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

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).

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.

Statistical code for the analyses completed for this study is available at the repository here: <https://osf.io/65ckh/>. The data are available upon request from the authors.

The authors declare that they have no competing interests.

Author 1 analyzed and interpreted the data and wrote the manuscript. Author 2 collected the data, wrote sections of the manuscript, and served as Author 1’s thesis co-chair. Author 3 served as Author 1’s thesis co-chair and wrote sections of the manuscript. All authors read and approved the final manuscript.

Abstract

In this study, we explored how 203 early adolescents (age range 10-16) from nine out-of-school STEM programs worked with data by forming questions, constructing measures, attending to precision, creating models, and representing data. Using the Experience Sampling Method (ESM) we measured individual students engagement during specific activities that involve work with data, to better understand profiles (or common patterns) of engagement during these activities. We found that work with data was a common occurrence in these out-of-school programs and that six distinct profiles characterized learners’ patterns of cognitive, behavioral, and affective engagement. We found that specific activities – generating data and modeling data – were significantly associated with full engagement profiles, though work with data, overall, explained a small amount of the variability in youths’ engagement. We found few relations between youth characteristics (gender, under-represented status, or pre-program interest) and how youth engaged in work with data. Implications of the are discussed with respect to work with data in STEM in general and to engagement in informal learning environments.

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

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

Today data are all around us, shaping how we work, socialize, teach, and learn. Data are not only ubiquitous but are also important. While there has been a lot of attention in higher and K-12 education on administrators’ and teachers’ use of data (e.g., McCaffrey, Lockwood, Koretz, & Hamilton, 2003; Horn, Kane, & Wilson, 2015), how learners, create, use, and assess the validity of data as a part of their learning has been explored less.

Researchers have studied students’ engagement in math and science (Fredricks, Filsecker, & Lawson, 2016; Authors and colleagues, 2018a; Schneider et al., 2016), but have not examined how students engage in work with data in STEM settings. Knowing more about how youth engage in work with data —independent of achievement and learning— is a meaningful outcome for STEM learners (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).

Because engaging in work with data seems to be so potentially beneficial to learners, knowing more about how work with data relates to learners’ engagement is needed. Through work with data, learners can transform from consumers of knowledge to creators of knowledge (Hancock, Kaput, & Goldsmith, 1992; Lehrer & Schauble, 2015; Lee & Wilkerson, 2018; Finzer, 2013). Working with data is also supported by recent reform efforts that emphasize not just knowing *about* key concepts, but participating in the practices of STEM disciplines, foci of both science and mathematics curricular standards (NGSS Lead States, 2013; National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010). Indeed, work with data presents an area of overlap between the two sets of standards.

Our purpose, then, is to examine a variety of learning activities that involve work with data and to explore how work with data and youth characteristics relate to youth engagement. We explore these topics in the context of outside-of-school STEM enrichment programs carried out during the summer that introduce youth to STEM-related data in a context that builds upon youths’ interests. Knowing more about youths’ experiences in working with data can also provide a foundation for subsequent work to explore how particular curricula and engaging experiences for youth spark their interest in STEM, including hobbies and occupations related to data science, but also pursuing future study and careers in science, mathematics, or engineering and computer science.

**Background**

# Defining Work with Data

Work with data has been conceived in varied ways by researchers (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. Hancock et al. (1992) provided a useful description of the process of working with data in general, namely, “using data to solve real problems and to answer authentic questions” (p. 337). Hancock et al. (1992) focus specifically on two over-arching activities, data creation and analysis, arguing that the former (data creation) is underemphasized 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, we focus on the core aspects that scholars have most often included in their conceptualizations. In particular, we focus on five key aspects that are most useful for investigating work with data all of the STEM content areas. 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, subsequently, making observations about other phenomena, as follows.

* *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 (or regression) models
* *Interpreting and communicating findings*: Discussing and sharing findings from the earlier processes of working with data

# Past Research on Work With Data

Working 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. This focus on phenomena is particularly relevant to those designing and enacting learning opportunities focused on work with data given the greater availability (or the ability to create data about) many aspects of the natural and social world (Lee & Wilkerson, 2018). 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 focused on mathematics contexts and 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., Finzer, 2013; Konold & Pollatsek, 2002;). There has been some research about work with data in science settings, too. However, what it means to work with data can vary greatly in actual classrooms and other learning environments (McNeill & Berland, 2017).

Past research on cognitive capabilities as outcomes from working with data 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). A key outcome of engaging in work with data has to do with how learners account for variability (Lee, Angotti, & Tarr, 2010; Lehrer, Kim, & Schauble, 2007; Lesh, Middleton, Caylor, & Gupta, 2008; Petrosino et al., 2003), 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).

What has not been explored, however, is how learners and youth participate in different aspects of work with data through the lens of engagement. 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). Engagement is often conceptualized as a meta-construct, that is, one that is made up of other constructs (Skinner & Pitzer, 2012; Skinner, Kindermann, & Furrer, 2009. 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, situational or momentary 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 in STEM (Greene, 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 (Berland et al. 2016). 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. While sometimes defined in terms of extra-curricular involvement or following directions, behavioral engagement can be considered working hard on learning-related activities (Fredricks et al., 2004; Singh, Granville, & Dika, 2002). Finally, affective engagement can be defined as emotional responses to activities, such as being excited, angry, or relaxed (Pekrun & Linnenbrink-Garcia, 2012).

Also, some contextual 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 is being both challenged by, and good at, a particular activity (Shernoff et al., 2016). Conceptualizing perceptions of challenge and competence as conditions, rather than factors that influence engagement, can be a 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, *challenge* and *competence* can be considered together with engagement, as in this the present study.

**Key Factors That May Affect Youths’ Engagement**

Past research suggests youths' characteristics that impact their cognitive, behavioral, and affective engagement. For example, interest in the domain of study has been found to impact students’ engagement (Shernoff et al., 2003; Shernoff et al., 2016; Author and colleagues, 2013). In addition, some studies have shown that there are gender-related differences in engagement (Author and colleagues, 2011).

How teachers support learning practices (Author and colleagues, 2017) also impacts how students’ engage. Particularly concerning work with data = (Lehrer & Schauble, 2015; Wilkerson, Andrews, Shaban, Laina, & Gravel, 2016), sustained support from those leading youth activities is an essential component of learners being able to work with data. Thus, how youth activity leaders plan and enact activities related to work with data can have a large impact on students' engagement. Furthermore, because of the importance of work with data across STEM domains, carrying out ambitious activities focused on work with data may plausibly have a substantial impact on the extent to which youth engage in summer STEM program settings.

**Studying Engagement as a Dynamic Construct and the Experience Sampling Method**

Because engagement is understood as both context-dependent and multi-dimensional, it is challenging to measure the engagement construct in authentic ways in empirical studies. One methodological approach that is well suited to the context-dependent and multidimensional nature of engagement is the Experience Sampling Method (ESM; Author and colleagues, 2007; see Author and colleagues, 2017; Turner & Meyer, 2000; and Sinatra et al., 2015 for recent applications to engagement research). ESM typically involves asking participants short questions about their experiences at random moments in response to a digital signaling device. ESM is particularly well-suited to understanding the context-dependent nature of engagement because students provide repeated brief reports about their experience as it is happening, minimally interrupting them from the activity they are engaged in (Hektner et al., 2007). The ESM approach is both sensitive to changes in engagement over time, as well as between learners and allows us to understand engagement and how factors impact it in more nuanced and complex ways (Turner & Meyer, 2000).

One powerful and increasingly widely used way to examine the multi-dimensional nature of learning-related constructs (such as engagement) is the use of profiles of or groups of variables that are measured. This profile approach is especially important given the multidimensional nature of engagement. We know from past research that engagement can be explained through different patterns among its individual dimensions (Bergman & Magnusson, 1997; Bergman et al., 2003). Because learners’ engagement includes cognitive, behavioral, and affective aspects experienced together at the same time, it can be experienced as a combined effect that is categorically distinct from the effects of the individual dimensions of engagement. This combined effect can be considered as profiles of engagement.

In past research, profiles are commonly used as part of what is described as person-oriented approaches (Bergman & Magnusson, 1997; Bergman, Magnusson, & El Khouri, 2003), those used to consider the way in which psychological constructs are experienced together and at once in the experiences of learners. There are some recent studies taking a profile approach to the study of engagement (i.e., Salmela-Aro, Moeller, Schneider, Spicer, & Lavonen, 2016a; Salmela-Aro, Muotka, Alho, Hakkarainen, & Lonka, 2016b; Van Rooij, Jansen, & van de Grift, 2017; Authors and colleague, 2018a), though none have done so to study youths' engagement in work with data.

Some past studies have considered profiles of cognitive, behavioral, and affective aspects of engagement (i.e., Salmela-Aro et al., 2016b; Authors and colleague, 2018a). A potential way to extend this past research is to account for not only engagement (cognitive, behavioral, and affective), but also the intricately connected perceptions of challenge and competence. This analytic approach is especially important since a profile approach emphasizes the holistic nature of engagement and the impact of not only external but also intra-individual factors. Thus, the profiles of engagement may usefully include not only the dimensions of engagement but also youths’ perceptions of how challenging the activity they were doing is and of how competent at the activity they are

**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) and a Latent Profile Analysis (LPA; Harring & Hodis, 2016) analytic approach in order to identify profiles representing commonalities in the ways students experience engagement. In particular, an ESM 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.

The framework for this study (Figure 1) conceptualizes engagement as a multi-dimensional construct consisting of cognitive, behavioral, and affective dimensions that can be modeled as *profiles of engagement*. How students engage in work with data may be impacted by specific activities, conceptualized here as *the five key aspects of work with data* and Youth’s characteristics such as gender, pre-program interest in STEM, and being an individual from a group that is under-represented (in STEM).

[INSERT FIGURE 1 ABOUT HERE]

Guided by this conceptual framework, we examined four research questions to better understand students’ engagement in working with data:

1. What profiles of engagement emerge from data collected via ESM in the programs?
2. How does work with data relate to the profiles of engagement?

**Method**

In this section, we describe the research design, context, and participants, followed by the procedure, measures, and their associated data sources, and the data analyses necessary to answer each of the research questions.

**Research Design**

This study makes use of an innovative ESM methodology (Hektner et al., 2007)within a correlational study design to understand youths’ experience in work with data in-the-moment. Using ESM, youth were provided with mobile phones that signaled them at random times (within intervals, so that the signals were not too near or far apart) to respond to short questions about their cognitive, behavioral, and affective engagement, and their perceptions of challenge and competence at the time of the signal.

**Context**

The setting for the present study was nine out-of-school STEM programs in the Northeastern United States during 2015. Youth spent around three hours per day, 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. As can be seen in the descriptions of the programs provided in Appendix A, many of the programs aimed at involving youth in work with data.

**Participants**

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

## [INSERT TABLE 1 ABOUT HERE]

## Procedure

Before the beginning of the programs, youth completed a pre-survey that included questions about their experience in STEM. In addition, demographic information for the youth was collected from the programs. At the programs’ beginning, youth were introduced to the ESM data collection method. ESM data were collected two days each week, for three weeks (weeks 2-4 of the programs). Each day, youth were signaled four times[[1]](#footnote-1). In addition to the collection of ESM data during the programs, the programs were recorded using portable video cameras by research team members only on the days during which ESM data were collected. In all of the programs, about equal video-recording time was dedicated to classroom and field experiences.

**Data Sources and Measures**

The data sources and measures for this study are the pre-program survey, ESM, video-recordings, and the demographic information we collected from the programs.

**ESM measures of engagement for the profiles**. Measures for engagement were created from five short ESM questions (Table 2). Youth were asked to respond to the five ESM questions each time they were. Each of the ESM items consisted of the item text and the following four item response options: (1) Not at all, (2) A little, (3) Somewhat, and (4) Very Much.

[INSERT TABLE 2 ABOUT HERE]

**Measures from the video recordings for work with data**. Codes for work with data were generated on the basis of the activity that the youth activity leaders were facilitating. The coding frame we used was adapted from the STEM-Program Quality Assessment (STEM-PQA; Forum for Youth Investment, 2012), an assessment of quality programming in after-school programs. In particular, the five aspects of work with data were identified through codes from the STEM-PQA because this coding frame included codes for work with data (See Appendix B for the alignment between the STEM-PQA and the aspects of work with data). The coding frame was originally designed to describe the types of activities the program leaders were facilitating**.**

### 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.

### [INSERT TABLE 3 ABOUT HERE]

Out of the 248 instructional episodes, 236 were code-able for work with data; for the 12 that were not code-able, 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

. Three items adapted from Vandell, Hall, O’Cadiz, and Karsh (2012) were used to measure youths’ pre-program interest in STEM. 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. Authors and colleagues (2018b) for more details on this (use of the maximum value) measurement approach. Like the ESM measures, youth were asked to report their agreement with the three items on a 1 (A little) to 4 (Very much) scale. Reliability and validity information for this scale is presented in Vandell et al. (2008). 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).

**Demographic measures.** Demographic information for youths’ gender and their racial and ethnic group were used to construct dichotomous variables for gender and being from 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, or Native American.

**Data Analysis**

**Latent profile analysis.** Latent Profile Analysis was used to identify engagement profiles. A key benefit of the use of LPA as used for the analysis related to this question is that it outputs the *probability* of an observation being a member of a profile. 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. For the present study, six model types were considered. The [name removed for peer review] package (Authors and colleagues, 2018c) and the MPlus software (Muthen & Muthen, 1998-2017) were used to carry out LPA through open-source statistical software we developed. 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 of the number of profiles (the bootstrapped LRT [BLRT]), and concerns of interpretability and parsimony.

**Multi-level models.** Mixed effects models were used to study the relations between the aspects of work with data and youth characteristics and engagement. We used mixed effects models to 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 research questions 3 and 4 use random effects for youth, instructional episode, and program effects , and the first set of models that we ran included only these effects. This was so that we could calculate the Intra-class Correlation Coefficient, the ICC, which we carried out individually for each profile. The ICC represents the proportion of the variation in each profile that is attributable to each youth, the instructional episode, or the program youth were in. These values provide insight into the sources of variability in youths’ engagement and are helpful for interpreting the presence (or absence) of the effects of variables at each of the levels (see Figure 1 for a depiction of which variables are at which levels). For all six profiles, the *ICC*s represent the variability (the proportion of variance explained) associated with each of the levels for each profile.

# Results

## Descriptive statistics for the engagement measures

Correlations (first-order Pearson) and the frequency, range, mean (*M*), and standard deviation (*SD*) are first reported for all variables. In addition, the frequencies of the codes for aspects of work with data and the numbers of responses by youth, program, and instructional episode are presented.First, descriptive statistics for the five engagement variables that were used to estimate the profiles were calculated. 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. 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)).

## Correlations among the study variables

In order to understand whether the relations between the variables were (generally) as expected, correlations between the variables that were used to create the profiles of engagement (cognitive, behavioral, and affective engagement [*Cog., Beh.,* and *Aff. Eng.*] and challenge and competence [*Chall.* And *Comp.*]), pre-program interest in STEM (*Pre-interest*), and the variables for the five aspects of work with data (abbreviated in Figure 1) were calculated, as presented in Table 4. In addition, relations between these variables and those for the five aspects of work with data were identified.

## [INSERT TABLE 4 ABOUT HERE]

## Research Question #1: How frequent is work with data in summer STEM programs?

### The codes for the aspects of work with data (described above in the section on the measures) were counted and presented as a proportion of the number of code-able instructional episodes. Results from this analysis provide initial insight into how often each of the aspects took place during the programs. Of the 236 instructional episodes used in the analysis, 170 (72%) involved one or more of the five aspects of work with data. As presented in Table 5, the five aspects of work with data occurred regularly.

[INSERT TABLE 5 ABOUT HERE]

### Making observations was the least frequent of the five aspects, occurring in 24% of instructional episodes. Data modeling was the next least frequent aspect, occurring in 29% of the episodes, followed by asking questions (38%), generating data (43%), and communicating findings (43%). As suggested by the frequencies presented in Figure 2, 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.

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

On the basis of multiple fit indices and criteria, a model with equal variances and covariances fixed to zero was found to be most reasonable. This is a common model type (and is the default for the MPlus software) that allows for mean differences between the profiles as well as the variables’ variances across all of the profiles to be estimated. 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 are presented in Table 6.

[INSERT TABLE 6 ABOUT HERE]

The profiles for this solution are 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. Figure 3 shows the profiles with the raw data (not transformed). 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 of the composition of the profiles.

[INSERT FIGURE 2 ABOUT HERE]

[INSERT FIGURE 3 ABOUT HERE]

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. The results for research questions 3 and 4 use this solution and the six profiles in subsequent analyses. In order to provide evidence for whether the profiles are distinct from one another, we used a MANOVA 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*-tests associated with each ANOVA were also statistically significant.

**Research Questions #3 and #4: How does work with data relate to the profiles of engagement?**

In advance of studying the relations between work with data and youth characteristics and the profiles of engagement, we explored how much variability was present for each of the profiles at the instructional episode, youth, and program levels through the ICCs output that is part of the mixed effects model output. The ICCs are presented in Figure 4. Note that these ICCs are from models without any predictors included, or null models, and are useful for understanding subsequent results from models that include predictor variables (for both work with data and youth characteristics).

[INSERT FIGURE 4 ABOUT HERE]

These results show how much variability in the profiles was systematic at these different levels and was potentially attributable to youth, instructional episode, and program. There could be systematic variability at the youth level for example for the *Full* profile, for example, but there could be very little variability for the, *Universally Low* profile at this level. 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 were larger: they ranged from .093 to .432. These *ICCs* 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.

To understand which aspects of work with data and youth characteristics were related to the profiles, 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 separately and together, the results from the combined model were used to provide answers to research questions 3 and 4. Specifically, pre-program interest in STEM, gender and URM status were included as predictors along with the aspects of work with data. The results are interpreted in terms of 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 percent more likely to report a response in that particular profile.

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 *Universally low* 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 (more conservative and recommended in recent research) Kenward-Rogers approximation (Halekoh & Hojsgaard, 2014).

[INSERT TABLE 7 ABOUT HERE]

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 also 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.

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.

There were not any statistically significant effects of youths’ URM status. This null finding may be a function of the large proportion of youth from under-represented (in STEM) racial and ethnic groups, which resulted in a lack of statistical power to detect a difference. 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 few youth who were *not* from under-represented groups in the sample. 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

In this study, engagement was used to understand the experience of youth working with data during summer STEM programs. We identified six profiles of engagement using an innovative technique (i.e., LPA) that aims to identify the ways that variables group together. The profiles represented different configurations of how youth were working hard, learning, enjoying themselves, and feeling challenged and competent at the time they were signaled during various STEM activities. 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. In the remainder of this section, we discuss why findings with respect to a) the frequency of work with data in summer STEM programs, b) the nature of youths’ engagement, and c) what relates to youths’ engagement. Also—and in light of the limited relations of work with data and youths’ characteristics to their engagement—we discuss some of the limitations of this study as well as recommendations for future research and implications for practice.

## Key Findings Regarding Work With Data in Summer STEM Program

First, results showed that work with data was quite common in the summer STEM programs. There was also 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 relative to other program activities—and differences between the aspects of work with data. They align with what may be expected given past research: Such programs are designed to engage youth in the practices of STEM domains, including, and as we argued earlier *especially* those relating to work with data (Dabney et al., 2012; Elam et al., 2012). Even still, this study is the first of its kind in that it actually affords the measurement the proportion of time youth spent in work with data. As such it reflects the value of combining video-recorded data to provide insight into key practices such as those related to work with data.

## Key Findings Related to Youth Engagement in Summer STEM Program

As presented earlier, six profiles of engagement were identified through the use of LPA. Little work has examined profiles of engagement, and so these results can provide new insight into the nature of youth engagement in summer STEM programs. Authors and colleague (2018a) have also examined profiles of engagement. In Author and colleague’s (2018a) work, however, the profiles were constructed from indicators cognitive, behavioral, and affective engagement—but not perceptions of challenge and competence (as in this study). They found six profiles, some of which partially overlap with those found in the present study: *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, these findings are noteworthy, but not yet suggestive of distinct configurations of engagement (such as high levels of cognitive engagement with high levels of challenge—yet moderate levels of the other variables) in STEM education.

On the other hand, the *Only behavioral*, *Only affective*, and *Engaged and competent but not challenged* profiles were found in the present study, but not in Author and colleague’s (2018) study. Specifically, the *Engaged and competent but not challenged* profile seems to suggest a type of engagement that is particularly unique to summer STEM programs. When responding in ways characterized by this profile, 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. 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 suggest that lower challenge may *not*, as would be anticipated given theory and past research (i.e., Csikszentmihalyi, 1990; Shernoff et al., 2016), be associated with lower engagement. This may suggest 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 engaging.

**The Relationships Between Work With Data, Youth Characteristics, and Engagement**

Some key relationships between work with data and youth characteristics and youths’ engagement were found, but these relations were minimal. 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. 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.

There are other possible reasons, though, too, for the minimal relations. Another possibility is that the novel analytic approach or the measures used also had impacts; but, again, the small variability at the instructional episode level is likely a greater factor than these, and a review of the correlations between the aspects of work with data and the variables used to create the profiles showed minimal relations. Taken together, it seems that the major reason for limited relations between work with data and youth engagement is that youth simply did not engage very differently from instructional episode to instructional episode.

Even with the limited findings *overall* in mind, 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. 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 perceived by youth to be demanding and important. Modeling, too, is important. It has been described as the central scientific and engineering practice (Lehrer & Schauble, 2015; Weisberg, 2012). 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 shift, from learning about STEM content to working to create STEM knowledge is one of the affordances of modeling (Berland et al., 2016; Schwarz et al., 2009).

To step back from specific relationships, these findings overall suggest that work with data may not be more engaging *per se*. Instead, it may be how youth engage in work with data that matters most. In the midst of science and mathematics reform efforts that emphasize *doing* what scientists and mathematicians do—by engaging in scientific and engineering practices—engaging in practices in merely procedural ways may not help students to become more capable in the domains that they study. This point is echoed in recent scholarship pointing out the importance of engaging in science in ways that are both ‘minds-on’ as well as ‘hands-on’ (Furtak & Penuel, 2018) and in the importance of engaging learners in practices that are meaningful to students and in terms of the discipline (Berland et al., 2016). So, engaging youth in work with data may, on its own, not be more engaging (apart from in a few cases, i.e., when generating or modeling data). Instead, *how* youth work with data may be the key: More substantial relationships between work with data and youth engagement may be found within STEM classes or programs that are focused explicitly on work with data or data science. In addition, how teachers support students to work with data through orchestrating whole-class discussions may have an effect on how engaging such work is to youth.

There were not many significant relations in terms of youths’ gender, URM status, or pre-program interest, which was surprising because there *was* substantial variability in the profiles of engagement at the youth level. This was also surprising as 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; Author and colleague, 2008). There was one significant relation: 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. This finding is in line with past research suggesting a relationship 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 suggests that it may be engaging to students to involve them in activities that present greater challenge to them.

## Limitations and Recommendations for Future Research

One limitation concerns the programs: The programs included 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. Future research may find it useful to explore the nature of work with data in programs (or activities) designed specifically around working with data: Some data science and statistics-related curricula have recently been developed for this purpose (see Unit 5 on Computing and Data Analysis of the *Exploring CS* curriculum; Exploring CS, 2018).

There are some 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. However, there were some divergences that may have had an impact on 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).

Finally, the little variability at the instructional episode level is 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.

## 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., 2016; Hancock et al., 1992; McNeill & Berland, 2017; 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. 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).

**Conclusion**

We explored how more than 200 youth from nine summer STEM programs experienced work with data through the lens of contemporary engagement theory. Using ESM and LPA, we identified six distinct profiles of youth engagement, which indicated how youth perceived their experiences (in terms of their cognitive, behavioral, and affective engagement, and challenge and competence) in-the-moment. While we found that the five aspects of work with data were common, these aspects were not highly related to youths’ engagement: apart from the generating and modeling data aspects (which were related only to the *Full* engagement profile), we found that work with data alone did not explain the differences in youths’ engagement. Some key reasons why work with data was not strongly related to youths’ engagement is the minimal variability at the instructional episode level (the level at which work with data was measured), the difficulty of measuring work with data, and how work with data was carried out in the context of the summer STEM programs. Future research—and the design and implementation of curricula in the future—may find it useful to explore programs (and classrooms) focused particularly on engaging youth in work with data, an aim which will be easier as more research and curriculum development related to data science education advances.

# References

Author and colleagues (2007). *Book.*

Author and colleague (2008). *Journal article*.

Author and colleagues (2011). *Journal article*.

Authors and colleagues (2013). *Journal article*.

Author and colleagues (2017). *Journal article*.

Authors and colleagues. (2018a). *Journal article.*

Authors and colleague (2018b). *Journal article.*

Authors and colleagues (2018c). *Journal article.*

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.

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. (1990). Flow: The psychology of optimal experience. New York, NY: Harper and Row.

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.

Exploring CS. (2018). Exploring Computer Science [educational curriculum]. Retrieved from http://www.exploringcs.org/

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

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.

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.

Konold, C., Finzer, W., & Kreetong, K. (2017). Modeling as a Core Component of Structuring Data. *Statistics Education Research Journal*, *16*(2).

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.

McCaffrey, D. F., Lockwood, J. R., Koretz, D. M., & Hamilton, L. S. (2003). *Evaluating Value-Added Models for Teacher Accountability. Monograph*. RAND Corporation. PO Box 2138, Santa Monica, CA 90407-2138.

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.

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.

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.

1. 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). [↑](#footnote-ref-1)