working with data were considered the same in the variables used to predict youths’ engagement. Future research can aim to understand youths’ engagement in outside-of-school data science programs and K-12 units, for example, that are focused more on work with data to understand better how work with data engages youth. Nevertheless, this study does provide insight into how work with data took place during *model* (i.e., designed around best practices for such programs) summer STEM programs and how such work relates to youths’ engagement.

In addition to limitations related to the focus of the programs and how work with data was enacted as part of a cycle, there were also some general measurement-related limitations. Work with data can be difficult to measure because, as the qualitative analysis revealed, there were a variety of ways in which youth can be involved in work with data. McNeill and Berland (2017) describe a similar type of disagreement across science education settings: While a limitation, the coding frame did represent agreement across a range of studies across STEM contexts for the aspects of work with data. In terms of the alignment of the measure with the conceptual framework for work with data, the dimensions of the STEM-PQA measure aligned closely with the aspects of work with data. However, there were some divergences that may have had an impact upon some of the findings. For example, for the interpreting and communicating findings code, the STEM-PQA codes for *Analyze* (“Staff support youth in analyzing data to draw conclusions”) and *Use symbols or models* (“Staff support youth in conveying STEM concepts through symbols, models, or other nonverbal language”) were used. In the case of the latter STEM-PQA code, conveying STEM concepts through symbols, models, or other nonverbal language could have reflected instructional episodes in which youth used, for example, mathematical equations or formulas, but did not do so as part of modeling data of a phenomena in the world: They could have simply been using an equation outside of the context of any particular phenomena. Future research may consider the usefulness of coding for this aspect of work with data (and this aspect of science curricular standards in particular; see NGSS Lead States, 2013).

## Implications for Practice

A few implications for practice can be drawn from this study, though these are somewhat restricted given the minimal findings. First, *generating data* and *modeling data*, in particular, may be beneficial in terms of engaging youth. Youth activity leaders (in summer STEM and other STEM enrichment contexts) and teachers (in *formal* learning environments) can best include the beneficial practices of generating and modeling data not in isolation, but rather through involving youth and learners in complete cycles of an investigation. This aligns with both foundational and contemporary research on work with data in education (Berland et al., 2018; McNeill & Berland, 2017; Hancock et al., 1992; Lee & Wilkerson, 2018).

Another implication concerns how work with data was enacted. As found in this study, work with data (and even specific aspects of work with data, such as asking questions) does not involve activities that are enacted in a universal way. An instructor instead of youth interpreting and communicating findings, for example, or learners asking general, conceptual questions about work with data, as another, are different from youth working to interpret findings and figuring out how to ask a question that can be answered with data, respectively. This heterogeneity suggests to those involved in planning and enacting engaging activities that involve data to consider *who* works with data carefully, *how* they do so, and *how much time and sustained focus* is required for such activities to be carried out. This implication aligns with recent curricular reform efforts, some of which suggest that involving work in STEM-related practices is most effective when it involves learner-driven (but instructor-supported) iterative processes of identifying a question or problem, marshaling sources of data that can be used to figure out what is happening, and developing model-based explanations that are shared with the learning community (National Governors Association, 2013; National Research Council, 2012; NGSS Lead States, 2013). While just two implications, youth activity leaders and teachers and those designing data-rich activities and evaluating the impacts of instruction based on such activities can use the findings from this study as a starting point to consider how engaging in work with data may also prepare learners to think of, understand, and take action based on data in education and in other areas of their lives.

# References

Berland, L. K., Schwarz, C. V., Krist, C., Kenyon, L., Lo, A. S., & Reiser, B. J. (2016). Epistemologies in practice: Making scientific practices meaningful for students. *Journal of Research in Science Teaching, 53*(7), 1082-1112.

Beymer, P. N., Rosenberg, J. M., Schmidt, J. A., & Naftzger, N. (2018). Examining relationships between choice, affect, and engagement in out-of-school time STEM programs. *Journal of Youth and Adolescence, 47*(6), 1178-1191. <https://doi.org/10.1007/s10964-018-0814-9>

Bielik, T., & Yarden, A. (2016). Promoting the asking of research questions in a high-school biotechnology inquiry-oriented program. *International Journal of STEM Education, 3*(1), 15.

Bystydzienski, J. M., Eisenhart, M., & Bruning, M. (2015). High school is not too late: Developing girls’ interest and engagement in engineering careers. *Career Development Quarterly, 63*(1), 88–95. <http://doi.org/10.1002/j.2161-0045.2015.00097.x>

National Governors Association Center for Best Practices, Council of Chief State School Officers. (2010). *Common Core State Standards for Mathematics*. Washington, DC: National Governors Association Center for Best Practices and the Council of Chief State School Officers.

Csikszentmihalyi, M. (1997). *Finding flow: The psychology of engagement with everyday life*. New York, NY: Basic Books.

English, L. D. (2012). Data modelling with first-grade students. *Educational Studies in Mathematics, 81*(1), 15-30.

Finzer, W. (2013). The data science education dilemma. *Technology Innovations in Statistics Education, 7*(2), p. 1-9.

Forum for Youth Investment. (2012). *Youth Program Quality Assessment*. Washington, DC: The Forum for Youth Investment

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report*. Alexandria, VA: American Statistical Association.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59-109.

Gelman, S. A., & Markman, E. M. (1987). Young children’s inductions from natural kinds: The role of categories and appearances. *Child Development, 58*(6), 1532-1541.

Gopnik, A., & Sobel, D. M. (2000). Detecting blickets: How young children use information about novel causal powers in categorization and induction. *Child Development, 71*(5), 1205-1222.

Hancock, C., Kaput, J. J., & Goldsmith, L. T. (1992). Authentic inquiry with data: Critical barriers to classroom implementation. *Educational Psychologist, 27*(3), 337-364.

Harring, J. R., & Hodis, F. A. (2016). Mixture modeling: Applications in educational psychology. *Educational Psychologist, 51*(3-4), 354-367.

Hasson, E., & Yarden, A. (2012). Separating the research question from the laboratory techniques: Advancing high‐school biology teachers’ ability to ask research questions. *Journal of Research in Science Teaching, 49*(10), 1296-1320.

Hektner, J. M., Schmidt, J. A., & Csikszentmihalyi, M. (2007). *Experience sampling method: Measuring the quality of everyday life*. Sage.

Konold, C., & Pollatsek, A. (2002). Data analysis as the search for signals in noisy processes. *Journal for Research in Mathematics Education, 33*(4), 259-289.

Lee, H., & Hollebrands, K. (2008). Preparing to teach mathematics with technology: An integrated approach to developing technological pedagogical content knowledge. *Contemporary Issues in Technology and Teacher Education, 8*(4), 326-341.

Lee, V. R., & Wilkerson, M. (2018). *Data use by middle and secondary students in the digital age: A status report and future prospects*. Commissioned Paper for the National Academies of Sciences, Engineering, and Medicine, Board on Science Education, Committee on Science Investigations and Engineering Design for Grades 6-12. Washington, D.C.

Lehrer, R., & Romberg, T. (1996). Exploring children’s data modeling. *Cognition and Instruction, 14*(1), 69-108.

Lehrer, R., & Schauble, L. (2004). Modeling natural variation through distribution. *American Educational Research Journal, 41*(3), 635-679.

Lehrer, R. & Schauble, L. (2015). *Developing scientific thinking*. In L. S. Liben & U. Müller (Eds.), Cognitive processes. Handbook of child psychology and developmental science (Vol. 2, 7th ed., pp. 671-174). Hoboken, NJ: Wiley.

Lehrer, R., Kim, M. J., & Jones, R. S. (2011). Developing conceptions of statistics by designing measures of distribution. *ZDM, 43*(5), 723-736.

Lehrer, R., Kim, M. J., & Schauble, L. (2007). Supporting the development of conceptions of statistics by engaging students in measuring and modeling variability. *International Journal of Computers for Mathematical Learning, 12*(3), 195-216.

Lesh, R., Middleton, J. A., Caylor, E., & Gupta, S. (2008). A science need: Designing tasks to engage students in modeling complex data. *Educational Studies in Mathematics, 68*(2), 113-130.

McNeill, K. L., & Berland, L. (2017). What is (or should be) scientific evidence use in k‐12 classrooms? *Journal of Research in Science Teaching, 54*(5), 672-689.

McNeill, K. L. & Krajcik, J. (2007). *Middle school students’ use of appropriate and inappropriate evidence in writing scientific explanations*. In Lovett, M & Shah, P (Eds.), Thinking with data. (pp. 233-265). New York, NY: Taylor & Francis Group, LLC.

Miller, E., Manz, E., Russ, R., Stroupe, D., & Berland, L. (2018). Addressing the epistemic elephant in the room: Epistemic agency and the next generation science standards. *Journal of Research in Science Teaching*, *55*(7), 1053-1075. <https://doi.org/10.1002/tea.21459>

Muthén, B. (2004). *Latent variable analysis*. The Sage handbook of quantitative methodology for the social sciences. Thousand Oaks, CA: Sage Publications, 345-68.

Muthén, L. K., & Muthén, B. O. (1997-2017). *Mplus User’s Guid*e. Los Angeles, CA: Muthén & Muthén.

NGSS Lead States. (2013). *Next generation science standards: For states, by states*. Washington, DC: National Academies Press.

Petrosino, A., Lehrer, R., & Schauble, L. (2003). Structuring error and experimental variation as distribution in the fourth grade. *Mathematical Thinking and Learning, 5*(2&3), 131-156.

Rosenberg, J. M., Schmidt, J. A., & Beymer, P. N. (2018). *tidyLPA: Easily carry out Latent Profile Analysis*. <https://cran.r-project.org/web/packages/tidyLPA/index.html>

Rosenberg, J. M., Xu, R., & Frank, K. A. (2018). *KonFound-It!: Quantify the robustness of causal inferences*. <http://konfound-it.com>.

Schmidt, J. A., Rosenberg, J. M., & Beymer, P. (2018). A person-in-context approach to student engagement in science: Examining learning activities and choice. *Journal of Research in Science Teaching*, 55(1), 19-43. <https://dx.doi.org/10.1002/tea.21409>

Shumow, L., & Schmidt, J. A. (2013). *STEM interest and engagement (STEM I.E.)*. National Science Foundation proposal for award number 1421198.

Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The challenges of defining and measuring student engagement in science. *Educational Psychologist, 50*(1), 1-13. <doi:10.1080/00461520.2014.1002924>

Stroupe, D. (2014). Examining classroom science practice communities: How teachers and students negotiate epistemic agency and learn science‐as‐practice. *Science Education, 98*(3), 487-516.

Vandell, D. L., Hall, V., O’Cadiz, P., & Karsh, A. (2012). *Piloting outcome measures for summer learning initiative programs*. Final report to the David and Lucile Packard Foundation, Children, Families, and Communities Program. Retrieved from <http://faculty.sites.uci.edu/childcare/files/2013/07/SL-Outcomes-2011-Pilot_Edited_8.19.pdf>

Weisberg, M. (2012). *Simulation and similarity: Using models to understand the world*. Oxford University Press: Oxford, England.

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review, 67*(3), 223-248.

Wilkerson, M. H. & Fenwick, M. (2017). *The practice of using mathematics and computational thinking*. In C. V. Schwarz, C. Passmore, & B. J. Reiser (Eds.), Helping Students Make Sense of the World Using Next Generation Science and Engineering Practices. Arlington, VA: National Science Teachers’ Association Press. pp. 181-204.