How engagement during out-of-school time STEM programs promotes the development of

youths' interest in STEM domains

Joshua M. Rosenberg, Patrick N. Beymer, and Jennifer A. Schmidt

Michigan State University

Recently, out-of-school-time programs focusing on science, technology, engineering, and mathematics (STEM) have proliferated in order to combat declines in STEM interest during adolescence (Brophy, 2008; National Academy of Engineering and National Research Council, 2014; OECD, 2016) and to meet the demands of a rapidly growing STEM workforce (Fayer et al., 2017). Though many have argued that contexts for learning outside of the school setting have an important role to play in youths' development of interest (Bell, Lewenstein, Shouse & Feder, 2009; Hidi & Renninger, 2006), relatively little is known about whether and how youths' interest develops in such contexts.

Contemporary motivational theory suggests that interests emerge from the interactions of an *individual* in a particular *environment*, rather than residing completely within the individual or the environment (Hidi et al., 2004; Prenzel, 1992). Thus it is important to understand the ways that individuals engage with STEM-focused environments in order to understand how STEM-related interests may emerge. The purpose of this paper, then, is to examine whether and how youths' sustained engagement in summer STEM programming promotes the development of interest in STEM domains. To explore this issue, we employ unique methods of data collection and analysis that allow us to examine the effect of sustained multidimensional engagement in STEM program on the development of youth's interest over time.

Out-of-school Time as a Context for Interest Development in STEM

Out-of-school time is an overarching term used to refer to summer programs as well as after-school programs where youth participate in voluntary learning (Lauer et al., 2006). These programs have been touted for their ability to provide youth with mentors as well as enable youth to further develop their identity (Hirsch et al., 2010). Recently, there has been an influx of out-of-school time programs focused on STEM domains: Such programs are developed with the

purpose of increasing youth interest in STEM careers (Dabney et al., 2012; Elam et al., 2012). An emerging body of research provides some evidence that such programs can be effective in achieving this aim. One study showed that youth who attend out-of-school time STEM programs have a higher likelihood of choosing STEM career paths (Dabney et al., 2012). Some have suggested that out of school time learning environments might be especially effective at developing interest in STEM (or other) domains because these programs are free from typical school constraints such as set curricula, and therefore are able to focus their time on engaging activities (Renninger, 2007). However, little is known about how one's level of actual engagement in informal learning settings such as summer STEM programs impacts the development of interest. It is critical to understand whether and how interest can develop in these informal programs and whether the degree of one's engagement in these programs may contribute to this development.

Interest Development

Hidi and Renninger (2006) describe individual interest as "a relatively enduring disposition to re-engage particular contents over time" (p. 111). They, and other contemporary theorists frame interest as the product of the interaction between a person and particular content, and as such, interest is content specific (Hidi et al., 2004; Krapp, 2000; Prentzel, 1992; Renninger & Wozniak, 1985). A person's enduring interest in STEM content (denoted as *individual interest* by Hidi and Renninger (2006)), for example, is the result of the way that person has interacted with STEM ideas, practices and content over time. These interactions will be shaped by features of the environments in which such interactions take place, by other individuals who facilitate these interactions, and by a person's own efforts, such as their self regulation and the way they choose to engage with particular content (Renninger, 2000;

Renninger & Hidi, 2002; Sansone & Smith, 2000). A more enduring individual interest can

developed as a result of the repeated experience of *situational* interest -- a more fleeting interest

in a specific activity or task that is typically characterized by immediate positive affective

reactions to the immediate activity at hand, but may or may not be sustained over time.

Facilitating the development of youths' enduring individual interest in STEM fields has

been an explicit goal of many of the summer STEM programs that have proliferated throughout

the united states in recent years, including the nine programs that are the focus of this study.

These programs are typically designed to support interest development by presenting STEM

content in novel ways (i.e., by appealing to situational interest), by involving local experts in the

field in programming, and by allowing opportunities for deeper and more focused interaction

with STEM content without many of the constraints, distractions, and pressures of traditional

school environments. As such, summer STEM programs often have many supports built in for

interest development. Youth will benefit differently from these supports though, depending on

how they choose to engage with the STEM content within the confines of these programs. Youth

who engage with STEM content in different ways may differently support the development of

their own interest in the content. This study is aimed at exploring the degree to which youth can

support their own interest development through their engagement in summer STEM programs.

Interest development is presumed to involve cognitive, affective, and behavioral

processes, such that a more enduring personal interest emerges as a result of repeated interaction

with content that is affectively positive (as is the case with situational interest), that contributes

to the accumulation of stored knowledge and value, and that requires personal investment of

effort (Hidi & Renninger, 2006). Thus in exploring the role of youth engagement as a facilitator

of interest in summer STEM programs, we focus on these multiple dimensions of youth

engagement (more on this when we discuss engagement below).

Of course, no one enters summer STEM programs in an "interest vacuum." Indeed, many

youth register for summer STEM programs precisely because they already have an interest in

STEM that they wish to develop further: As such, summer STEM programs may support interest

development through a recursive processes. Alternatively, youth may become involved in

summer STEM programs for a variety of other reasons (i.e., because a parent or caretaker

decided they should do it; Beymer, Rosenberg, & Schmidt, 2018), in which case individual

interest may not be a driving force behind one's participation in a summer program. The point is,

youth enter summer programs with a certain amount of individual interest in the content

(whether high or low), which itself is the result of prior opportunities for interaction with said

content. To understand how youth's level of engagement in summer STEM programs supports

the development of their individual STEM interest, it is necessary to account for their individual

level of interest at program entry. Regardless of their level of STEM interest at program entry,

youth who are more deeply *engaged* during their time in the program may recursively impact the

level of their individual STEM interest (Hidi & Renninger, 2006). As specified in theoretical

accounts of the development of individual interest, situational interest plays an important role in

their momentary interactions. Therefore, in order to focus on *changes* in individual interest over

the period of the program we include youths' initial individual interest. To begin to account for

the recursive nature of interest development we also attempt to account for youth's experienced

situational interest during their daily engagement in program activities.

The interest of any individual youth in STEM may also be influenced by societal norms

and stereotypes about what types of people are interested in and qualified for particular domains.

Despite the parity that exists in males' and females' coursetaking and achievement in STEM areas, there continue to be persistent stereotypes about many STEM fields as being "male" fields, and of girls and women being less capable than boys and men in STEM domains (see Hill, Corbett and St. Rose, 2010 for a review). These persistent stereotypes and the implicit and explicit biases that emerge from them, can influence the way boys and girls interact with STEM content, and their trajectories of interest development over time. Given existing gender-related stereotypes about STEM disciplines, it is important to take gender into account in our examination. We also note that a similar argument might be made with respect to youth race and ethnicity. Our analyses do not control for this particular youth characteristic because nearly all of our study participants are black and/or Hispanic and would be considered underrepresented minorities in STEM fields.

Measuring and Understanding the Impacts of Engagement

Contemporary frameworks for engagement highlight the *multiple dimensions of* of engagement and its *dynamic* nature due its malleability that can change based on individual characteristics and pre-dispositions, as well as environmental factors (Shernoff & Schmidt, 2008; Skinner & Pitzer, 2013). Most commonly, scholars refer to cognitive, affective, and behavioral components (Christenson et al., 2012; Fredricks et al., 2004) -- dimensions that correspond to the processes proposed to be involved in interest development (see Hidi & Renninger, 2006). *Behavioral engagement* refers to one's involvement in activities in terms of their effort and participation. *Cognitive engagement* refers to one's mental investment in his or her own learning. *Affective engagement* refers to one's positive and negative feelings one has toward learning activities (Fredricks et al., 2004, 2011; Sinatra et al., 2015).

Engagement has been linked to a number of critical outcomes including persistence, achievement, and interest (Sinatra et al., 2015). Importantly, engagement has been shown to vary over time and context (Shernoff & Schmidt, 2008). For example, certain activities may lead to higher or lower levels of engagement, which in turn may impact the development of interest. Considering the wide variety of activities that characterize both formal and informal learning settings, it is important to measure engagement repeatedly over time and consider engagement *in situ*, using data collection methods (e.g., experience sampling) that allow for examining repeated measures over time. Examining engagement as situational affords the opportunity to ultimately identify factors influencing engagement that are under the control of educators (Schmidt et al., 2018). t

Scholars have developed methods to assess engagement in a way that makes it possible to understand its changes and dynamics over time and across activities. Among the approaches for measuring engagement in such a way, intensive longitudinal methods that use Experience Sampling Method and diary studies have shown their utility (Bolger & Laurencau, 2013; Hektner, Schmidt, & Csikszentmihali, 2007). The present study adds to this body of literature investigating how sustained engagement over a period of time in summer STEM programs may have cumulative effects on individual outcomes (interest) as measured before and after the programs.

The Present Study

This study explores how interest develops as a result of youth momentary engagement measured using the experience sampling method (Hektner et al., 2007). The research question driving this study is, *How is youth's sustained engagement across three weeks of summer STEM programming related to changes in their STEM interest over the duration of the program?* The

relationships examined in this paper are represented in Figure 1. To address this research

question,we use multiple data collection methods to examine the development of interest over

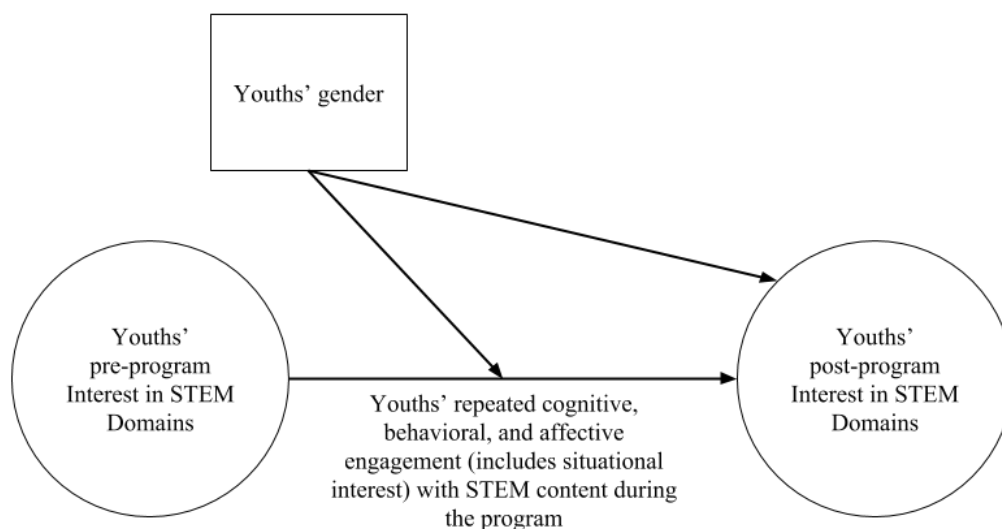time via youth engagement across a number of summer STEM programs.

Youths' gender

Youths'
pre-program
Interest in STEM
Domains

Youths' repeated cognitive,
behavioral, and affective
engagement (includes situational
interest) with STEM content during
the program

Youths'
post-program
Interest in STEM
Domains

Figure 1. Conceptual framework for how engagement during summer STEM programs impacts

youths' development of interest in STEM domains.

Method

Context

The present study was part of a larger study focused on the engagement and interest of

youth while participating in one of nine summer STEM programs in the Northeast United States.

These nine summer programs are part of two larger organizations that provide opportunities to

low-income youth to participate in summer programs where cost is not a barrier due to the support from external and internal funding. While attending the programs, adolescents participate in activities designed to create interest and engagement in STEM. These programs are designed so that adolescents spend about half of their time participating in activities in the classroom and half of their time participating in field experiences. These field experiences take place with community partners where youth engage in activities including designing computer games, exploring the ecology of islands, and planting community gardens. Each program lasted between four and six weeks. Throughout those weeks, adolescents participated Monday through Thursday for three hours each day.

Participants

Participants consist of 203 youth (50% female). The average age of participants was 12.71 (SD = 1.18) and ranged from 10 to 16 and the demographic makeup was 6% White, 36% Black, 48% Hispanic, 7% Asian, and 3% multiracial (as presented in Table 1).

Table 1. Demographic characteristics of youth.

Youth ($N = 203$)	% of Youth
Sex	
Male	50%
Female	50%
Race/Ethnicity	
Hispanic	48%
White	6%

Black	36%
Multi-racial	3%
Asian/Pacific Islander	7%

Age

10	4%
11	28%
12	31%
13	21%
14	12%
15	3%
16	1%

Parent Education (N = 171)

High School or Below	79%
Graduated from College (B.A. or B.S.)	21%

Procedure

Prior to the beginning of the program, youth completed a self-report survey for collecting demographic information and their interest in STEM. Over the course of the each program, data were collected on their overall engagement using the experience sampling method (Hektner et al., 2007). Data were collected using the experience sampling method for three weeks of the

program during two days each week. Data collection took place so that field experiences and classroom experiences were equally represented. The average attendance rate across programs was 83.08% (SD = 0.16). Youth completed a second self-report survey at the end of the program to assess their post-program interest in STEM.

   Experience sampling method. The experience sampling method is an intensive longitudinal method of data collection that is used to assess real-time experiences in responded to randomly emitted signals (see Hektner et al., 2007 for a description of this method). Adolescents were signalled four times per day throughout six days of the program. The only condition for signalling was that each signal must occur at least 15 minutes apart. Adolescents responded to each signal via mobile phones that they were given at the beginning of the day. Overall, 2,968 experience sampling responses were collected ($M = 14.6$) and the completion rate was 63%. Just over half of all missed signals were due to youth absence from the program.

Measures

   Engagement. Indicators of youth engagement were collected using the experience sampling method. A composite measure was formed consisting of four items (alpha = 0.85). Consistent with theoretical models of how sustained individual interest develops, our measure of engagement includes an indicator of momentary situational interest, which captures the extent to which particular program activities engaged youth. The four items asked youth to indicate on a four-point Likert scale (1 = *not at all*, 4 = *very much*) their agreement with the following statements:

- *Working hard*: As you were signaled, how hard were you working?

- *Concentrating*: As you were signaled, how well were you concentrating?

- *Enjoying*: As you were signaled, did you enjoy what you are doing?

● *Interest*: Was the main activity interesting?

Individual Interest. At the outset and conclusion of each program, a 4-point Likert Scale (1 = *not at all*, 4 = *very much*) was used to assess youth's individual interest in the STEM area or areas that were the focus of the program they enrolled in. For each relevant area, youth responded to three questions about their interest:

● I have always been fascinated by science/math/building

● I am interested in Science/math/building

● At school, science/math/building things is fun

The individual interest measure represented the mean of interest items across all relevant domains. Thus for some students, the mean was based on 3 items, while for others it was based on as many as 9 items representing all three domains (with Cronbach alpha values ranging from .77 - .86 for each domain specific interest scale).

Data Analysis

The goal of the data analysis was to understand how youths' in-the-moment engagement, relates to their post-program STEM interest, accounting for their pre-program STEM interest and gender. In order to carry out this analysis, we conduct a two-step modeling approach. The first step uses mixed effects (or multi-level) models using the lme4 (Bates, Machler, Bolker, & Walker, 2015) package in R (R Core Team, 2018) to estimate the youth-specific mean levels of engagement across *all* of their ESM responses, accounting for the effects of their pre-program interest and gender (Model 1A). Then, we use the predicted values for engagement as variables in the second model to predict youths' post-program STEM interest, controlling for their pre-program interest and with gender as a predictor (Model 1B). These predicted values are called

Best Linear Unbiased Predictors (BLUPs), and represent a compromise between a) simply calculating the mean for each student and b) calculating the overall level of engagement for all students, weighting how much systematic information is available for each youth to estimate a youth-specific effect. They are helpful as a way to concisely summarize students' engagement over time, as measured through repeated measures ESM responses. This approach is superior to simply calculating the mean, which likely introduces bias into the estimation of models that use it (i.e., results that are statistically significant may be spurious (Gelman & Hill, 2007). For youth *j* and ESM response *i,* the models are:

Model 1A:

$$Y_{\text{predicted-engagement-}i} = {}_{0} + \text{ESM-engagement}_i * {}_1 + \text{gender}_j * {}_2 + \text{pre-interest}_j * + {}_i$$

$$_{0} = {}_{00} + \text{engagement}_j * \mu_1$$

Model 1B:

$$Y_{\text{post-interest-}i} = {}_{0} + \text{predicted-engagement}_i * {}_1 + \text{gender}_j * {}_2 + \text{pre-interest}_j * {}_3 + {}_i$$

Because this study uses an observational design, we use sensitivity analysis to quantify how robust our inferences are in light of potentially omitted, confounding variables, and other potential sources of bias (see Frank, 2000, for a description of the approach). Using the konfound R package (Rosenberg, Xu, & Frank, 2018), we describe what percentage of an effect would need to be due to invalidate inferences for effects.

The mixed effects modeling approach does not account for the uncertainty in the predictions of youths' ESM engagement when used to predict their post-program interest

(Houslay & Alastair, 2017). Accordingly, we were concerned about the possibility of the results appearing stronger than they should be. Therefore, we pursued a one-step modeling approach. We used a Markov Chain Monte Carlo (MCMC) approach in a Bayesian framework, estimating a model predicting youths' ESM engagement and post-interest as part of a single model, with the relationship between these two outcomes (after accounting for youths' pre-program interest and gender) being used to explain how their ESM engagement related to changes in their post-program interest. This method requires priors, or distributions representing the plausibility of candidate parameters; we briefly describe the "non-committal" priors we used in the absence of past research results or theory to guide their use in a more informative way, and to more simply compare the estimated effects of interest between the two approaches. We discuss the priors we specified in Appendix A. The results, which were very similar to those obtained from the two-step approach using mixed effects models, are presented in Appendix B and are briefly referenced in the discussion section.

Results

Preliminary Results

First, we report descriptive statistics and Pearson correlations for all study variables. Table 2 shows that youths' pre and post-program interest had similar means (pre-interest: $M =$ 3.04 ($SD = 0.90$); post-interest: 3.10 (0.91), measured on a scale with a range of one-four, indicating higher-than-average interest in STEM. The mean difference ($M_{\text{pre-interest}} - M_{\text{pre-interest}} =$ 0.03) in interest, using a paired t-test, was not significant ($t = 0.485$, p = 628), indicating that--not accounting for youth' engagement *during* the programs--youths' average interest does not change from before to after their involvement in the program. The correlation between pre and post-interest ($r = .59$) was high, indicating that youth who enter the program with high interest

are likely to leave with high interest. Note that this does not suggest that these youths' interest

has changed, but rather that there is a high degree of alignment between youths' pre- and post-

interest in STEM domains. Youths' mean level of engagement was 2.86 (0.86), indicating

moderately high youth engagement as measured by the approximately 2,700 ESM responses

youth completed throughout the programs. The correlation between ESM engagement and pre-

interest and post-interest was moderate ($r = .15$ and $.30$, respectively).

Table 2. Descriptive statistics for study variables.

Variable	Number of responses	Mean	Standard Deviation	Median
Pre-interest	193	3.044	0.90	3.33
Post-interest	160	3.08	0.92	3.33
ESM engagement	2714	2.862	0.864	2.83

Models Predicting Momentary Engagement and Post-Program STEM Interest

  First, we present results from mixed effects (or multilevel) models. Model 1A specified

youths' momentary engagement, measured via the ESM, as the outcome, with fixed effects for

youths' gender and pre-program STEM interest, and a random intercept--representing youth-

specific differences from the overall, mean levels of engagement--for each youth. The first

model, predicting youths' ESM engagement (their engagement as indicated by their responses

over the course of their time in the programs) show that females' engagement is estimated to be lower than males' (B = -.06 (SE = .09), p = .45). But, this effect did not reach the criterion for statistical significance. The effect of youths' pre-program interest was positive and statistically significant (B = 0.10 (0.05), p = .033), suggesting that youth with relatively higher levels of pre-program interest tended to be more engaged during daily program activities. However, only 6.81% of this inference would need to be due to bias to be invalidated, indicating that this effect is likely not very robust due to potential sources of bias. Model 2B predicted students' post-program interest on the basis of their gender, pre-program interest, and their predicted engagement (which is a youth-specific prediction of their engagement throughout their time in the program). The results for this outcome showed that female interest is estimated to be lower than for males (B = -0.14 (SE = 0.11), p = .239), though it did not reach statistical significance. Not surprisingly, the effect of youths' pre-program interest upon post-program interest is positive and statistically significant (B = 0.62 (SE = 0.06), p < .001) indicating that youth with high levels of pre-program interest tend to have high levels of post-program interest when they leave the program This effect was robust, with 78.27% needing to be due to bias in order to be invalidated. The effect of youths' predicted engagement during the course of programming upon their post-program interest was positive and statistically significant (B = 0.46 (SE = 0.10), p < .001), even after accounting for the effects of their pre-program interest and gender. This suggests that youth who sustain high levels of momentary engagement over the course of the program tend to experience increases in STEM interest over the course of the summer program, regardless of their level of STEM interest when the program began. This effect was also robust, with 53.34% needing to be due to bias for it to be invalidated. We also calculated a partial R^2 for this effect, finding that it indicates a moderate-sized effect (partial R^2 = .340).

Table 3. Results from mixed effects models with engagement.

Model 1A (Predicting Engagement)	B (SE)
Fixed parts	
Intercept	2.55 (0.17)
Gender (Female)	-0.06 (0.09, p = .45)
Pre-interest	0.10 (0.05, p = .033)
Random parts	
Youth (var)	.33
Model 2A (Predicting Post-interest)	
Intercept	1.24 (0.23)
Gender (Female)	-0.14 (0.11, p = .239)
Pre-interest	0.62 (.06, p < .001)
Predicted engagement	0.46 (0.10, p < .001); *r* = .340

Discussion

We discuss key findings and limitations and recommendations for research.

Key Findings

We sought to understand how youths' engagement in summer STEM programs impacted their development of interest. Instead of using youths' mean level of engagement across their ESM responses, we first used BLUPs from mixed effects models by modeling ESM engagement as an outcome, which were estimated along with the other parameters estimated in the first model. For this model, we found that youths' pre-program interest predicted their in-the-moment engagement (though the effect was small), suggesting that youth who begin summer STEM programs with higher initial interest have a (modest) tendency to be more engaged while

attending the program. Gender did not predict youths' in-the-moment engagement. Second, we used the BLUPs (as students' predicted interest over their time in the program) as predictors in the model with youths' post-program interest as the outcome. In alignment with Hidi and Renninger's (2006) model of interest development and in support of research (i.e., Dabney et al., 2012) on the impact of OST STEM programs, we found that students' sustained engagement over the course of the summer programs strongly predicted *changes* in their interest: For every one-unit change in their predicted engagement (when engagement is measured on a one-four scale), interest in STEM (also measured with a one-four scale) after the program was .46 units greater.

Particularly in light of the overall lack of change in youths' interest (the not statistically significant change in pre-post program interest in STEM), these results show that is not participation in summer STEM programing, but level of engagement in these program impacts their interest in STEM domains. These results suggest that youths' experiences in summer STEM programs strongly and positively relate to changes in interest in STEM domains.

Notably, when we used a more conservative MCMC approach (Appendix B), the effects were essentially the same, though with a slightly smaller estimated effect for the relationships observed. For example, the partial R^2 (calculated to more easily compare results between the two approaches) for the effect of youths' predicted engagement for the effect of youths' predicted engagement was moderate--and statistically significantly different from zero--using both modeling approaches (Partial R^2 from the mixed effects modeling approach = .34; Partial R^2 from the mixed effects modeling approach = .27). While both approaches yielded similar results in this case, in cases with less strong effects, the more conservative MCMC approach could be used to avoid making potentially spurious inferences (Houslay & Wilson, 2017). Particularly for

complex data structures, such an approach can compare and in some cases have advantages (i.e., in data with complex random effects structures) over other approaches, such as multilevel Structural Equation Modeling. While this was a study designed to ask and answer a substantive question about youths' engagement and how it relates to their interest, it demonstrates the use of a data analytic approach (i.e., the use of youths' predicted engagement from mixed effects models--and, in the appendix, models estimated with MCMC) that is particularly suited to exploring how in-the-moment experiences relate to their antecedents and outcomes. While we used ESM data, this approach could also be used for other methods of data collection, such as log-trace data, such as that from intelligent tutoring systems and online microworlds (Gerard, Ryoo, McElhaney, Liu, Rafferty, & Linn, 2015; Gobert, Baker, & Wixon, 2015).

Limitations and Recommendations for Future Research

While the use of ESM is a strength of the present study, one limitation concerns the specific items used and how we used them to measure youths' engagement. We measured engagement using a composite, aiming to tap the multiple dimensions of engagement in past research, but not distinguishing them in the analysis. As a consequence, we do not understand whether the changes are a consequence of behavioral, cognitive, and affective engagement (which includes an indicator of situational interest)--or some particular configuration of these dimensions. Future research can better explore *which types* of engagement matter in which ways in terms of their impact on youths' interest. To this end, person-oriented analyses, and analytic approaches such as Latent Profile Analysis, hold promise for understanding how the dimensions of engagement are experienced by youth. Researchers can also explore other outcomes, such as changes in youths' future goals and plans and their competence.

Another limitation concerns the sample of participating youth in the present study. The sample was collected in a highly purposive manner, such that programs designed and implemented with best practices for OST STEM programming were identified and selected as the context for this study. While the results use students' pre-interest to account for how their initial inclination toward STEM domains impacts their post-program interest, their experiences during the programs are highly contingent upon the activities they are involved in and the quality of the guidance and instruction of the youth activity leaders. Additionally, the sample was made up of youth almost completely from underrepresented (in STEM) groups of individuals. Accordingly, we are not yet able to say how these findings might generalize to other summer STEM programs, and considering how youths' engagement impacts their interest in other, different OST contexts may be a worthwhile aim of future research. Related to this limitation, there was a large degree of missing data, though not substantially higher than that in other studies that use ESM (Hektner et al., 2007). Our use of sensitivity analysis showed that more than 50% of the effect of youths' engagement upon the changes in their interest would need to be due to bias for it to be invalidated, suggesting that issues related to missing data would have to be very substantial (i.e., the effect of survivorship bias, in that only the youth who were most interested attended and completed all of the survey measures) for this effect to be negated were the data that are missing to be included. Nevertheless, particularly in OST settings, where data collection can present some challenges, future research can carefully consider and aim to mitigate potential impacts of missing data.

Conclusion

The aim of this study was to explore whether and how youths' engagement in summer STEM programs promotes their development of interest in STEM. Using ESM as a data

collection methodology and a unique mixed effects modeling approach for the data analysis, we found that youths' experiences in terms of their in-the-moment engagement led to changes in their post-program interest. The method of data collection (ESM) is an example of how intensive data methods (i.e., Bolger & Laurencau, 2013) can be used to ask questions that are difficult to answer simply using self-report surveys, which do not lend (as much) insight into the experience of individuals *during* the moment, activity, or programming we are interested in. These findings are in line with previous research suggesting that out-of-school time STEM programs can lead to increased STEM interest (Dabney et al., 2012; Elam et al., 2012).

In this study, we sought to add to this literature by examining one possible mechanism, engagement, for the development of youth interest. Engagement is widely understand by educators and scholars as important to many adaptive academic outcomes, including interest. As out-of-school time programs are well-suited to supplement STEM learning among youth with engaging activities, it is important for researchers to continue to focus on these informal learning environments as they are particularly promising vehicles for interest development.

These findings have some implications for practice as well as the research literature. The experiences that youth have in summer STEM programs have an impact on their STEM interest after their involvement in these programs; accordingly, youth activity leaders should seek to design activities that are highly engaging to youth. Research on engaging activities in summer STEM programs suggests expansion of initiatives to provide opportunities for youth, particularly those with less opportunities to develop interest in STEM domains, to participate in out-of-school time STEM programs.

References

Bates, D., Maechler, M., Bolker, B., & Walker, S. (201). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1-48. doi:10.18637/jss.v067.i01.

Bell, P., Lewenstein, B., Shouse, A.W., & Feder, M.A. (Eds.) (2009). Learning science in informal environments: People, places and pursuits. Washington, DC: National Academy Press.

Beymer, P. N., Rosenberg, J. M., Schmidt, J. A., & Naftzger, N. J. (2018). Examining Relationships among Choice, Affect, and Engagement in Summer STEM Programs. Journal of Youth and Adolescence, 1-14.

Bolger, N., & Laurenceau, J. P. (2013). Intensive longitudinal methods. New York, NY: Guilford.

Brophy, J. (2008). Developing students' appreciation for what is taught in school. Educational Psychologist, 43(3), 132-141. doi:10.1080/00461520701756511

Brophy, J. (2008). Developing students' appreciation for what is taught in school. Educational psychologist, 43(3), 132-141.

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

Fayer, S., Lacey, A., & Watson, A. (2017). STEM occupations: past, present, and future.

Frank, K. 2000. "Impact of a Confounding Variable on the Inference of a Regression Coefficient." Sociological Methods and Research, 29(2), 147-194 https://msu.edu/~kenfrank/papers/impact%20of%20a%20confounding%20variable.pdf

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

Gelman, A., & Hill, J. (2007). Data analysis using regression and multilevel, hierarchical models (Vol. 1). New York, NY, USA: Cambridge University Press.

Gerard, L. F., Ryoo, K., McElhaney, K. W., Liu, O. L., Rafferty, A. N., & Linn, M. C. (2016). Automated guidance for student inquiry. Journal of Educational Psychology, 108(1), 60.

Gobert, J. D., Baker, R. S., & Wixon, M. B. (2015). Operationalizing and detecting disengagement within online science microworlds. Educational Psychologist, 50(1), 43-57.

Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R Package. Journal of Statistical Software, 33(2), 1-22. http://www.jstatsoft.org/v33/i02/.

Harackiewicz, J. M., Smith, J. L., & Priniski, S. J. (2016). Interest matters: The importance of promoting interest in education. Policy Insights from the Behavioral and Brain Sciences, 3(2), 220-227.

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

Hidi, S. & Renninger, K. A. (2006). The four-phase model of interest development. Educational Psychologist, 41(2), 111-127.

Hidi, S., Renninger, K.A., & Krapp, A. (2004). Interest, a motivational variable that combines affective and cognitive functioning. In D. Y Dai & R.J. Sternber (eds.), Motivation, emotion and cognition: Integative perspectives on intellectual functioning and development (pp. 89-115). Mahwah, NJ: Lawrence Erlbaum.

Hill, C., Corbett, C., & St Rose, A. (2010). Why so few? Women in science, technology, engineering, and mathematics. American Association of University Women. 1111 Sixteenth Street NW, Washington, DC 20036.

Hirsch, B. J., Mekinda, M. A., & Stawicki, J. (2010). More than attendance: The importance of after-school program quality. American Journal of Community Psychology, 45(3), 447–452. https://doi.org/10.1007/s10464-010-9310-4.

Houslay, T. M., & Wilson, A. J. (2017). Avoiding the misuse of BLUP in behavioural ecology. Behavioral Ecology, 28(4), 948-952.

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.

National Academy of Engineering and National Research Council (2014). STEM integration in K-12 Education. Washington, DC: National Academies Press.

OECD (2016), PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematics and Financial Literacy, OECD Publishing, Paris. http://dx.doi.org/10.1787/9789264255425-en

Prenzel, M. (1992). Selective persistence of interests. In K.A. Renninger, S. Hidi & A. Krapp (eds.), The role of interest in learning and development (pp. 71-98). Hillsdale, NJ: Lawrence Erlbaum.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rosenberg, J. M., Xu, R., & Frank, K. A. (2018). Konfound-It: Interactive application (and Stata procedure and R package) to carry out sensitivity analysis. http://konfound-it.com.

Shernoff, D. J., & Schmidt, J. A. (2008). Further evidence of an engagement–achievement paradox among US high school students. Journal of Youth and Adolescence, 37(5), 564-580.

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

Skinner, E., Pitzer, J., & Steele, J. (2013). Coping as part of motivational resilience in school: A multidimensional measure of families, allocations, and profiles of academic coping. Educational and psychological measurement, 73(5), 803-835.

Appendix A: Results using Markov Chain Monte Carlo (MCMC) estimation

MCMCglmm methods, unlike ML, requires priors. One view of these priors is that they constrain the possible values that parameters may take in order to set the modeling up for success (Houslay & Wilson, 2017). Another way to look at priors is to consider them as part of a Bayesian approach, in which they represent the degree of belief in different parameter values (Gelman & Hill, 2007). There are also cases when the prior values can be estimated from the data in the sample. Gelman and Hill (2007) describe multi-level models in these terms: For the "random" effects, usually "grouping" variables like the classroom students are in, for example, the prior for the classroom-specific effects is estimated on the basis of the mean and variance in the dependent variable from the whole sample or data set collected. In these cases (in which the prior for "random" effects can be estimated from the data), the priors for the other variables can be set to be neutral, with the aim to constrain the possible values that parameters may take rather than to be used as part of a fully Bayesian approach. Accordingly, for the models specified for this study, we specified weak priors that are consistent with a wide range of possible parameter values. Some authors refer to these as "uninformative" priors; for identical model specifications, the results from use of uninformative priors with MCMCglmm and ML are identical. When we run the MCMC procedure, we follow guidelines by Kruschke (2015) for checking the representativeness (by checking the burn-in period and the convergence of multiple chains), accuracy (by checking the effective sample size and the the Monte Carlo standard error (MCSE) for estimates), and efficiency of the estimation (by running multiple chains in parallel and by way of using the efficient MCMCglmm package.

Appendix B: MCMC results

In this appendix, we present results from models estimated with MCMC. Because this approach accounts for the uncertainty in the estimates of youths' ESM engagement, it is expected to yield results that are less biased than those obtained from the two-step approach. In this approach, using the MCMCglmm (Hadfield, 2012) we first specify the model. More information on the model specified (and the priors required) is described in Appendix A.

Here is the model estimated with MCMC, for individual j and ESM response i:

$$Y_{engagement-i} = \beta_{01} + esm\text{-}engagement_i * \beta_{11} + gender_j * \beta_{21} + pre\text{-}interest_j * \beta_{31} + \epsilon_i$$

$$\beta_0 = \beta_{00} + engagement_j * \mu_1$$

$$Y_{post-interest-i} = \beta_{02} + gender_j * \beta_{12} + pre\text{-}interest_j * \beta_{22} + \epsilon_i$$

$$Y_{engagement-i} \sim Y_{post-interest-i}$$

This is specified as a multivariate model, with the relationship between ESM engagement and post-interest examined through their covariance ($Y_{engagement-i} \sim Y_{post-interest-i}$). The results are presented in Table B1. To determine statistical significance, we used the Highest Density Intervals to determine whether 95% of the posterior distribution did not include zero.

Table B1. Results from models with engagement estimated with MCMC.

Fixed parts (Predicting ESM engagement)	B (SE)
Intercept	2.53 (.24)
Gender	-.06 (.09)
Pre-interest	.11 (.07)
Fixed parts (Predicting post-interest)	

Intercept	1.71 (.85)
Gender	-.20 (.16)
Pre-interest	.47 (.20)*

Random parts

ESM Engagement intercept (var.)	.33
Post-interest intercept (var.)	.61
ESM Engagement and Post-interest (corr.)	.27***

* = p < .05

Appendix C: Code

The code to reproduce the analysis is available here:

https://github.com/jrosen48/mcmcglmm/blob/master/mcmgglmm-example-3.Rmd