Examining youth engagement during learning activities that involve work with data: An Experience Sampling approach

Joshua M. Rosenberg

2018-05-20

Contents

Chapter 1

Introduction

How citizens plan our day-to-day lives, communicate, and learn are increasingly impacted by data. These sources of data–either quantitative *or* qualitative–are created by us, for us, and about us, although at present opportunities for learners to analyze data in educational settings remain limited. Work with data includes broad processes of collecting, creating, modeling data, and eveb asking questions that can be answered with data.

Working with data, then, is more than just crunching numbers, or interpreting a figure created by someone else. Rather, it is about making sense of phenomena in the world (or solving problems), a point particularly relevant to those interested in the educational role and implications of working with data (Lee & Wilkerson, in press; Singer, Hilton, & Schweingruber, 2006; Wild & Pfannkuch, 1999). Aspects of work with data cut across STEM domains and are recognized as core competencies across recent curricular documents. For example, the *Next Generation Science Standards* and the *Common Core State Standards* (in mathematics; National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010; NGSS Lead States, 2013) both highlight the role of authentic work work with data. Scholars have pointed out the benefits of working with data for learners as young as two years old (Gopnik, & Sobel, 2000).

In supporting teachers and learners' data analysis efforts, scholars have examined a variety of learning-related practices. In particular, past research has focused on mathematical practices, or activities, like generating measures of phenomena and creating data models (English, 2012; Lehrer & Romberg, 1996; Lesh, Middleton, Caylor, & Gupta, 2008). Findings from this area of research suggest that engaging in these practices "has an exceptionally high payoff in terms of students' scientific reasoning" (Lehrer & Schauble, 2015, p. 696) and can highlight the utility of mathematics for students' lives (Lesh, Middleton, Caylor, & Gupta, 2008).

Because engaging in data-related practices seems to be so potentially beneficial to learners (Lee & Wilkerson, in press; National Research Council, 2012), we need to better understand the nature of learners' engagement in learning activities that involve various aspects of work with data. To date, past research shows that using an engagement framework to characterize students' learning activities is highly informative. One's engagement in learning tasks is a key outcome in its own right and may be an antecedent of changes in other outcomes, such as their well-being, achievement and pursuit of an area of study or career (Sinatra, Heddy, & Lombardi, 2015; Wang, Chow, Hofkens, & Salmela-Aro, 2015; Wang & Eccles, 2012). However, research has not examined engagement in these data-related activities in particular.

The purpose of this study, then, is to examine youth engagement in a variety of learning activities that involve work with data. Engagement in work with data is explored in the context of outside-of-school STEM enrichment programs carried out during the summer and work with data is considered in terms of specific aspects identified from past research, such as asking questions and generating and modeling data. 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 et al., 2015). 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.

Chapter 2

Literature Review

The framework for this study is informed by work on STEM-related learning practices, or activities, student engagement, and analytic approaches to modeling multidimensional constructs. In this review of literature, I define work with data as a key practice across STEM domains. I also describe and justify a multi-dimensional framework for understanding engagement, and then review an approach to analyzing data that is ideal for capturing this multidimensionality.

2.1 Defining Work with Data

Some scholars have focused on a few key pieces of data analysis, connected through the use of "data to solve real problems and to answer authentic questions" (Hancock et al., 1992, p. 337). This focus on solving real problems or answering authentic questions–rather than being taught and learned as isolated skills–is an essential part of work with data having the most educational benefits to learners (National Research Council, 2012; see Lehrer and Schauble [2012] Windschitl, Thompson, & Braaten [2018] for excellent, practice, in-depth examples of work with data being used as part of instructional approaches). This approach has primarily been taken up by mathematics educators and is reflected in statistics curriculum documents (Franklin et al., 2007). In science settings, where answering questions about phenomena serve as the focus of activities, it shares features of the process of engaging in scientific and engineering practices, but has been less often studied.

Work with data has been conceived in different ways. For some specific examples from different domains, see Lee and Wikerson's (in press) forthcoming summary report for the National Academy of Sciences and Wild and Pfannkuch (1999), Franklin et al. (2007), and Lehrer and Schauble (2004). Because there is not an agreed upon definition of work with data–particularly across subject area domains–I 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 (Figure 1) 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 with respect to 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 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

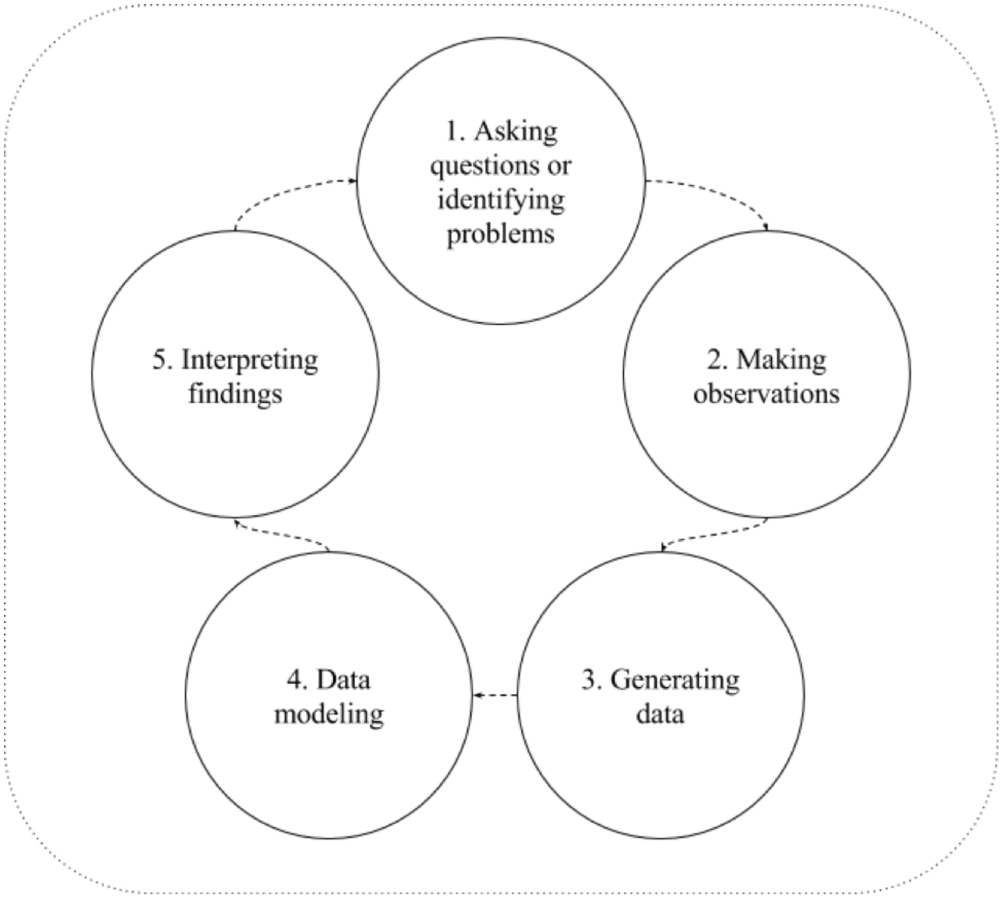


Figure 2.1: Work with data in STEM education settings

- *Interpreting and communicating findings*: Activities related to identifying a driving question regarding the phenomena that the question is about

These five synthesized aspects of work with data are a part of a cycle because not only does each part follow that before it, but also because the overall process is iterative: interpreting findings commonly 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 intermediate findings are critiqued and subject to critique, and revised over time (McNeill & Berland, 2017; Lee & Wilkerson, in press).

2.2 The role of working with data in STEM learning environments

Working with data can serve as an organizing set of practices for engaging in inquiry in STEM learning settings (Lehrer & Schauble, 2015). Data are both encountered and generated by learners, and so opportunities for learners to work with data provide many opportunities to leverage their curiosity because processes of inquiry can be grounded in phenomena that learners themselves can see and manipulate or phenomena that learners are interested in. Also important, becoming proficient in work with data can provide learners with an in-demand capability in society, owing to the number of occupations, from education to entrepreneurship, that demand or involve taking action based on data (Wilkerson & Fenwick, 2017). 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.

Recent educational reform efforts emphasize work with data (i.e., the scientific and engineering practices in the NGSS and the standards for mathematical practice in the Common Core State Standards). However, work with data is uncommon in many classroom settings (even classrooms emphasizing recent science education reform efforts; McNeill & Berland, 2017; Miller, Manz, Russ, Stroupe, & Berland, in press), and so learning environments suited to engaging in work with data, but not explicitly designed to support it, may be valuable to study because they may serve as incubators of these rare and challenging learning activities.

Outside-of-school programs are a potentially valuable setting to explore engagement in work with data because of the combined pedagogical and technical expertise of their staff and the activities learners do during their participation in them. Staff for these programs includes educators and scientists, engineers, and others with the technical experience. Additionally, the programs were designed to involve learners in the types of real-world practices experienced by experts in STEM disciplines. Attendance in such programs is associated with many benefits to learners (Green, Lee, Constance, & Hynes, 2013; see Lauer, Akiba, Wilkerson, Apthorp, Snow, & Martin-Glenn, 2006, for a comprehensive review). These programs are also selected because little research has examined how data are part of the experiences of youth in out-of-school-time programs, despite its place as one of a few core practices in STEM. While these reasons to study work with data focus on outside-of-school programs, they are also germane and applicable to more formal learning environments, such as classrooms, in which teachers want to design opportunities for their learners to work with data.

2.3 What We Know (And Do Not Know) About How Youth Work with Data

Scholars have researched cognitive capabilities as outcomes related to work 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 et al., 2001; Gopnik & Sobel, 2000), often from a Piagetian, individual-development focused tradition (i.e., Piaget & Inhelder, 1969). A key outcome of engaging in work with data has to do with how learners account for variability (Lehrer, Kim, & Schauble, 2007; Petrosino, Lehrer, & Schauble, 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 (and covariability).

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, such as the role of measurement in exposing learners in a direct way to sources of variability (Petrosino et al., 2003), role of simulation to learn about sampling distributions (Stohl & Tarr, 2002), and use 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, 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 (Lee & Hollebrands, 2008; Lehrer, Kim, & Schauble, 2007).

Despite this past research, how learners and youth participate in different aspects of work with data in terms of engagement theory has not been examined. Consider the practice of modeling data, commonly described as a—-or *the*—-key part of many applied 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.

2.4 Engagement in General and in STEM Domains

The nature of engagement is discussed in terms of general features that have been identified across content area domains, conditions that support engagement, and differences between engagement in general and in STEM settings. This is followed by a discussion of two key features of engagement: its dynamic, or context-dependent, characteristics and its multidimensional nature. Finally, methods for capturing these two features empirically through the Experience Sampling Method, or ESM–and how this (multidimensional) data can be analyzed–are described.

Engagement is defined in this study as active involvement, or investment, in activities (Blumenfeld et al., 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 exemplary 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, 2012) 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 of 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; Miller et al., in press).

The emphasis on developing new knowledge and capabilities through 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, cognitive engagement is defined for this study in terms of learning something new or getting better at something. While sometimes defined in terms of extra-curricular involvement or following directions, behavioral engagement is defined in this study as working hard on learning-related activities (Fredricks et al., 2004; Singh, Granville, & Dika, 2002). Finally, affective engagement is defined as affective responses to activities, such as being excited, angry, or relaxed (Pekrun & Linnenbrink-Garcia, 2012).

Finally, there are some key conditions that facilitate engagement. Emergent Motivation Theory (EMT; Csikszentmihalyi, 1990), provides a useful lens. From EMT, a key moment-to-moment condition for engagement is how difficult individuals perceive an activity to be, or its *perceived challenge*. Another key condition is how good at an activity individuals perceive themselves to be, or their *perceived competence*. What is most important–and necessary in terms of being engaged–is being both challenged by and good at a particular activity. Thus, these two conditions are analytically considered together with engagement, as described in the section below on measuring engagement.

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.

2.5 Individual factors that may effect youths' engagement

Past research suggests learners or youths' characteristics, such as their interest in the domain of study, impact their cognitive, behavioral, and affective engagement (Shernoff et al., 2003; Shernoff et al., 2016; Shumow, Schmidt, & Zaleski, 2013). These are both moment-to-moment, context-dependent conditions that support engagement (like those discussed above, perceptions of challenge and competence) as well as youth-specific factors. These factors are at the level of individual differences (i.e., youths' more stable interest in STEM domains), and may impact engagement, as described in this section.

A factor that can support engagement concerns teacher support for specific learning-related practices (Strati, Schmidt, & Maier, 2017). Particularly concerning work with data, which is demanding not only for learners but also teachers (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 large impact on the extent to which youth engage in summer STEM program settings. Consequently, this study considers work with data through the use of a coding frame that characterizes the extent to which teachers are supporting specific STEM practices in their instruction, including aspects of work with data.

Other factors that impact youths' engagement are individual characteristics and differences. In recognition of differences among learners in their tendency to engage in different (higher or lower) ways in specific activities based in part on individual differences (Hidi & Renninger, 2006), learners' interest in STEM before the start of the programs is also considered as a factor that can impact engagement. Knowing about whether and to what extent youths' interest *before* participating in summer STEM programs explains their engagement *during* them is a key question on its own right. It is also important in terms of properly understanding the effects of other factors, such as working with data, above and beyond the effect of pre-program interest. In addition to this interest, gender and the racial and ethnic group of students is also considered, as past research has indicated these as factors that influence engagement in STEM (Bystydzienski, Eisenhart, & Bruning, 2015; Shernoff & Schmidt, 2008). To include the racial and ethnic group of students, being part of an under-represented minority (URM) group is considered. To sum up, youths' pre-program interest, gender, and URM group membership are considered as individual factors that may impact youths' engagement.

2.6 Challenges of Measuring Engagement as a Contextually-Dependent and Multidimensional Construct

Because of the way engagement has been thought of as having context-dependent characteristics and being multi-dimensional, it is challenging to use engagement (when conceptualized in such a way) in empirical studies. One methodological approach that has benefits in terms of both the context-dependent and multi-dimensional nature of engagement is the ESM. A number of scholars have explored or extolled benefits to its use in their recent work (e.g., Strati et al., 2017; Turner & Meyer, 2000; Sinatra et al., 2015).

This study employs the Experience Sampling Method (ESM; Hektner, Schmidt, & Csikszentmihalyi, 2007) where learners answer short questions about their experience when signaled. ESM involves asking—usually using a digital tool and occasionally a diary—to ask participants short questions about their experiences. ESM is particularly well-suited to understanding the context-dependent nature of engagement because students answered brief surveys about their experience when they were signaled, minimally interrupting them from the activity they are engaged in and also seeking to collect measures about learners' experience when signaled (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).Though time-consuming to carry out, ESM can be a powerful measure that leverages the benefits of both observational and self-report measures, allowing for some ecological validity and the use of closed-form questionnaires amenable to quantitative analysis (Csikszentmihalyi & Larson, 1987). Despite the logistic challenge of carrying out ESM in large studies, some scholars have referred to it as the *gold standard* for understanding individual's subjective experience (Schwarz, Kahneman, & Xu, 2009).

Research has shown us how the use of ESM can lead to distinct contributions to our understanding of learning and engagement. This work also suggests how ESM can be put to use for the purpose of the present study. For example, Shernoff, Csikszentmihalyi, Schneider, and Shernoff (2003) examined engagement through the use of measures aligned with flow theory, namely, using measures of concentration, interest, and enjoyment (Csikszentmihalyi, 1997). In a study using the same measures of engagement (Shernoff et al. (2016) used an observational measure of challenge and control (or environmental complexity) and found that it significantly predicted engagement, as well as self-esteem, intrinsic motivation, and academic intensity. Schneider et al. (2016) and Linnansaari et al. (2015) examined features of optimal learning moments or moments in which students report high levels of interest, skill, and challenge, as well as their antecedents and consequences. Similar to ESM in that through its use engagement can be studied in a more context-sensitive, still other scholars have used daily diary studies to examine engagement as a function of autonomy-supportive classroom practices (Patall, Vasquez, Steingut, Trimble, & Pituch, 2015; Patall, Steingut, Vasquez, Trimble, & Freeman, 2017). This past research that used ESM (or daily diary studies) to study engagement has shown that ESM can be used to understand fine-grained differences in learning activities, such as the aspects of work with data that are the focus of this study.

Other research shows us that there are newer approaches to analyzing ESM data that can contribute insights into the context-dependent nature of engagement in a more fine-grained way. For example, Strati et al. (2017) explored the relations between engagement to measures of teacher support, finding associations between instrumental support and engagement and powerfully demonstrating the capacity of ESM to understand some of the context-dependent nature of engagement. Similarly, Poysa et al. (2017) used a similar data analytic approach as Strati et al. (2017), that is, use of crossed effects models for variation within both students and time points, both within and between days. These studies establish the value of the use of ESM to understand the context-dependent nature of engagement and that such an approach may be able to be used to understand engaging in work with data. Additionally, these recent studies (particularly the study by Strati and colleagues) show that how effects at different levels are treated, namely, how variability at these levels is accounted for through random effects as part of mixed effects models, is a key practical consideration for the analysis of ESM data.

One powerful and increasingly widely used way to examine context-dependent constructs, such as engagement holistically is the use of *profiles*, or groups of variables that are measured. This profile approach is especially important given the multidimensional nature of engagement. Profiles are commonly used as part of what are described as person-oriented approaches, those used to consider the way in which psychological constructs are experienced together and at once in the experiences of learners. In the context of the present study, this approach can help to identify naturally occurring profiles of engagement, or engagement as reported by youth via ESM during particular moments. Note that in the present study, ESM involves asking youth about to report on their experience at the time they were signaled (rather than, for example, before or after the program, which traditional surveys are well-suited for).

In this study, profiles are used in the service of understanding how students engage in work with data in a more holistic way. 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; Schmidt, Rosenberg, & Beymer, 2018), though none have done so to study youths' engagement in work with data.

The profile approach has an important implication for how we analyze data collected from ESM about youths' engagement, in particular when we consider how to understand engagement as a multi-dimensional construct, and one with momentary, or instructional episode-specific, conditions (Csikszentmihalyi, 1990). We know from past research that engagement can be explained in terms of different patterns among its individual components (Bergman & Magnusson, 1997), in the present case its cognitive, behavioral, and affective components. 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.

Past studies have considered profiles of cognitive, behavioral, and affective aspects of engagement. For example, to account for the context-dependent nature of engagement, some past studies have used other measures to predict engagement, such as use of in-the-moment resources and demands (Salmela-Aro et al., 2016b) and the use of instructional activities and choice (Schmidt et al., 2018). A potential way to extend this past research is to account for not only engagement (cognitive, behavioral, and affective), but also its key conditions (challenge and competence). This 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. Accordingly, youths' perceptions of the challenge of the activity and of their competence at it are used along with the measures of engagement to construct profiles of engagement. Thus, the profiles of engagement include youths' responses to five ESM items for their cognitive, behavioral, and affective engagement and their perceptions of how challenging the activity they were doing is and of how competent at the activity they are.

2.7 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 better understand learners' experiences working with data. The present study does this through the use of contemporary engagement theory and innovative methodological and analytic approaches. Doing this can help us to understand work with data in terms of learner's experience, 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 addition to this need to study engagement in work with data through the lens of engagement, no research that I am aware of 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.

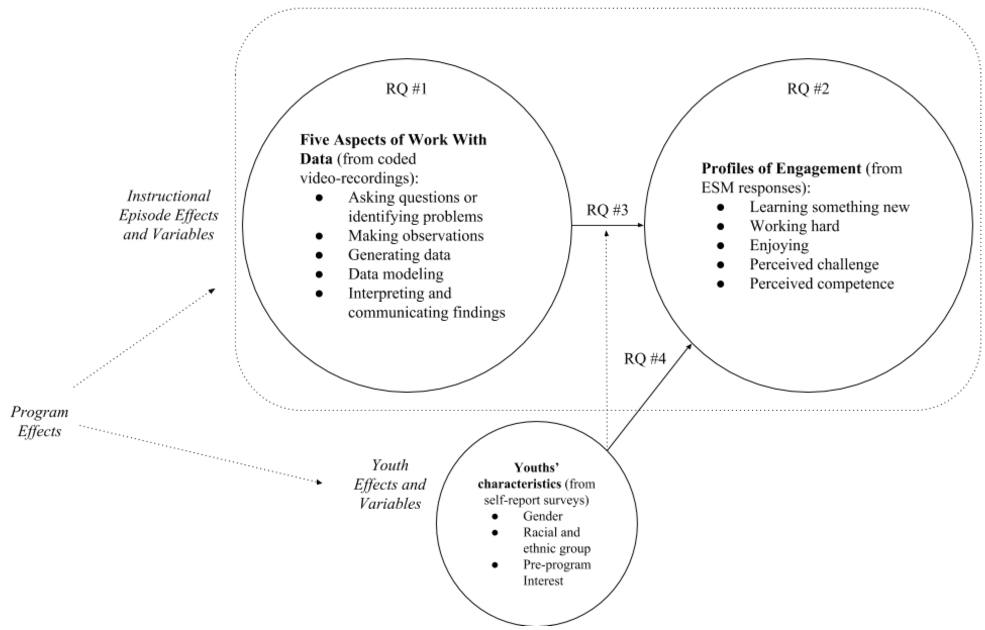


Figure 2.2: A conceptual framework for this study with research questions labeled

2.8 Conceptual Framework and Research Questions

To summarize, the present study is about how learning activities involving various aspects of work with data can be understood in terms of engagement. Its context is out-of-school-time STEM enrichment programs designed to meet guidelines for best practices. The conceptual framework in the present study is presented in Figure 2 and is laid out in the remainder of this section.

There are five aspects of work with data synthesized from past research (i.e., Hancock et al., 1992; Lehrer & Romberg, 1996; Wild & Pfannkuch, 1999):

1. Asking questions or identifying problems
2. Making observations
3. Generating data
4. Data modeling
5. Interpreting and communicating findings

In Figure 2, 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. In addition, a pre-program measure of youths' pre-program interest in STEM, along with youths' gender and URM status, are hypothesized to be associated with the profiles and the relations of work with data and the profiles.

The ESM responses that make up the profiles are associated with different groups. These groups, which may introduce dependencies that violate statistical assumptions of the independence of the responses, are commonly considered in the Hierarchical Linear Modeling (also known as multi-level or mixed effects modeling) literature as *random effects* (Gelman & Hill, 2007; West, Welch, & Galecki, 2015). In this study, there are three groups that can be modeled as random effects to account for the dependencies they introduce: They are at the youth, instructional episode (which are indicators for the moments–or segments–in which youth are asked to respond to the ESM signal), and program levels. Pre-program interest, gender, and URM status are predictor variables at the youth level and the aspects of work with data are predictor variables at the instructional level; while there are not any predictor variables considered at the program level in the study, dependencies among the responses from youth within each program are accounted for in the random effect. To summarize, the three groups (youth, instructional episode, and program) and predictor variables present given the data collection and sampling strategy are modeled using random effects in a multi-level modeling approach.

As depicted in the Figure 2, the four research questions are formalized as follows:

1. What is the frequency and nature of opportunities for youth to engage in each of the five aspects of work with data in summer STEM programs?
2. What profiles of engagement emerge from data collected via ESM in the programs?
3. How do the five aspects of work with data relate to profiles of engagement?
4. How do youth characteristics relate to profiles of engagement?

Chapter 3

Method

3.1 Context

The setting for the present study is nine out-of-school STEM programs designed around best practices in urban areas in the Northeast United States during the summer of 2015. These are described in the Appendix with pseudonyms for the program names. Two intermediary organizations contracted by the urban area 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 provide training and professional development for the program's staff. A key difference between the intermediary organizations was that one separated academic and enrichment-related activities, whereas, in another, which was more closely involved in the day-to-day activities of the program, the academic and enrichment components were more integrated, which may have program-specific effects on youths' engagement. Many of the programs aim to involve youth in work with data. 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.

3.2 Participants

Participants consist 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: $M = 12.71$, $SD = 1.70$, $min. = 10.75$, $max. = 16.36$). Detailed demographic characteristics of youth are presented in the table.

3.3 Procedure

Youth completed a pre-survey before the program including questions about their experience in STEM, intention to pursue a STEM major or career, and questions for other motivation and engagement-related measures. At the beginning of the programs, youth were introduced to the study and the phones used for data collection related to the ESM. As indicated in the earlier section, ESM is a method of data collection that involves signalling youth to respond to short questions on phones that they were provided. Youth are signaled at random times (within intervals, so that the signals were not too near or far apart) in order to obtain a sample of youths' experiences 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. This detail is important because programs associated with

Table 3.1: Demographic characteristics of youth

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

one of the intermediaries rotated between classroom and field experience days, while the other used the first half of each day for one (i.e., classroom activities) or the other (i.e., field experience days).

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). All of the programs were video-recorded by research team members. So that measures corresponding to the video and ESM data can be matched, videos include a signal from the video-recorder identifying the ESM signal to which youth responded at that point in the video.

3.4 Data Sources and Measures

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

3.4.1 ESM measures of engagement for the profiles

Measures for engagement were constructed from three ESM responses that served as indicators for the experience of engagement and two ESM responses for the conditions of engagement. The three variables for engagement are for learning (for the cognitive engagement construct), working hard (for behavioral engagement), and enjoying (for affective engagement). The variables for the conditions are for perceived challenge and perceived competence. All five items are used to construct profiles. 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), as presented in Table 2. Note that because these three dimensions of engagement are measured using single-item indicators (which is common in studies using ESM; Hektner et al., 2007), information about the reliability and validity information for these measures is not included.

Table 3.2: ESM measures for profiles

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?

Table 3.3: Measure for pre-program interest in STEM

Construct	Items.text
Individual interest in STEM	I am interested in science / mathematics / engineering.
	At school, science / mathematics / engineering is fun
	I have always been fascinated by science / mathematics / engineering)

3.4.2 Survey measures of pre-interest in STEM

Measures of youths' pre-interest are used as youth-level influencers of the profiles. 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: Not at all (associated with the number 1 on the survey), A little (2), Somewhat (3), and Very Much (4). Reliability and validity information on this scale is presented in Vandell et al. (2008). The 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 (taking the maximum value) measurement approach. The items are presented in Table 3. Overall levels of this measure were high (M = 3.044 (SD = 0.901).

3.4.3 The five key aspects of work with data

Different aspects of work with data are 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. I then identified the specific activities that corresponded to the five aspects of work with data, as defined in Table 4. Details on the reliability of this measure is described next; more information on how the measure aligns with the original STEM-PQA on which this measure is based are presented in the Appendix.

In February, 2017, 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. Raters completed a four-hour online training module on the overall PQA tool and then attended an in-person two-day training led by a trainer from the David P. Weikart Center for Youth Program Quality, the tool's publisher, where they learned about the instrument, trained on its use, and then established inter-rater reliability with a master coder. 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 all 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.

Table 3.4: Coding Frame for Work With Data

Code	Description
Asking questions	Discussing and exploring topics to investigate and pose questions.
Making observations	Watching and noticing what is happening with respect to the phenomena or problem being investigated.
Generating data	Figuring out how or why to inscribe an observation as data and generating coding frames or measurement tools.
Data modeling	Understanding and explaining phenomena using models of the data that account for variability or uncertainty.
Interpreting and communicating findings	Discussing and sharing and presenting findings.

Note that these codes were unique to each signal to which youth responded (but were not unique to each youth, as youth in the same program were signaled at the same time). I discuss limitations to use of the STEM-PQA for work with data in the discussion.

3.4.4 Demographic variables used

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 are identified on the basis of youths' racial and ethnic group being Hispanic, African American, Asian or Pacific Islanders, or native American.

3.5 Data Analysis

The steps for both preliminary and the primary analyses are described in this section.

3.5.1 Preliminary analyses

First-order Pearson correlations and the frequency, range, mean, and standard deviations are first examined 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 examined. Note that 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 are used for all of the analyses involving measures for the aspects of work with data.

3.5.2 Analysis for Research Question #1 (on the frequency and nature of work with data)

There are two primary steps taken to answer this question, one more quantitative in nature and one more qualitative. Specifically, first, the frequency of the codes for the individual aspects of work with data from the STEM PQA measure of work with data (described above in the measures) are calculated. Note that this coding frame focused on the degree of instructional support the activity leaders provided for youth to work with data.

However, this framework provided no context on how particular types of work with data were enacted. Qualitative differences in *how*, for example, youth were asking questions are not evident from the STEM-PQA codes used to find the frequencies of the aspects of work with data. In order to provide more context in the description of how work with data in the context of summer STEM programs, all of the segments were coded using an open-ended, qualitative approach. Three research assistants were trained for approximately eight hours over four meetings. Then, each research assistant coded all of the segments associated with one of the videos. Two coders coded every segment, except for the 77 (out of the total 248) segments that the

STEM-PQA coding that indicated no aspects of work with data were present. For these 77 segments, only one coder coded each segment.

The coders used the following five guiding questions, associated with each of the five aspects of work with data, for the qualitative coding:

- When asking questions or defining problems is coded, what, if any are the questions or problems? Who is asking the question (i.e teacher or student)
- When making observations is coded, what are youth doing?
- When generating data is coded, how, if they are, are youth collecting or recording data?
- When analyzing or modeling data is coded, what analysis are they doing, or what models are they using? Are they talking about variability or uncertainty? If so, how?
- When interpreting and communicating findings is coded, what are youth interpreting or how are they communicating?

This coding took around 75 hours of coding by the research assistants. After coding all of the segments for each program, the coders and I met to discuss potential issues that emerged throughout the coding, and to clarify how they applied the coding frame (so the coders and I met nine times during the process to discuss the coding). I then read through all of the codes for all of the segments then made notes associated with each of the five aspects of work with data. I used these notes to write detailed descriptions of each of the aspects of work with data, which I grouped into the themes. I present these themes in the section of the results for this question.

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

To answer this question, Latent Profile Analysis (LPA; Harring & Hodis, 2016; Muthen, 2004) is used. LPA allows for capturing the multidimensional nature of engagement through profiles. 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 10 profiles were considered. As part of LPA, the model type selection–where the type refers to which parameters are estimated–is a key topic. For the present study, six model types were considered:

1. Varying means, equal variances, and covariances fixed to 0
2. Varying means, equal variances, and equal covariances
3. Varying means, varying variances, and covariances fixed to 0
4. Varying means, varying variances, and equal covariances
5. Varying means, equal variances, and varying covariances
6. Varying means, varying variances, and varying covariances

The MPlus software (Muthen & Muthen, 1998-2017) is used to carry out LPA as part of this study. In order to more flexibly carry out LPA, an open-source tool, tidyLPA (Rosenberg, Schmidt, & Beymer, 2018), was developed. This tool provides interfaces to both the MPlus software and to the open-source mclust software. In addition to being used as part of this study, this package is provided free of use to other analysts as the first tool dedicated to carrying out LPA as part of the R software. More details on the statistical software developed and included in the Appendix.

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, sample adjusted BIC [SABIC], consistent AIC [CAIC]), statistics about the quality of the profile assignments (entropy, which represents the mean posterior probability), statistical tests (Vu-Lo-Mendell-Rubin LRT [VLMR], Lo-Mendell-Rubin LRT [LMR], and the bootstrapped LRT [BLRT]), and concerns of interpretability and parsimony are used. As described in more detail in the section of the results pertaining to this question, on the basis of these criteria, the *model one type, six profiles* solution is selected and used as part of subsequent analyses.

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

Broadly, this question is focused on how work with data as coded from video-recordings of the programs, relates to the profiles. 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) is used. All of the models for this and the subsequent research question use random effects for youth, instructional episode, and program effects. Youth and instructional episode can be considered to be crossed with both nested within the program.

The probability of a response belonging to the profile is the dependent variable and the aspects of work with data are the independent variable. There are six models, for each of the six profiles. Because the outcome from LPA is not a hard classification (i.e., an observation is in a profile—or not) but a probability, the dependent variable is treated as a continuous variable.

First, null models with only the random parts (i.e., random youth, instructional episode, and program effects) are specified. Then, the five aspects of work with data are 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 associated with these five predictors. For example, if the coefficient for the effect of the asking questions aspect of work with data upon one of the profiles is 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 only with the aspects of work with data yielded very similar results (see the Appendix for more detail). The models with both together were also used as part of research question #4, though they are presented here (and interpreted in the sections for both results). In specific, mixed effects models, predicting the probability of membership in each of the six profiles as the dependent variable–using the work with data codes as predictors–were specified.

Because the results were found to be identical when the aspects of work with data and the youth characteristics are considered in separate and in the same model, the results from the two sets of variables being in the same model are used for both to provide answers to both this and the next research question. 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 yielded one (small) statistically significant result, the results for this analysis are not presented in the results.

3.5.5 Analysis for Research Question #4 (how youth characteristics relate to engagement)

Research question #4 is focused on how the relationships of work with data differ on the basis of youth characteristics–their pre-program interest, gender and URM status. Like for the previous research question, models that account for the cross-classification of the instructional episodes and the youth are used. The dependent variable is again the probability of a response being in the profile. The three youth characteristics (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) are added as predictors. Like for the previous research question, the statistical significance and the magnitude and direction of the coefficients associated with each predictor are interpreted to answer this question. For example, and similar to the interpretation of the predictors associated with RQ #3, if the relationship between pre-program interest and a profile is 0.05, then for each one-unit increase in pre-program interest, then youth are are five percentage points more likely to report a response in a particular profile.

Models with the youth characteristics separate from and together with the aspects of work with data were fit. Like for the results for the previous question, the models only with the youth characteristics yielded very similar results; see the Appendix for the results from the model with the youth characteristics included as predictors in separate (without the aspects of work with data) models. Thus, the models presented in the previous section with both youth characteristics and the aspects of work (see the table above) with data are interpreted here.

As described in the previous sub-section, because the results were very similar when the aspects of work with data and the youth characteristics were added in *separate* models compared to when they were included in the same model, the results for both sets of predictors in the same model are presented and interpreted. In addition, interactions between statistically significant aspects of work with data and all of the youth characteristics are examined, though because none of these interactions were found to be statistically significant, they are not included with the results.

3.6 Sensitivity Analysis

For observational studies, such as the present study, it can be important to determine how robust an inference is to alternative explanations. One approach to addressing this is sensitivity analysis, which involves quantifying the amount of bias that would be needed to invalidate an inference. Using the approach described in Frank, Maroulis, Duong, and Kelcey (2013), I carried out sensitivity analysis for inferences made relative to key findings. I used the R package konfound (Rosenberg, Xu, & Frank, 2018). The result, and what is used to interpret and contextualize findings, is a numeric value for each effect that indicates the proportion of the estimate that would have to be biased in order to invalidate the inference. I use these to interpret and contextualize the statistically significant findings. Higher values indicate more robust estimates in that the inferences would still hold even if there were substantial bias in the estimate and that are interpreted as robust findings, while lower values, when present, indicate less robust findings that I interpret with more caution.

Chapter 4

Results

In this section, results associated with the preliminary analysis and the four research questions are presented.

4.1 Preliminary results

4.1.1 Descriptive statistics for the engagement measures

First, descriptive statistics for the engagement measures (the five variables that were used to estimate the profiles) are presented in Table 5. These values suggest moderately high levels of the three dimensions of engagement (with mean values between 2.768 (SD = 1.063) for cognitive engagement, and 2.863 (SD = 1.044), for behavioral engagement, on one-four scales) and high perceptions of competence (M = 3.000 (SD = 0.952)); the only continuous variable that was not measured using ESM) and lower perceptions of challenge (M = 2.270 (SD = 1.117)).

4.1.2 Correlations among the study variables

Next, correlations between the variables that are used to create the profiles are presented in Table 6. These correlations among the variables used to construct the profiles, which range from r = .08 through r = .60 (all statistically significant), show moderate relations.

Table 4.1: Descriptive statistics for study variables

	n	Mean	SD
Cog. eng.	2969	2.768	1.063
Beh. eng.	2959	2.863	1.044
Aff. eng.	2970	2.831	1.051
Challenge	2970	2.270	1.117
Competence	2970	3.000	0.952

Table 4.2: Correlations among study variables

	Pre-interest	Cog. eng.	Beh. eng.	Aff. eng.	Challenge	Competence
Pre-interest						
Cog. eng.	.14					
Beh. eng.	.13	.60				
Aff. eng.	.12	.59	.57			
Challenge	.15	.30	.27	.27		
Competence	.06	.40	.41	.47	.08	

Table 4.3: Proportion of signals for which each of the aspects of work with data was present

Aspect of Work With Data	Proportion	N
Asking Questions	0.389	92
Making Observations	0.258	61
Generating Data	0.453	107
Data Modeling	0.288	68
Communicating Findings	0.470	111

4.2 Results for Research Question #1

4.2.1 Frequency of work with data

Of the 236 instructional episodes, 170 (72%) were coded as involving *any* of the aspects of work with data. Table 7 includes the frequency of the specific aspects of work with data, with interpreting and communicating findings being the most present (occurring in 47% of the coded instructional episodes), followed by generating data (in 45% of the instructional episodes), asking questions (in 39%), data modeling (29%), and then making observations (26%).

Note that these results are for codes applied to approximately ten-minute (video-recorded) instructional episodes and that the aspects of work with data could co-occur. On average, there were 1.86 ($SD = 1.61$) aspects of work with data present in each 10-minute instructional episode. This indicates that, on average, youth were engaged in around two of aspects of the work with data during each instructional episode. There was considerable variation in the extent to which these types of work with data were supported in each program (see the Appendix).

4.2.2 Nature of work with data

For these results, the different aspects of work with data were looked at in more detail using an open-ended, qualitative approach in order to better understand the nuance and the specific nature of what was going on during these episodes in terms of how students work with data. This coding, which showed there to be distinct, qualitative differences in the particular ways youth worked with data when, for example, they were data modeling, resulted in approximately three to four sentence notes from each of two raters for every instructional episode and showed the specific nature of work with data.

4.2.2.1 Asking questions or identifying problems

Among the instructional episodes that included the asking questions aspect of work with data (as determined through coding with the STEM-PQA measure), creating qualitative descriptions revealed that around one-third (36/92, or 39%) were focused on asking questions focused on youth working to understand the phenomenon or problem they were investigating. For example, in an instructional episode during the *Ecosphere* program in which youth constructed inclined tables to study how water moved throughout the ecosystem, the youth activity leader prompted youth to generate hypotheses of what would happen when water was poured onto the table, before pouring the water. Other instructional episodes involved questions that were not focused on predicting, conjecturing, or hypothesizing. In such cases, the code was applied to instances in which the youth were asking generic questions (i.e., about how they do an assignment) or when the instructor was asking youth questions (i.e., math-related questions). For example, in the *Marine Investigators* program, youth visited a water treatment site, and were provided opportunities to ask questions about what they observed.

4.2.2.2 Making observations

In the instructional episodes during which the STEM-PQA revealed that youth were making observations, the vast majority (49/57, 86%) of these were focused on observing phenomenon in the field, or, in the case of engineering-focused programs, noticing what was going on with a particular design. For example, in the *Building Mania* program, youth constructed Rube Goldberg machines; youth were prompted by the activity leaders to notice how changes in their design led to differences in how far objects were launched or rolled. In just a few cases, making observations was focused on making observations not of phenomena, but of the instructor. For example, in the *Adventures in Mathematics* program, instances in which youth observed other youth or the youth activity leader solving a mathematics problem was often coded as involving making observations.

4.2.2.3 Generating data

Around half (48/102, or 47%) of these episodes, for youth generating sources of data (as indicated by the STEM-PQA), the instructional episodes were focused around writing down observations made of a phenomenon, recording information from experiments, or recording the results of a trial (in engineering contexts). For example, in the *Marine Investigators* program, youth collected pieces of recyclable plastic, bringing them back to the classroom and counting them for each location they were collected. The other half of the cases were most often instructional episodes in which youth were writing down what the youth activity leader was saying or were focused on collecting specimens (but not writing them down) entering them into a spreadsheet, or otherwise recording them as data. For example, again in the *Marine Investigators* program, youth used nets to collect saltwater organisms, which they then transported in buckets back to the classroom setting for subsequent analysis. While these specimens could be considered as data, at least in the instructional episode described, youth did not inscribe notes or any other observations on the specimens they were collecting, and so data was not generated (at this stage).

4.2.2.4 Data modeling

A large majority (49/68, 72%) of the instructional episodes coded (with the STEM-PQA) for data modeling were focused on youths' uses of statistical and mathematical models. For example, in the *Comunidad de Aprendizaje* program, youth accessed nationally-representative data and were tasked to solve problems, like finding out what percentage of people engage in particular activities, like donating to charity. In a small number of instructional episodes, this aspect of work with data was present when the youth activity leader, rather than students, was doing the modeling, or the model was not one that could generate data. For example, in the *Marine Investigators* program, a youth activity leader used a plush toy seal designed to teach youth about anatomy and the dangers of aquatic mammals consuming trash and recyclables.

4.2.2.5 Interpreting and communicating findings

In around half (49/103, 48%) of the instructional episodes in which youth were interpreting and communicating findings (as coded by the STEM-PQA), youth were sharing what they found from an investigation or the results of using the product they designed. For example, in the *Comunidad de Aprendizaje* program, youth participated in an activity designed to support their thinking about creating a product to bring to market; the youth activity leaders described this as being akin to the television show the *Shark Tank*. In one instructional episode, the youth activity leader asks youth to think of an idea that would make an investor willing to invest in; students shared their ideas, describing what their ideas was, why it was a good idea, how much they could sell it for, and what their profit would be, while fielding questions from youth activity leaders and their peers. Interpreting and communicating findings was also commonly present in instructional episodes in which youth were debating the findings of an investigation, such as the results of calculations for the amount of recyclables entering waterways (in *Marine Investigators*). In the other half of the responses, youth were most commonly communicating about topics other than the results of an investigation or design process, such as trying to find out the answer to a question posed by the youth activity leader, or the youth activity leader was who was doing the interpreting and communicating. For example, in the *Adventures in Mathematics* program, the youth activity leader helped youth to solve problems on a worksheet, asking guiding questions to help youth start to solve problems on their own.

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

A relatively simple model (model *one* type; with varying means, equal variances, and covariances fixed to 0) with six profiles was selected for use in subsequent analyses. This model has profiles characterized by both varying levels on both the dimensions of engagement–cognitive, behavioral, and affective–and youths' perceptions of challenge and competence. In addition, the number of observations across the profiles is relatively balanced.

This means that *six distinct profiles* were identified in the data. This selection was on the basis of fit statistics, statistical tests, and concerns of interpretability and parsimony. The solution demonstrated superior fit on the basis of the information criteria (AIC and BIC) and on the basis of the measure of classification accuracy (entropy). Note that a seven profile solution with the same specifications regarding means, variances and covariances was also a similarly good fit (and is presented in the Appendix), but the 6 profile solution was ultimately chosen on the basis of parsimony and interpretability.

For the six profiles, presented below in Figure 3, the *first* plot shows the profiles with variables that were centered to have a mean equal to 0 and a standard deviation of 1. Thus, the y-axis for this plot is labeled "Z-score"). The *second* plot shows the profiles with the raw data (not transformed). Thus, the *y*-axis for this plot is labeled "Value."

This solution is characterized by:

- A *universally low* profile, characterized by low levels of engagement and challenge and competence
- An *only behaviorally engaged* profile, with moderate levels of behavioral engagement, very low affective engagement, and moderately (low) levels of cognitive engagement and challenge and competence
- An *only affectively engaged* profile, with moderate levels of affective engagement, low levels of behavioral engagement, and moderately (low) levels of cognitive engagement and challenge and competence
- A *all moderate* profile, with moderate levels of the three dimensions of engagement and challenge and competence
- An *engaged and competent but not challenged* profile, characterized by high levels of each of the three dimensions of engagement and of competence, but with low levels of challenge
- A *full* profile, with high levels of engagement, challenge, and competence

The number of observations associated with each of the profiles is somewhat balanced, with the universally low profile with the largest number of observations ($n = 667$), followed by the all moderate profile ($n = 638$). Each of the other four profiles were associated with 300 to 400 observations.

4.3.1 Sources of variability in profiles of engagement

The remaining analyses use the six profiles described above. Sources of variability in these profiles can be used as additional information in their own right for interpreting the profiles and in order to anticipate the effects of predictor variables at the youth, instructional episode, and program levels.

First, the proportion of the variability at each of these levels is explored through the use of null, or variance components, in Table 4.3. Again, these are models that only include grouping (i.e., the variable identifying which youth a response is from, what signal the response is associated with, and from which program the youth and signal were from) factors. These models provide insight into at which of these "levels" predictors may be able to explain the outcome. For all six profiles, the ICCs at the program level were very small, from 0.00 to 0.023. This suggests that very little variability can be explained simply by the program. For the instructional episode level, the ICCs were also very small, ranging from 0.004 to 0.011. Finally, the youth-level ICCs ranged from .099 to .427.

Table 4.4: Intra-class correlation (ICC) values for each of the three levels

Instructional Episode	Youth	Program
0.006	0.093	0.009
0.006	0.267	0.023
0.015	0.310	0.000
0.009	0.100	0.000
0.004	0.262	0.003
0.031	0.432	0.019

Looking across these values, most of the explained variability in the responses is associated with youth; the program and instructional episode levels were associated with very small values, suggesting that variables at these levels have minimal variability to explain. In turn, this suggests that these variables, including those for work with data, may not have strong effects in terms of their relations with the profiles.

In terms of specific ICCs at the youth level, the value for the youth-level ICC was highest for the *Full* profile ($ICC = .432$), suggesting that some youth have a strong tendency to be fully engaged (possibly due to their initial interest or other individual characteristics and differences). The other profile characterized by a consistent pattern across all of the variables–the *Universally low* profile–had a modest value ($ICC = .265$). Finally, a large amount of variability is associated with the residual (variance that is not associated with the program, instructional episode, or youth levels). This suggests that there is wide variation in students' responses that may not be readily explained or predicted by variables *at one level alone*. Remaining unexplained variability is the product of youth, instructional episode, and program effects together. Some youth from particular programs may engage during some episode instructional episodes in very high or low ways that are not captured by modeling the variability at each of these levels alone.

Variability in terms of the number (and proportion) of profiles each youth reports can also be considered. When the breakdown of responses in each of the six profiles by youth is explored: As presented in Figure X, the value of the mean for the proportion of responses for each youth in the profile they reported most was .540 ($SD = .194$, $min = .182$, $max = 1.00$). This indicates that on average youths' most reported profile comprises just over one-half of their responses, with substantial variability in their responses. Apart from this overall mean proportion, the mean proportion for specific profiles can be considered. For example, when *Full* engagement was reported by a youth more than any other profile, they reported it, on average, in just over 60% of their responses. No other profile that youth reported most was associated (on average) with a larger proportion of their responses. This suggests that even when youth report (relatively) stable engagement over the course of their time in the programs, they still engage in a variety of different ways. When youth reported *All moderate* engagement (the profile with the lowest proportion of responses, on average, when reported more than any other), they reported it in just less than 40% of their responses. This suggests that though youth may report one profile far more than others–and that youth who report particular (i.e., *Full*) profiles more than others may reflect somewhat stable engagement, there still exists substantial variability in youths' engagement.

4.4 Results for Research Question #3: Aspects of work with data and engagement

The results for this and the next research question are presented in Table 4.5. In this table, each column represent the output from one of the six different models. For example, the first column includes the results for the model with the probability of a response being associated with the *Only behavioral* profile as the dependent variable. The cells down the rows contain the coefficients (and their standard errors and (p-values)) for each of the predictor variables.

Only the coefficients associated with the aspects of work with data are interpreted in this section in order to provide results for this research question, because the youth characteristics are interpreted for research question #4. As described earlier, the results were practically the same for these sets of predictors being included in either separate or the same model, so they are included in the same model.

See the Appendix for the results from the model with the aspects of work with data included as predictors in separate (without the youth characteristics) models. Note that there were only significant relations with the *Full* profile (see the column with the column name *Full* for these results). The only relations that were statistically significant were for the relations between modeling data and the *Full* profile ($\beta = 0.034$ (0.017), $p = .020$) and between generating data and the *Full* profile ($\beta = 0.027$ (0.015), $p = .033$): When youth were either modeling or generating data, they were more likely to be fully engaged.

After finding there to be few relations between the aspects of work with data and engagement, sensitivity

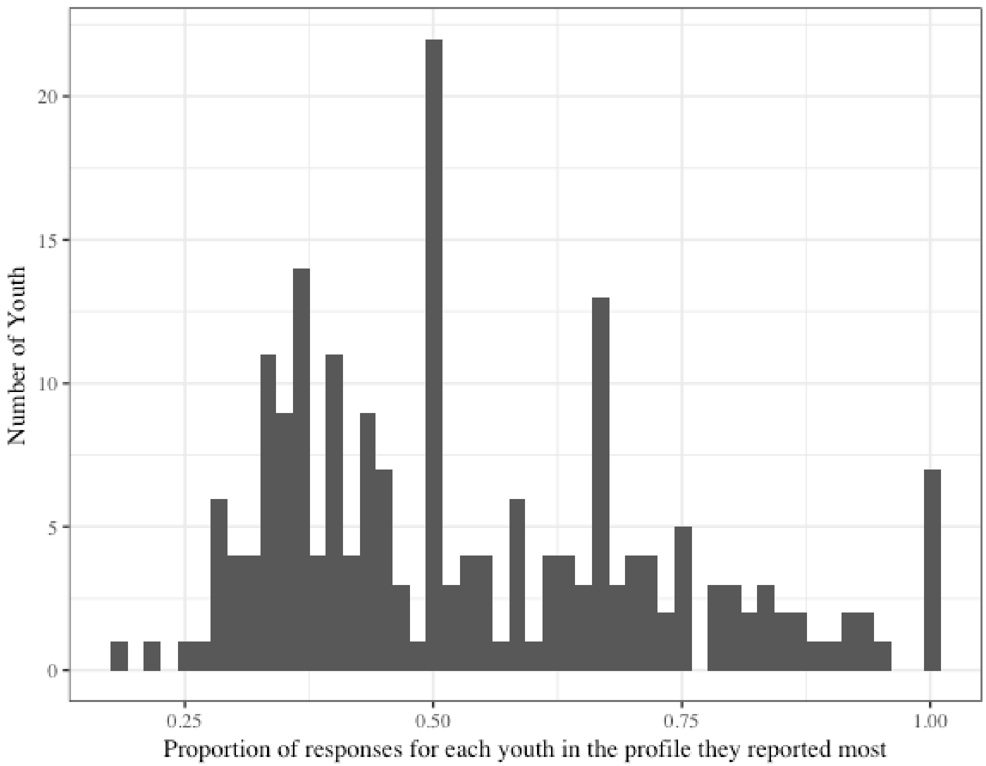


Figure 4.1: Histogram of the proportion of responses for each youth in the profile they reported most

analysis for the statistical significant effects was carried out. This revealed that the effect of modeling data on *Full* engagement was more robust than that for generating data (also upon *Full* engagement): 9.835% of the effect of modeling would have to be due to bias to invalidate the inference about its effect, whereas only 1.884% of the effect of generating data would need to be due to bias to invalidate the inference about its effect. Further explanations and investigations of these effects are the focus on research question #4 (in terms of the effect of youth characteristics) and are discussed in the next chapter.

Table 4.5: Results of mixed effects models with the interactions between interest and other characactistics and the composite for work with data

Profile	Universally low	Only behavioral	Engaged and competent but not challenged	All moderate	Only affective	Full
Intercept	0.356 (0.086) (p < .001)	0.107 (0.045) (p = 0.01)	0.022 (0.063) (p = 0.362)	0.354 (0.075) (p < .001)	0.09 (0.04) (p = 0.015)	0.094 (0.083) (p = 0.132)
Pre-interest	-0.047 (0.022) (p = 0.982)	-0.013 (0.012) (p = 0.873)	0.039 (0.016) (p = 0.009)	-0.012 (0.019) (p = 0.743)	0.007 (0.01) (p = 0.264)	0.018 (0.021) (p = 0.195)
Gender-Female	0.06 (0.037) (p = 0.051)	0.019 (0.019) (p = 0.159)	0.025 (0.028) (p = 0.188)	-0.038 (0.033) (p = 0.875)	-0.02 (0.018) (p = 0.867)	-0.035 (0.037) (p = 0.827)
URM status	-0.01 (0.052) (p = 0.575)	0.031 (0.026) (p = 0.122)	-0.012 (0.04) (p = 0.614)	-0.076 (0.046) (p = 0.95)	0.018 (0.025) (p = 0.234)	0.043 (0.053) (p = 0.207)
Asking	-0.015 (0.018) (p = 0.789)	0.015 (0.015) (p = 0.158)	-0.011 (0.015) (p = 0.763)	0.023 (0.017) (p = 0.09)	0.004 (0.014) (p = 0.397)	-0.019 (0.016) (p = 0.887)
Observing	0.003 (0.018) (p = 0.427)	0.013 (0.015) (p = 0.191)	0.009 (0.015) (p = 0.266)	0.007 (0.017) (p = 0.342)	-0.017 (0.014) (p = 0.886)	-0.025 (0.016) (p = 0.94)
Generating	-0.014 (0.017) (p = 0.789)	0.014 (0.014) (p = 0.17)	-0.014 (0.014) (p = 0.833)	0.012 (0.016) (p = 0.233)	-0.02 (0.013) (p = 0.938)	0.027 (0.015) (p = 0.033)
Modeling	0.004 (0.019) (p = 0.407)	-0.023 (0.016) (p = 0.929)	0 (0.015) (p = 0.504)	-0.004 (0.018) (p = 0.593)	-0.012 (0.015) (p = 0.8)	0.034 (0.017) (p = 0.02)
Communicating	0.002 (0.018) (p = 0.461)	0.018 (0.015) (p = 0.115)	0.004 (0.015) (p = 0.404)	-0.011 (0.017) (p = 0.747)	0.016 (0.014) (p = 0.124)	-0.027 (0.016) (p = 0.956)

4.5 Results for Research Question #4: Youth characteristics and engagement

Like for the results for research question #3, each column is associated with the results for a single model, again in Table 4.5. For example, the first row is again associated with the results for the model predicting the probability of the *Only behavioral* profile, with the cells across the columns containing the coefficients, their standard errors, and their p-values. These results show that overall pre-interest is associated with the *Engaged and competent but not challenged* profile (β = 0.039 (0.016), p = .009). For this effect, 17.879% would be needed to invalidate the inference, suggesting a moderately robust effect. The effect of being a female was not statistically significant but has a relation of 0.060 (0.037, p = .051) upon the probability of a response being associated with the *Universally low* profile. For the effect of gender upon the *Universally low* profile, 17.843% of the bias would need to be removed (or the effect would need to be larger by this percentage) to sustain the inference.

These few, small findings were more surprising than the similarly minimal relations observed for work with data: as the null models indicate, there were large ICCs (a large proportion of the variability in the outcome variables) at the youth-level (as pre-interest, gender, and URM status were variables associated with this level). However it appears that the youth level variables of interest to this study were not effective at explaining much of this variability. This is discussed further in the next chapter.

Chapter 5

Discussion

5.1 Key Findings

5.1.1 Key findings for research question #1 (on the frequency and nature of work with data)

In terms of the frequency and nature of work with data, work with data was found to be common in the summer STEM programs that made the context for this study. A coding frame synthesized from past research carried out in STEM domains on work with data was used to find that work with work with data occurred from around one-quarter of the time of the program's time (making observations) to around one-half of the program's time (generating data and communicating findings). Data modeling was, like making observations, less common, whereas asking questions and generating data, like communicating findings, were relatively more common. These findings are as expected based on past research (Lee & Wilkerson, in press) and given the design and goals of summer STEM programs (Dabney et al., 2012; Elam et al., 2012), including those participating in the present study.

In-depth qualitative showed that asking questions, generating data, and interpreting and communicating findings, the three aspects that were *more frequent* in the programs, also were more consistent with general aspects of learner-centered, hands-on STEM activities, and not as consistent with work with data in particular. This suggests that while work with data is somewhat common, qualitative analysis is an important part of understanding youths' engagement in work with data. More veridical forms of it are somewhat less common, occurring in around 25% of the programs' time. While descriptive in nature, these results present the first insight that I am aware of of the extent of work with data in STEM enrichment programs. They suggest that, as past scholarship (National Research Council, 2009, 2012) can provide a context for youth to be involved in the type of scientific and engineering practices-focused activities that can be particularly powerful for youth (and students) in terms of their learning.

5.1.2 Key findings for research question #2 (what profiles of engagement emerge)

Six profiles of engagement were identified. These were selected using a rigorous model selection approach and through use of a sophisticated modeling approach (LPA) and statistically software developed for this analysis (tidyLPA). These profiles included those that were strongly negative (*Universally low*) and strongly positive (*Full*), as well as those characterized by different levels of engagement (*Only behavioral* and *Only affective*) and by different levels of the conditions of engagement (Engaged and Competent but not Challenged). An *All moderate* profile was also identified. The profiles suggest that the experiences of youth in summer STEM programs are variable and that the use of ESM can aid in the study of youths' engagement. Little research has examined profiles of engagement, though Schmidt et al. (2018) examined profiles of engagement, constructed from items for cognitive, behavioral, and affective engagement (but not perceptions of challenge and competence), and 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.

5.1.3 Key findings for research question #3 (how work with data relates to engagement)

Before relations between the groups of "predictor" variables, work with data and youth characteristics, and the profiles, were explored, the amount of variability that could be explained at the program, youth, and instructional episode levels were explored; use of cross-classified mixed effects models were particularly helpful for this goal. The amount of variability that could be explained at the program and instructional episode level was small (no larger for any profile than .023, and as low as .00 at the program level and .004 at the instructional episode level for some profiles), while the amount of variability that could be explained at the youth level was moderate to large (between .099 and .427). This suggests that while there is variability in the composition of the profiles that were identified, youth characteristics–their pre-program thoughts, beliefs, and characteristics and their inclination to engage in particular ways throughout the program–largely explains the prevalence of the profiles. This also suggests that what youth do during the programs, and the design and implementation of the programs themselves, have little to do with how youth engage in them. This implies that even the strongest predictor variables at these (instructional episode and program) levels would likely not explain much variability in the profiles (though this is not always the case, as there are cases in which adding variables at one level can increase the amount of variability that can be explained at another; Gelman & Hill, 2007).

In line with what the preliminary analysis of the amount of variability that could be explained at the youth, instructional episode, and program levels, relations between work with data were largely not found, though some small, statistically significant relations were identified. Importantly, both generating and modeling data were found to be positively related to the *Full* profile, suggesting that when youth are involved in these practices, then they are more likely to report high levels of cognitive, behavioral, and affective engagement, and high perceptions of competence and challenge. The effect of data modeling was more robust than that for generating data, the latter which should be interpreted with caution. In short, this suggests that these activities are beneficial to youths' engagement. Both communicating and interpreting findings and the composite measure for work with data were positively related to the *Only behavioral* profile and these findings were fairly robust. This profile may indicate that students are experiencing a routine engagement (and not particularly adaptive) when they are communicating findings and being involved in work with data in general. As there is no research on how work with data relates to youths' engagement, the findings associated with this research question provide some, albeit limited, evidence (and directions for future research) for how some aspects of work with data relate to youths' engagement.

5.1.4 Key findings for research question #4 (how youth characteristics relate to engagement)

Not as much in line with expectations given the preliminary analysis, relations between youth characteristics and the profiles were found to be small. In this way, these small relations were similar (in magnitude) to those between work with data and the profiles. Youth with higher pre-program interest were more likely to be *Engaged and competent but not challenged*, suggesting that youth with higher interest in STEM are inclined to be highly engaged and good at what they are doing, but are not challenged by the activities they experience. This could be a function of the relationship between youths' interest and their competence before the program, which are often strongly related ([add]); these youth, as a result of their higher interest and competence, need more challenging activities to be more fully engaged. This effect was fairly robust. The interaction with gender and the work with data composite revealed a positive relationship with *Full* engagement, suggesting that the more that female youth work with data, the more likely they are to be positively engaged. However, sensitivity analysis revealed that this effect was not very robust, which, along with its small magnitude, suggests that it should be interpreted with some caution. Finding that female youth who are engaged in work with data are more likely to be fully engaged is important, given that past research has suggested that female students are less likely to be engaged in STEM classes but we have limited information about what types of instruction may best support female students to be engaged and successful (e.g., Patall et al., 2017).

5.2 Limitations of the Study and Recommendations for Future Research

This study examines youths' engagement as an outcome. Accordingly, outcomes from engaging, such as the products of neither youths' work or the specific cognitive capabilities they develop through their participation, are not the focus. Thus, while some findings about how work with data and youth characteristics were found to be associated with different profiles of engagement, we do not have an understanding of how engaging in more or less adaptive ways relates to these outcomes. Examining how work with data and engagement relate to key learning, motivational, and future goals and plans-related outcomes is a topic for future research.

Another limitation concerns the context of the study, summer STEM programs. While the programs that were involved in the study have many affordances for work with data and for being highly engaging for youth, they have some limitations, too, particularly with respect to support work with data. Importantly, these were not programs explicitly designed to support work with data; while such contexts are being developed, they are not yet widespread. Learning environments that deliberately support work with data over a long period may demonstrate different patterns of engagement than those examined in this study because of the focus on and sequencing of the aspects of work with data, which may make it more (or less) cognitively, behaviorally, or affectively engaging than is determined in this study. As Miller, Manz, Russ, Stroupe, and Berland (2018) highlight, truly engaging STEM activities are not easily come by; they require students to take ownership over and to make decisions about their explorations or designs. Thus, future research may study work with data in contexts designed to support it. A key part of this future research may be studying both work with data and how work with data is supported (most importantly by the instructor but also by the curriculum and technological tools).

A related limitation is that the programs that were the focus of this study were model programs, or those based on characteristics of exemplary STEM enrichment programs. As a result, engagement may be different in other STEM enrichment programs depending on characteristics of the programs and their activities, and findings from this study should be interpreted in terms of programs that share similar characteristics.

A potential issue concerns the analytic approach. As noted above, the profiles demonstrated very little variability at the program and instructional episode level, suggesting that factors at this level would likely not strongly predict the profiles. This could be a function of the use of profiles and the specific variables selected. It may also be the result of the outcome (engagement and its conditions) selected. Other analytic approaches can be carried out to determine the viability of the profiles approach and use of the items for engagement and its conditions for understanding work with data.

A final limitation concerns the measures used. In particular, the qualitative coding revealed alignment but also discrepancies between work with data as determined from the PQA codes and the conceptual framework for work with data. While these issues were small, they suggest that the coding frame for work with data is a limitation of the present study.

While these are important limitations, it is worth noting that the modeling strategy (with the mixed effects models) in inherently a conservative approach. Thus, while the findings detected are small, they can be considered to be trustworthy on the basis of the way the ESM data were analyzed. This trustworthiness is enhanced by the use of sensitivity analysis, which showed how much of the effects could be due to bias for them to be invalidated.

5.3 Implications for Practice

5.3.1 Engage youth in key aspects of work with data

While limited evidence, this study suggests that generating and modeling data in particular may be beneficial in terms of engaging youth. Generating data in particular may be a key practice because it involves making work with data concrete; as Lehrer and Schauble (2015) describe, recording data in the form of "inscriptions" can serve as commitments that youth make (in terms of what data were chosen to be collected and recorded). This implication, in particular, should be interpreted with caution, however, given the very small magnitude of the effect. Similarly, data modeling has been described as *the* central scientific and engineering practice (Schwarz et al., 2009; 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.

Practically, 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 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). Recent curricular reform efforts also suggest that the best way to engage learners in particular practices is through the process 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 then communicated (or even used in an argument; National Governors Association, 2013; National Research Council, 2012; NGSS Lead States, 2013). With respect to work with data in particular, youth activity leaders and teachers 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 their day-to-day lives. Many questions or problems learners face may involve data that can be meaningfully incorporated into engaging learning activities.

5.3.2 Leverage the affordances of summer STEM and other STEM enrichment programs

Another implication for practice concerns the affordances (and constraints) of summer STEM and other STEM enrichment programs. One affordance of these programs relevant to these informal and to K-12 learning environments concerns selecting activities that are engaging to youth. For example, in the *Marine Investigators*, youth participated in activities designed to help them understand water quality in their ecosystem. Youth collected trash from sites around their community (in different "districts") and then brought the trash and recyclable plastic back to the classroom. Then, the youth activity leaders asked students to figure out how much plastic enters local waterways. As a part of this activity, youth activity leaders asked students not only to determine the quantity of trash that entered the waterways, but asked students about *why* they used math in particular ways (i.e., adding the quantity of trash collected and then extrapolating from this quantity to the amount from across the entire city over the course of the year). This appeared to be a powerful activity, one that was coded as involving all five aspects of work with data according to the measures for instructional support for work with data; this type of activity seemed to suggest that instructional support for work with data may impact youth's engagement.

Another affordance concerned the relevance of the program to youth's lives. For example, in the *Building Mania* program, youth are involved in engineering design (i.e., identifying a problem and designing a solution), particularly around the use of simple machines. In a day in the classroom setting, youth are creating, testing, and revising catapults. In the next day, youth visit an area University, and are led in a discussion by a physicist who works with particle colliders. In this example, the expertise of the physicist, who explicitly mentions the benefits of engaging in the engineering design process and the importance of combining engineering to addressing problems (such as mitigating the damage of earthquakes), seems to be highly relevant to what youth are doing in their class. In these two days of class, youth are engaged in different aspects of work with data as indicated by the codes for instructional support for work with data (collecting data on the efficacy of their designs in the classroom day, and asking questions in the subsequent day, particularly); these seem to suggest, like the example of work work with data from the *Marine Investigators* program, affordances of work with data for summer STEM programs.

5.3.3 Consider the constraints of summer STEM and other STEM enrichment programs

There are also constraints to summer STEM and other STEM enrichment programs. For example, youth activity leaders faced challenges linking activities as part of a complete cycle of investigation. For example, in the *Ecosphere* program, youth collected water samples in the field. They then brought these samples to the classroom and tested the water, involving students in both collecting and, to a degree, generating data (by noting the pH levels of the water). However, later in the day, youth created a small-scale model (with inclined trays of dirt, rocks, and plants) of an ecosystem, in which they added food coloring to determine the impacts of chemicals and acid rain. Youth then interpreted and discussed these findings, but did not connect the discussion to the water samples youth collected and tested earlier. This activity presented an opportunity for deeper engagement, in which youth could interpret and communicate findings related to the state of the water in their ecosystem, but, instead, it was potentially limiting in terms of youth's engagement in work with data.

Another constraint related to the challenge of linking activities concerned what the programs focused on. For example, the mathematics-focused programs, such as the *Adventures in Mathematics* program, the youth activity leaders recognizing that youth had difficulty solving equations, used duct tape and a "hippity hoppity", building on an earlier activity in which youth considered what constituted a rate, on how many "hops" it would take someone to move from one end of the line of duct tape to the other; the youth activity leader than asked youth to consider how far they could move in one hop and to consider how they could find out many hops it would take, using a mathematical equation. In this activity, youth were supported to approach mathematics problem-solving in creative ways. However, apart from data modeling, other aspects of work with data were rarely present, and most of the data that youth worked with was provided by the teacher or considered in the abstract. Programs focused on science or engineering, similarly, emphasized other aspects of work with data: The science-focused programs (*Island Explorers*, *The Ecosphere*, and *Marine Investigators*) all emphasized collecting and generating data, but data, particularly the data collected or generated, was rarely modeled or interpreted. In the engineering-focused programs (*Uptown Architecture*, *Crazy Machines*, and *Dorchester House*, youth often collected data that resulted from their engineering designs, and communicated and interpreted their findings, but, did not generate data, and, accordingly, (and like the science-focused programs) did not model data as a regular part of their activities. This finding suggests that while work with data may have been common overall, different aspects of instructional support for work with data were emphasized to different degrees based on the focus of the program.

5.4 Conclusion

Each of the disciplines that contribute to STEM learning involve work with data and how youth and students work with data in engaging ways is a concern of researchers and practitioners. While past research has focused on what aspects of work with data learners are involved in with respect to work with data, or specific conceptual outcomes from working with data, little research has considered youths' experience of working with data. In this study, engagement was used as a lens to understand the experience of youth working with data in the context of nine 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 identified from video-recordings of the programs. These codes were then used to predict profiles, or distinct groups on the basis of different levels, of youths' cognitive, behavioral, and affective engagement, and two other variables, youths' perceptions of challenge and competence. These measures were obtained using an innovative method, ESM, that provides some access to youths' experience in-the-moment of the activities they were involved in during the program.

Findings indicate that work with data occurs regularly in the programs and that there are some examples of ambitious activities centered on working with real-world data (and examples in which the work with data is not fully aligned with youth-driven work with data). Six profiles of engagement were identified, representing different configurations of the three dimensions of engagement and its conditions. Relations 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, one characterized by high levels of all five of the variables used to create the profiles. Female youth who were involved in work with data (at the instructional episode level) to a greater extent were also more likely to be fully engaged. This study suggests that work with data has purchase as an organizing set of activities for STEM can have some benefits in terms of understanding the nature of what youth do in summer STEM programs. In addition, this study shows that ESM and engagement can be used to understand youths' experiences. Data–and who is able to work with data–have important roles in STEM learning and in society; efforts to understand and support learners engaging in these ambitious activities should be encouraged and expanded.

Chapter 6

References

Akiva, T. (2005). Turning training into results: The new youth program quality assessment. High/Scope Resource, 24(2), 21-24.

Bergman, L. R., & Magnusson, D. (1997). A person-oriented approach in research on developmental psychopathology. Development and psychopathology, 9(2), 291-319.

Bergman, L. R., Magnusson, D., & El Khouri, B. M. (2003). Studying individual development in an interindividual context: A person-oriented approach. Psychology Press.

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.

Breckenridge, J. N. (2000). Validating cluster analysis: Consistent replication and symmetry. Multivariate Behavioral Research, 35(2), 261-285.

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

Cohen, J. (1992). A power primer. Psychological Bulletin, 112(1), 155.

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.

Corpus, J. H., & Wormington, S. V. (2014). Profiles of intrinsic and extrinsic motivations in elementary school: A longitudinal analysis. The Journal of Experimental Education, 82(4), 480-501.

Csikszentmihalyi, M. (1990). Flow: The psychology of optimal performance. Cambridge, England: Cambridge University Press.

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

Creswell, J. W., Plano Clark, V. L., Gutmann, M. L., & Hanson, W. E. (2003). Advanced mixed methods research designs. In A. Tashakkori & C. Teddlie (Eds.), Handbook of mixed methods in social and behavioral research (pp. 209–240). Thousand Oaks, CA: Sage.

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., & McColskey, W. (2012). The measurement of student engagement: A comparative analysis of various methods and student self-report instruments. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), The handbook of research on student engagement (pp. 763–782). New York: Springer Science. https://doi.org/10.1007/978-1-4614-2018-7_37

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.

Fredricks, J. A., Filsecker, M., & Lawson, M. A. (2016). Student engagement, context, and adjustment: Addressing definitional, measurement, and methodological issues. Learning & Instruction, 43, 1-4.

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.

Gopnik, A., Sobel, D. M., Schulz, L. E., & Glymour, C. (2001). Causal learning mechanisms in very young children: two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. Developmental Psychology, 37(5), 620.

Greene, B. A. (2015). Measuring cognitive engagement with self-report scales: Reflections from over 20 years of research. Educational Psychologist, 50(1), 14-30.

Greene, K. M., Lee, B., Constance, N., & Hynes, K. (2013). Examining youth and program predictors of engagement in out-of-school time programs. Journal of Youth and Adolescence, 42(10), 1557-1572.

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.

Hayenga, A. O., & Corpus, J. H. (2010). Profiles of intrinsic and extrinsic motivations: A person-centered approach to motivation and achievement in middle school. Motivation and Emotion, 34(4), 371-383.

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

Jahnukainen, M. (2010). Extreme cases. Encyclopedia of Case Study Research. Thousand Oaks, CA: 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.

Lauer, P. A., Akiba, M., Wilkerson, S. B., Apthorp, H. S., Snow, D., & Martin-Glenn, M. L. (2006). Out-of-school-time programs: A meta-analysis of effects for at-risk students. Review of educational research, 76(2), 275-313.

Lee, H. S., Angotti, R. L., & Tarr, J. E. (2010). Making comparisons between observed data and expected outcomes: students' informal hypothesis testing with probability simulation tools. Statistics Education Research Journal, 9(1), 68-96.

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.

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.

Linnansaari, J., Viljaranta, J., Lavonen, J., Schneider, B., & Salmela-Aro, K. (2015). Finnish Students Engagement in Science Lessons. NorDiNa: Nordic Studies in Science Education, 11(2), 192-206. Retrieved from https://www.journals.uio.no/index.php/nordina/article/view/2047

Lovett, M. C., & Shah, P. (2007). Preface. In M. C. Lovett & P. Shah (Eds.), Thinking with data (pp. x-xx [requested book through ILL to confirm page #s]). New York, NY: Lawrence Erlbaum.

Magnusson, D., & Cairns, R. B. (1996). Developmental science: Toward a unified framework. Cambridge, England: Cambridge University Press.

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 Guide. Los Angeles, CA: Muthén & Muthén.

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

Nolen, S. B., Horn, I. S., & Ward, C. J. (2015). Situating motivation. Educational Psychologist, 50(3), 234-247. Patall, E. A., Vasquez, A. C., Steingut, R. R., Trimble, S. S., & Pituch, K. A. (2016). Daily interest, engagement, and autonomy support in the high school science classroom. Contemporary Educational Psychology, 46, 180-194.

Patall, E. A., Steingut, R. R., Vasquez, A. C., Trimble, S. S., Pituch, K. A., & Freeman, J. L. (2017). Daily Autonomy Supporting or Thwarting and Students' Motivation and Engagement in the High School Science Classroom. Journal of Educational Psychology. Advance online publication. http://dx.doi.org/10.1037/edu0000214

Pekrun, R., & Linnenbrink-Garcia, L. (2012). Academic emotions and student engagement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), Handbook of research on student engagement (pp. 259-292). New York, NY: Springer.

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.

Piaget, J., & Inhelder, B. (1969). The psychology of the child. New York, NY: Basic Books.

Pöysä, S., Vasalampi, K., Muotka, J., Lerkkanen, M. K., Poikkeus, A. M., & Nurmi, J. E. (2017). Variation in situation-specific engagement among lower secondary school students. Learning and Instruction. http://dx.doi.org/10.1016/j.learninstruc.2017.07.007

Rosenberg, J. M. (2018). Comparing mplus and mclust output. Retrieved from https://jrosen48.github.io/r-markdown/comparing-mplus-mclust.html

Salmela-Aro, K., Moeller, J., Schneider, B., Spicer, J., & Lavonen, J. (2016). Integrating the light and dark sides of student engagement using person-oriented and situation-specific approaches. Learning and Instruction, 43, 61-70.

Salmela-Aro, K., Muotka, J., Alho, K., Hakkarainen, K., & Lonka, K. (2016). School burnout and engagement profiles among digital natives in Finland: A person-oriented approach. European Journal of Developmental Psychology, 13(6), 704-718.

Schneider, B., Krajcik, J., Lavonen, J., Salmela-Aro, K., Broda, M., Spicer, J., … & Viljaranta, J. (2016). Investigating optimal learning moments in US and Finnish science classes. Journal of Research in Science Teaching, 53(3), 400-421.

Schmidt, J. A., Rosenberg, J. M., Beymer, P. (advance online publication). A person-in-context approach to student engagement in science: Examining learning activities and choice. Journal of Research in Science Teaching. https://dx.doi.org/10.1002/tea.21409

Schwarz, N., Kahneman, D., & Xu, J. (2009). Global and episodic reports of hedonic experience. In R. Belli, D. Alwen, & F. Stafford (Eds.), Using calendar and diary methods in life events research (pp. 157-174). Newbury Park, CA: Sage.

Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. Educational Researcher, 27(2), 4-13.

Shernoff, D. J., Csikszentmihalyi, M., Schneider, B., & Shernoff, E. S. (2003). Student engagement in high school classrooms from the perspective of flow theory. School Psychology Quarterly, 18(2), 158-176.

Shernoff, D. J., Kelly, S., Tonks, S. M., Anderson, B., Cavanagh, R. F., Sinha, S., & Abdi, B. (2016). Student engagement as a function of environmental complexity in high school classrooms. Learning and Instruction, 43, 52-60.

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

Singh, K., Granville, M., & Dika, S. (2002). Mathematics and science achievement: Effects of motivation, interest, and academic engagement. The Journal of Educational Research, 95(6), 323-332.

Shernoff, D. J., & Schmidt, J. A. (2008). Further Evidence of an Engagement–Achievement Paradox Among U.S. High School Students. Journal of Youth and Adolescence, 37(5), 564–580. http://doi.org/10.1007/s10964-007-9241-z

Shumow, L., Schmidt, J. A., & Zaleski, D. J. (2013). Multiple perspectives on student learning, engagement, and motivation in high school biology labs. The High School Journal, 96(3), 232-252.

Skinner, E. A., & Pitzer, J. (2012). Developmental dynamics of engagement, coping, and everyday resilience. In S. Christenson, A. Reschly, & C. Wylie (Eds.), Handbook of Research on Student Engagement (pp. 21-45). New York: Springer Science.

Skinner, E. A., Kindermann, T. A., & Furrer, C. J. (2009). A motivational perspective on engagement and disaffection: Conceptualization and assessment of children's behavioral and emotional participation in academic activities in the classroom. Educational and Psychological Measurement, 69(3), 493-525.

Skinner, E., Furrer, C., Marchand, G., & Kindermann, T. (2008). Engagement and disaffection in the classroom: Part of a larger motivational dynamic? Journal of Educational Psychology, 100(4), 765.

Smith, C., Akiva, T., Sugar, S., Lo, Y. J., Frank, K. A., Peck, S. C., Cortina, K. S., & Devaney, T. (2012).Continuous quality improvement in afterschool settings: Impact findings from the Youth Program Quality Intervention study. Washington, DC: The Forum for Youth Investment.

Steinley, D., & Brusco, M. J. (2011). Evaluating mixture modeling for clustering: recommendations and cautions. Psychological Methods, 16(1), 63.

Stohl, H., & Tarr, J. E. (2002). Developing notions of inference using probability simulation tools. The Journal of Mathematical Behavior, 21(3), 319-337.

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.

Strati, A. D., Schmidt, J. A., & Maier, K. S. (2017). Perceived challenge, teacher support, and teacher obstruction as predictors of student engagement. Journal of Educational Psychology, 109(1), 131-147.

Trevors, G. J., Kendeou, P., Bråten, I., & Braasch, J. L. (2017). Adolescents' epistemic profiles in the service of knowledge revision. Contemporary Educational Psychology, 49, 107-120.

Turner, J. C., & Meyer, D. K. (2000). Studying and understanding the instructional contexts of classrooms: Using our past to forge our future. Educational Psychologist, 35(2), 69-85.

van Rooij, E. C., Jansen, E. P., & van de Grift, W. J. (2017). Secondary school students' engagement profiles and their relationship with academic adjustment and achievement in university. Learning and Individual Differences, 54, 9-19.

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

Wang, M. T., & Eccles, J. S. (2012). Social support matters: Longitudinal effects of social support on three dimensions of school engagement from middle to high school. Child Development, 83(3), 877-895.

Wang, M. T., & Holcombe, R. (2010). Adolescents' perceptions of school environment, engagement, and academic achievement in middle school. American Educational Research Journal, 47(3), 633-662.

Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. Journal of Experimental Psychology: General, 143(5), 2020-2045.

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

Wilkerson, M. H., Andrews, C., Shaban, Y., Laina, V., & Gravel, B. E. (2016). What's the technology for? Teacher attention and pedagogical goals in a modeling-focused professional development workshop. Journal of Science Teacher Education, 27(1), 11-33.

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.

Witherington, D. C. (2015). Dynamic systems in developmental science. In W. F. Overton & P. C. M. Molenaar (Vol. Eds.) & R. M. Lerner (Ed.), Handbook of child psychology and developmental science. Vol. 1: Theory & method (7th ed., pp. 63-112). Hoboken, NJ: Wiley.

Wormington, S. V., & Linnenbrink-Garcia, L. (advance online publication). A new look at multiple goal pursuit: The promise of a person-centered approach. Educational Psychology Review. doi:10.1007/s10648-016-9358-2

Chapter 7

Appendix

7.1 Appendix A: STEM-PQA alignment

Table 7.1: Alignment of codes for instructional support for work with data and the STEM-PQA

Work.With.Data	Description	STE
Asking questions or defining problems	Discussing and exploring topics to investigate and pose questions.	Pred
Making observations	Watching and noticing what is happening with respect to the phenomena or problem being investigated.	Class
Generating data	Figuring out how or why to inscribe an observation as data and generating coding frames or measurement tools.	Colle
Data modeling	Understanding and explaining phenomena using models of the data that account for variability or uncertainty.	Simu
Interpreting and communicating findings	Discussing and sharing and presenting findings.	Anal

7.1.1 Appendix B: Program descriptions

Table 7.2: Program (with pseudonyms) descriptions

Program.Name	Program.Description
Island Explorers	A science-focused program that aims to help youth develop expertise on one species found in the local ecosystem by reading and writ
The Ecosphere	A science-focused program that aims to help youth to explore the marine life of Narragansett Bay. Efforts were undertaken to build y
Zoology Partners	A science-focused program that aims to support youth's development of content knowledge related to the issue of endangered species
Marine Investigators	A science-focused program that aims to provide youth with opportunities to learn about and experience Narragansett Bay; examine l
Comunidad de Aprendizaje	A STEM-focused program that aims to help youth improve basic skills in mathematics and develop an interest in STEM content and
Jefferson House	A STEM-focused program that aims to support youth's development of basic math skills, the program was primarily focused on help
Uptown Architecture	An engineering-focused program that aims to support youth's participation in a process to design and build an outdoor learning spac
Building Mania	An engineering-focused program that aims to provide youth with the opportunity to experiment with designing and using simple mac
Adventures in Mathematics	A mathematics-focused program that aims to help youth to develop the basic math skills and prevent summer learning loss among pa

Table 7.3: Proportion of signals for which each of the aspects of work with data was present

Aspect of Work With Data	Proportion	N
Asking Questions	0.389	92
Making Observations	0.258	61
Generating Data	0.453	107
Data Modeling	0.288	68
Communicating Findings	0.470	111

7.1.2 Appendix C: Work with data by program

Table 7.4: Proportion of signals for which each of the aspects of work with data was present by program

Variable	Asking	Observing	Generating	Modeling	Communicating	Total Segments
Island Explorers	0.312	0.375	0.438	0.250	0.375	16
The Ecosphere	0.625	0.417	0.500	0.292	0.500	24
Zoology Partners	0.250	0.167	0.125	0.167	0.208	24
Marine Investigators	0.458	0.333	0.250	0.375	0.542	24
Comunidad de Aprendizaje	0.327	0.182	0.400	0.273	0.327	55
Jefferson House	0.167	0.083	0.542	0.458	0.750	24
Uptown Architecture	0.375	0.208	0.708	0.167	0.292	24
Building Mania	0.333	0.208	0.375	0.333	0.500	24
Adventures in Mathematics	0.583	0.292	0.542	0.458	0.750	24

7.1.3 Appendix D: Model specifications details

Here, the six models that are possible to specify in LPA are described in terms of how the variables used to create the profiles are estimated. Note that p represents different profiles and each parameterization is represented by a 4 x 4 covariance matrix and therefore would represent the parameterization for a four-profile solution. In all of the models, the means are estimated freely in the different profiles. Imagine that each row and column represents a different variable, i.e., the first row (and column) represents broad interest, the second enjoyment, the third self-efficacy, and the fourth another variable, i.e., future goals and plans. Models 1 and 3 meet the assumption of independence, that is, that, after accounting for their relations with the profile, the variables used to estimate the profiles are independent (Collins & Lanza, 2010). They estimate variable variances but do not estimate covariances (i.e., as can be seen, the covariance matrices are "diagonal," without any off-diagonal parameters that are estimated). These models are estimated by default in MPlus, although these assumptions can be relaxed (Muthen & Muthen, 2017). Importantly, this does not mean the variables used to create the profile are assumed to be not related; as Collins and Lanza (2010) explain:

> The local independence assumption refers only to conditioning on the latent variable. It does not imply that in a data set that is to be analyzed, the observed variables are independent. In fact, it is the relations among the observed variables that are explained by the latent classes. An observed data set is a mixture of all the latent classes. Independence is assumed to hold only within each latent class, which is why it is called "local".

Despite the assumption of independence, as Collins and Lanza (2010), Muthen and Muthen (2017), and others (i.e., Pastor et al., 2007; Vermunt & Magidson, 2002) note, it can be lifted to improve model fit, though these models without the assumption of independence may be better described as general or Gaussian mixture models (Fraley et al., 2017).

7.1.3.1 Varying means, equal variances, and covariances fixed to 0 (model 1)

In this model, which corresponds to the mclust model wit the name "EEI", the variances are estimated to be equal across profiles, indicated by the absence of a p subscript for any of the diagonal elements of the matrix. The covariances are constrained to be zero, as indicated by the 0's between every combination of the variables. Thus, this model is highly constrained but also parsimonious: the profiles are estimated in such a way that the variables' variances are identical for each of the profiles, and the relationships between the variables are not estimated. In this way, less degrees of freedom are taken used to explain the observations that make up the data. However, estimating more parameters–as in the other models–may better explain the data, justifying the addition in complexity that their addition involves (and their reduction in degrees of freedom).

$$\begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{bmatrix}$$

7.1.3.2 Varying means, equal variances, and equal covariances (model 2)

This model corresponds to the mclust model "EEE". In this model, the variances are still constrained to be the same across the profiles, although now the covariances are estimated (but like the variances, are constrained to be the same across profiles). Thus, this model is the first to estimate the covariance (or correlations) of the variables used to create the profiles, thus adding more information that can be used to better understand the characteristics of the profiles (and, potentially, better explain the data).

$$\begin{bmatrix} \sigma_1^2 & \sigma_{21} & \sigma_{31} & \sigma_{41} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{13} & \sigma_{12} & \sigma_3^2 & \sigma_{33} \\ \sigma_{14} & \sigma_{12} & \sigma_{12} & \sigma_4^2 \end{bmatrix}$$

7.1.3.3 Varying means, varying variances, and covariances fixed to 0 (model 3)

This model corresponds to the mclust model "VVI" and allows for the variances to be freely estimated across profiles. The covariances are constrained to zero. Thus, it is more flexible (and less parsimonious) than model 1, but in terms of the covariances, is more constrained than model 2.

$$\begin{bmatrix} \sigma_{1p}^2 & 0 & 0 & 0 \\ 0 & \sigma_{2p}^2 & 0 & 0 \\ 0 & 0 & \sigma_{3p}^2 & 0 \\ 0 & 0 & 0 & \sigma_{4p}^2 \end{bmatrix}$$

7.1.3.4 Varying means, varying variances, and equal covariances (model 4)

This model, which specifies for the variances to be freely estimated across the profiles and for the covariances to be estimated to be equal across profiles, extends model 3. Unfortunately, this model cannot be specified with mclust, though it can be with MPlus; this model *can* be used with the functions to interface to MPlus described below.

$$\begin{bmatrix} \sigma_{1p}^2 & \sigma_{21} & \sigma_{31} & \sigma_{41} \\ \sigma_{12} & \sigma_{2p}^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{13} & \sigma_{12} & \sigma_{3p}^2 & \sigma_{33} \\ \sigma_{14} & \sigma_{12} & \sigma_{12} & \sigma_{4p}^2 \end{bmatrix}$$

7.1.3.5 Varying means, equal variances, and varying covariances (model 5)

This model specifies the variances to be equal across the profiles, but allows the covariances to be freely estimated across the profiles. Like model 4, this model cannot be specified with mclust, though it can be with MPlus. Again, this model *can* be used with the functions to interface to MPlus described below.

$$\begin{bmatrix} \sigma_1^2 & \sigma_{21p} & \sigma_{31p} & \sigma_{41p} \\ \sigma_{12p} & \sigma_2^2 & \sigma_{23p} & \sigma_{24p} \\ \sigma_{13p} & \sigma_{12p} & \sigma_3^2 & \sigma_{33p} \\ \sigma_{14p} & \sigma_{12p} & \sigma_{12p} & \sigma_4^2 \end{bmatrix}$$

7.1.3.6 Varying means, varying variances, and varying covariances (model 6)

This model corresponds to the mclust model "VVV". It allows the variances and the covariances to be freely estimated across profiles. Thus, it is the most complex model, with the potential to allow for understanding many aspects of the variables that are used to estimate the profiles and how they are related. However, it is less parsimonious than all of the other models, and the added parameters should be considered in light of how preferred this model is relative to those with more simple specifications.

$$\begin{bmatrix} \sigma_{1p}^2 & \sigma_{21p} & \sigma_{31p} & \sigma_{41p} \\ \sigma_{12p} & \sigma_{2p}^2 & \sigma_{23p} & \sigma_{24p} \\ \sigma_{13p} & \sigma_{12p} & \sigma_{3p}^2 & \sigma_{33p} \\ \sigma_{14p} & \sigma_{12p} & \sigma_{12p} & \sigma_{4p}^2 \end{bmatrix}$$

7.1.4 Appendix E: Additional details on the model selection process

Looking across the statistics presented, some general ideas about which models are to be preferred emerge. Solutions are interpreted first for each model individually and then across models with the goal of choosing a smaller number of models to investigate in more detail.

Table 7.5: Solutions for models that converged with replicated LL

Number of Profiles	LL	AIC	BIC	SABIC	CAIC	Entropy	VLMR	LMR	BLRT
Model 1									
2	-19894.14	-19894.14	39916.16	39865.32	39820.47	0.807	3468.199 (0)	3397.353 (0)	3468.199 (0)
3	-19453.38	-19453.38	39082.59	39012.69	38951.11	0.794	881.519 (0.0126)	863.512 (0.0136)	881.519 (0)
4	-19196.33	-19196.33	38616.44	38527.47	38449.21	0.811	514.107 (0)	503.605 (0)	514.107 (0)
5	-18817.93	-18817.93	37907.60	37799.57	37704.68	0.913	756.788 (0)	741.329 (0)	756.788 (0)
6	-18648.78	-18648.78	37617.26	37490.17	37378.70	0.888	338.296 (0)	331.386 (0)	338.296 (0)
7	-18407.23	-18407.23	37182.11	37035.95	36907.95	0.886	523.141 (0.0112)	512.455 (0.0121)	523.141 (0)
9	-18186.35	-18186.35	36836.25	36651.96	36491.06	0.899	171.674 (0.1322)	168.167 (0.1359)	171.674 (0)
Model 2									
2	-19107.73	-19107.73	38423.27	38340.65	38267.95	0.924	850.304 (0)	832.934 (0)	850.304 (0)
3	-18897.06	-18897.06	38049.88	37948.20	37858.85	0.880	421.343 (0)	412.736 (0)	421.343 (0)
4	-18659.68	-18659.68	37623.06	37502.32	37396.37	0.922	474.773 (0)	465.075 (0)	474.773 (0)
5	-18474.83	-18474.83	37301.33	37161.52	37039.03	0.901	304.938 (0)	298.709 (0)	304.938 (0)

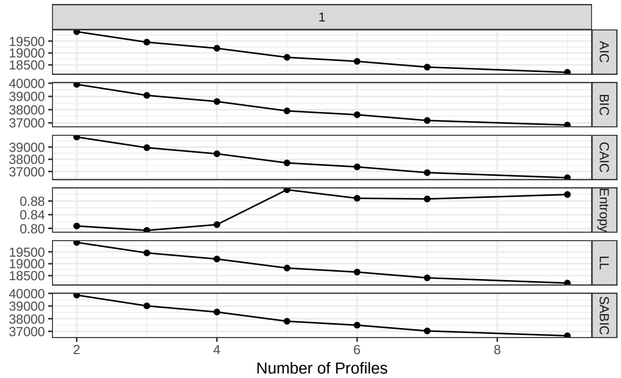


Figure 7.1: Fit statistics for model 1 solutions

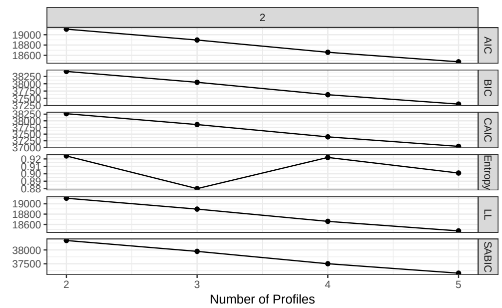


Figure 7.2: Fit statistics for model 2 solutions

For solutions associated with model 1, the decrease (indicating a preferred model) in information criteria becomes smaller as the number of profiles increases from 5 to 6 and 6 to 7. A solution associated with 8 profiles did not replicate the log-likelihood and the VLMR and LMR suggest that the solution associated with 9 profiles did not fit better than that with 8 profiles, suggesting that models with 7 or fewer profiles be preferred. Considering these models, the entropy statistic increases by a large amount between the solution associated with 4 and 5 profiles (and then decreases slightly between 5 and 6 and 6 and 7 profile solutions), suggesting (but not providing conclusive evidence) that models 5, 6, or 7 may be preferred. The bootstrapped LRT suggests that, until the log-likelihood is not replicated, every more complex model be selected. Taking these pieces of evidence into conclusion, for model 1, solutions associated with 4 through 7 may be considered in more depth, with an emphasis on solutions associated with profiles with 5 and 6 profiles on the basis of the slowing of the decrease in the information criteria associated with the solutions with greater profiles than these, and the increase in the entropy from 4 to 5 (and 6) profile solutions.

For solutions associated with model 2, only those associated with 2-5 profile solutions were associated with log-likelihoods that were replicated. For these four models, the log-likelihood decreased in a mostly consistent way, such that changes in the decrease are not as evident as those associated with model 1. The entropy statistic decreases from 2 to 3 profile solutions, increases from 3 to 4 profile solutions, and then decreases slightly from 4 to 5 profile solutions, providing some information that models associated with 4 profiles be preferred to the others. All of the LRTs suggest that the more complex model be selected, not providing clear information about which solutions are to be preferred. On the basis of these pieces of evidence, models with 3, 4, and 5 solutions may be considered in more depth. However, there is a lack of consistent evidence favoring more or less complex models.

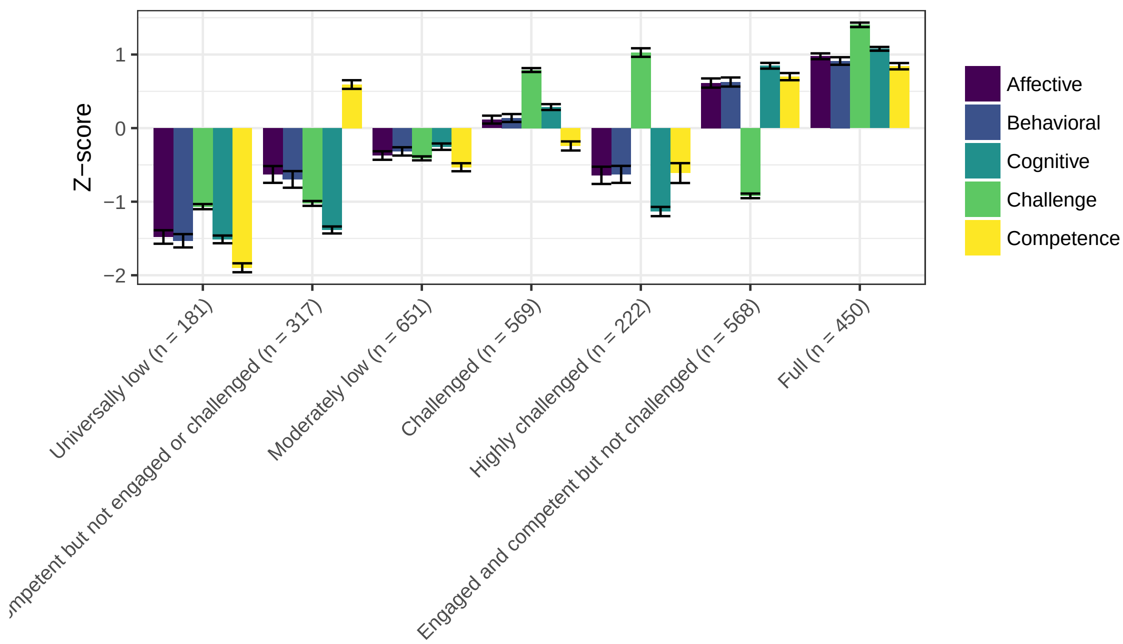
The model 1, six and seven profile solutions are compelling because both show profiles that are distinguished by dimensions of engagement and its conditions (challenge and competence). Note that for this model, only the means and variances are estimated (and so no covariances are estimated), and the variances are constrained to be the same across the profiles. While this is a very restrictive model, it, along with the model 3 type (which did not lead to solutions for any of the numbers of profiles specified) also is a standard model for LPA, in that it meets the assumption of local independence (of the variables that make up the profiles–unlike for models in which covariances are estimated) typical common to LPA (see Muthen & Muthen, 2016). While some of the solutions associated with the model 2 type did reach solutions, these demonstrated less appealing properties in terms of their fit statistics as well as their interpretability and with respect to concerns of parsimony. Thus, while no covariances are estimated for the model 1 type solutions, there is no requirement that these be specified; their benefit, when models associated with them are preferred, is that they can provide better fit: they can be used to better explain or predict the data in a sample, but their inclusion also means that over-fitting the model to the data can become a greater concern.

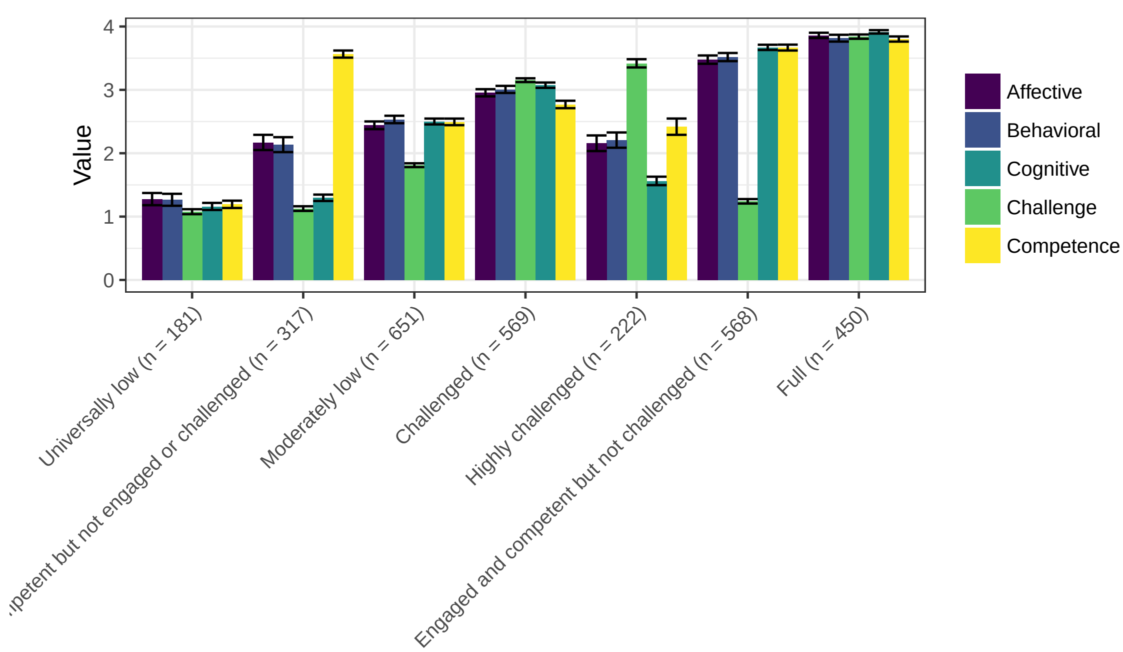
For each solution, alternate solutions associated with higher log-likelihoods were explored. One advantage of the six profile solution is that most of its profiles can also be identified in solutions with fewer profiles. For the six profile solutions, this alternate solution was very different, whereas for the seven profile solutions, this alternate solution was highly similar. The model solutions exhibit a less clear pattern in terms of which profiles appear when. All else being equal, on the basis of parsimony, the model 1, six profile solution is preferred and was selected for use in subsequent analyses.

7.1.5 Appendix F: Alternate model selected (model type 1, seven profile solution)

This solution is characterized by:

- A *full* profile, profile 7
- A *universally low* profile, profile 1
- A *competent but not engaged or challenged* profile, profile 2, characterized by high competence and moderate (low) or low levels of engagement and challenge
- A *moderately low* profile, profile 3, characterized by moderately low levels of all of the variables
- A *challenged* profile, profile 4, characterized by high challenge, moderate (high) levels of engagement, and moderate (low) levels of competence
- A *highly challenged* profile, profile 5, characterized by patterns similar to those of the challenged profile, but with higher challenge and with low levels of both engagement and challenge
- A *challenged but not engaged or competent* profile, profile 6, characterized by low levels of challenge, and high levels of engagement and competence





The number of observations associated with each of the profiles is not very balanced, with few ($n = 181$) observations associated with the universally low profile and few ($n = 222$) observations associated with the highly challenged profile. The number of observations associated with the other profiles ranged from 317 to 651. Distinct from other solutions, none of the other five profiles were found in the other model 1 solutions. Two pairs of the profiles–challenged and highly challenged and universally low and moderately low–exhibited similar patterns among the variables that were distinguished by different mean levels. The log-likelihood was replicated twice, with the next lowest log-likelihood being replicate four times, possibly warranting further investigation. Taken together, this solution raises questions about whether it may be too complex, possibly suggesting preference for model 1 five and six profile solutions.