

RESEARCH

Open Access

Measuring informal STEM learning supports across contexts and time



Bradley J. Morris^{1*} , Whitney Owens², Kirsten Ellenbogen³, Sibel Erduran⁴ and John Dunlosky⁵

Abstract

Background: Informal science activities are critical for supporting long-term learning in STEM fields. However, little is known about the kinds of activities children and their families engage in outside of formal settings and how such activities foster long-term STEM engagement. One gap in the literature is the lack of data that document self-designated STEM activities and measure their impact on later engagement with learning opportunities that are distributed over time and contexts (i.e., the informal learning ecology). One reason for this gap is that there has been little measurement during the events, because using only a few measures (which can be completed briefly) may reduce psychometric validity. We developed an instrument, the STEMwhere app, to measure four informal science learning supports (interest, engagement, identity, and goal setting), across the informal learning ecology. For a period of 2 months, 26 children ages 7–14 used the app to check-in during STEM activities and answer eight questions about each activity.

Results: The results demonstrated that most STEM activities occurred in the home, often consisted of hands-on activities, suggesting that the family home provides more opportunity for engagement than other locations. Child interest and engagement ratings were high in all settings and activities suggesting that high situational interest was relatively common during these activities. Further, user ratings suggested relations between different learning supports. For example, increases in interest were related to increases in subsequent engagement and “fun” goals, while increases in engagement were related to increases in learning goals. By collecting participant-generated check-ins, we identified periods of increasing activity and their likely triggers, which is a novel measure we refer to as topical runs. We operationally defined a run as a pattern of check-ins that were unlikely to occur by chance and shared a topic or location.

Conclusions: Our results serve as both a proof-of-concept for a novel tool for measuring informal STEM activity in the wild that provides data consistent with existing measures and provide novel findings that contribute to our understanding of where and how informal science activity occurs.

Keywords: Informal learning, STEM, App-based measurement, Engagement, Interest

Introduction

Measuring informal STEM activity across multiple time points distributed across multiple contexts (hereafter, the informal learning ecology; Erdogan & Stuessy, 2015) is important because surprisingly little is known about what types of experiences outside of school ignite and sustain children’s STEM learning (Science, Technology, Engineering, and Mathematics; National Academies of Sciences, Engineering, and Medicine (NASEM), 2016). In this paper, we describe the results of a preliminary investigation using a novel tool, the STEMwhere app, to measure informal

STEM activity across the informal learning ecology. Specifically, we asked 26 children between ages seven and fourteen to log into the app whenever they were engaged in a STEM activity, as defined by the participant, and to rate their interest, engagement, identity, and learning goals. The data allowed us to measure STEM activity at multiple points in time and across multiple contexts, which provided new findings regarding where STEM learning occurs and how these activities influenced support for learning over time.

Why is informal STEM learning important?

A well-trained STEM workforce is necessary for economic success in the twenty-first century because business, industry, and government agencies increasingly rely on workers with STEM skills (NASEM, 2016a). We follow the National Science Foundation (USA) in defining STEM as any activity

* Correspondence: bmorri20@kent.edu; <http://bmorri20.ehhs.kent.edu/bmorri20>

¹Department of Educational Psychology, Kent State University, 405 White Hall, Kent, OH 44242, USA

Full list of author information is available at the end of the article

within one of the four content areas denoted by the acronym (see Gonzalez & Kuenzi, 2012; National Academies of Sciences, Engineering, and Medicine, 2018). Further, we anticipate that many activities could represent multiple areas of STEM content, though these activities may or may not be integrated within the activity (Kelley & Knowles, 2016). STEM education is the primary driver for helping children become adults with STEM skills; however, criticism of STEM education has arisen within the US educational system because little progress has been made on international test scores (Provasnik, Malley, Stephens, Landeros, Perkins, & Tang, 2016; Grønmo, Lindquist, Arora, & Mullis, 2015). Moreover, although interest in STEM is common in early childhood, a steady decline in the number of students engaging in STEM courses and activities occurs through secondary school to universities (NASEM, 2016a). Thus, only a fraction of those interested in STEM as children begin a STEM career as adults (NASEM, 2016b).

Informal STEM learning has the potential to be an important part of the educational support system for helping young children move from nascent interest in STEM to a career in STEM (i.e., the STEM pipeline; NASEM, 2016a). A recent estimate suggests that children spend only 20% of their time learning in formal educational environments (Eshach, 2007; Falk & Dierking, 2010; Sacco, Falk, & Bell, 2014), which suggests that informal environments provide the means to augment learning in formal settings (Eshach, 2007; Morris, Masnick, Baker, & *Junglen, A., 2015) and may support long-term engagement (Falk, Storksdieck, & Dierking, 2007; Jones & Stapleton, 2017). Anecdotally, biographies of notable scientists suggest that their interest in their field emerged from informal, rather than formal experiences. For example, Neil deGrasse Tyson's interest in astronomy began with a trip to the Hayden planetarium (Farmer & Shepherd-Wynn, 2012). Even if such experiences do not increase STEM knowledge, they might support learning by increasing interest or identity, which are non-cognitive supports for learning (Fenechel & Schweingruber, 2010). To promote such learning, it is important to understand what types of activities lead to later engagement. Although some activities (e.g., a trip to the planetarium) might fit traditional definitions of STEM, other types of activities (e.g., playing with a drone) might not fit traditional definitions. In either case, both types of activities may trigger an increase in STEM engagement and are both included in our use of STEM activity. In summary, a promissory note of informal STEM learning is that it will contribute to building a healthier pipeline to science careers by promoting STEM learning.

Supporting STEM learning

STEM learning is traditionally defined as the accumulation of relevant knowledge and processes of science (e.g., creating unconfounded experiments; Zimmerman, 2007), which emerge by augmenting cognitive mechanisms through

experience with culturally transmitted knowledge (Morris, Croker, Masnick, & Zimmerman, 2012). However, this accumulation of knowledge occurs within STEM-related activities and requires non-cognitive supports such as interest, engagement, and identity (see Table 1 in the "Method" section for conceptual and operational definitions; Fenechel & Schweingruber, 2010). Learning supports are critically important for acquiring STEM knowledge because learning opportunities alone are not sufficient, as demonstrated by students who attend science class but fail to learn relevant content. However, there is a dearth of data measuring children's informal STEM activity across the informal learning ecology that furthers our understanding of how specific learning supports during these activities facilitate later science learning (Dorph, Schunn, & Crowley, 2017).

We focus on four learning supports for knowledge acquisition. *Interest* refers to attention to content over time (Hidi & Renninger, 2006; Maltese, Melki, & Wiebke, 2014). *Engagement* is involvement or participation in content, which results in positive emotional reactions (Eberbach & Crowley, 2017; Milne & Otieno, 2007). Although interest and engagement can be mutually influential (e.g., being interested in a topic often drives engagement), interest is a psychological state of which a student may be unaware (e.g., interest triggered by a phenomenon), whereas engagement generally refers to participation in an activity or event, so the two need not be related (e.g., one may be engaged without being interested; Renninger & Bachrach, 2015). Moreover, interest is often generated by and limited to captivating phenomena (i.e., situational interest) but can be maintained over long periods of time (well-developed individual interest; Renninger & Hidi, 2015). Well-developed STEM interest often manifests itself into sustained STEM engagement with STEM topics, which leads to knowledge acquisition (Thoman, Sansone, & Geerling, 2017).

Identity is recognition by a person and others that she or he can contribute to a STEM field. The recognition from another person (e.g., a parent or teacher) appears to be important in acquiring and sustaining STEM identity (Barton & Tan, 2010). Positive STEM identity is a predictor of higher levels of engagement and motivation (e.g., "I could become a scientist"; Calabrese Barton & Berchini, 2013). In a recent investigation, students who read stories about the intellectual struggles of famous scientists (e.g., Einstein had trouble in school) were more likely to view their own struggles positively (a component of identity) and improved their grades more than students who read about only the intellectual achievements of these scientists (Lin-Siegler, Ahn, Chen, Fang, & Luna-Lucero, 2016). Positive STEM identity helps to sustain the drive to pursue STEM over the long-term, a critical factor in choosing a career in STEM (Barton & Tan, 2010).

In addition to the motivational supports identified above (interest, engagement, and identity), goal setting is a central

Table 1 STEM learning supports: conceptual and operational definitions

Learning support	Conceptual definition	Operational definition	
		Self-report	Behavioral
Interest	Engaging with or reengaging with content over time (Hidi & Renninger, 2006)	Rating interest on a 10-point Likert scale	Change in the number of and duration between check-ins; increasing rates suggest increasing interest
Engagement	Involvement or participation in content, resulting in positive emotional reactions (Milne & Otieno, 2007; Wang et al, 2016)	Rating engagement on a 10-point Likert scale	Patterns in user-initiated check-ins; High frequency check-ins on the same topic
Identity	A recognition by a person and others that she or he can contribute to a STEM field (Calabrese Barton & Berchini, 2013)	Separate child and parent rating agreement with statements about seeing themselves in STEM fields on a 10-point Likert scale	Counting the number of user-initiated check-ins related to identity ratings; counting specific terms related to identity (see Table 2 for coding rubric)
Goal setting	A key component of self-regulated learning and improves learning outcomes (Pintrich, 2000)	Rating statements about three types of goals (fun, entertainment, or learning) on a 10-point Likert scale	Counting the number of user-initiated check-ins related to previous learning goals

component of information-processing models of self-regulated learning (Dunlosky & Metcalfe, 2008; Winne, 2001; for varying theoretical perspectives on self-regulation, see Zimmerman & Schunk, 2001). Self-regulated learners are active agents who set goals for their learning and then attempt to evaluate and regulate their learning, interest, motivation, and performance based on their knowledge about how best to meet these learning goals (Pintrich, 2000; Winne, Hadwin, Schunk & Zimmerman, 2008). People who set goals are more productive and effective, so measuring the kinds of goals that children set while engaging in informal STEM learning can offer insights into students' interest, engagement, and ultimate achievement. In summary, the combination of learning supports (goal setting, interest, engagement, and identity) provides important links between informal experiences and subsequent persistence and achievement in STEM. Although all have been investigated separately, there is a gap in our knowledge of the relation between STEM experiences and increasing STEM activity. One reason for this gap is the methodological limitation in collecting data that reveal the types of events that ignite STEM interest and reveal their impact over time.

One goal of the present project is to develop a tool to measure STEM activities and their impact on interest, engagement, identity, and learning goals (hereafter *learning supports*¹). The interaction between how learners engage with STEM activities will influence what one takes away from a particular experience—whether it be attending a science center or working with a home weather station (Winne & Hadwin, 1998; Bjork, Dunlosky, & Kornell, 2013; for reviews, see Zimmerman & Schunk, 2001). For example, a child interested in baking might be more likely to set learning goals in this context (e.g., seeing the need to learn about fractions to double a recipe) than a child less interested in baking. As a result, each support might influence other supports, depending on previous experiences, and relevant contexts. A child with a high interest in baking might stay engaged with this topic for a longer period of time as

compared to a child with little interest. In this way, these learning supports may benefit or hinder learning in a dynamic fashion in which changes in one support influences others, which may lead to additional engagement or learning opportunities (Eberbach & Crowley, 2009). Returning to our previous example, a child who had a positive experience learning fractions while baking with a parent might be more interested in learning fractions while solving relatively unengaging worksheets in a classroom. In a survey of university students enrolled in STEM programs, many students reported their interest in STEM being sparked by a range of activities that might not be perceived as being true STEM activities, such as watching television programs, playing outdoors, and doing activities with their families (Maltese et al., 2014). Although informal experiences have the *potential* to improve formal learning, the evidence suggests that translating the enthusiasm from informal experiences into classroom learning is difficult and often unsuccessful (Nasir, 2008; Saxe, 1988; Stevens, Satwicz, & McCarthy, 2008). Thus, it is critical to follow what participants consider STEM experiences, because it is unclear which types of experiences lead to increases in activity and improvements in learning.

Measuring the informal learning ecology and activity within this ecology

STEM activity occurs within a learning ecology, which refers to learning opportunities that are distributed over time and space (Erdogan & Stuessy, 2015). To understand the impact of the learning ecology on informal STEM activity over a period of time, measures are needed to address the following questions: *where* do informal activities take place, with *what* kinds of STEM activities are children engaged, *why or how* are children engaged (as reflected by the learning supports),

¹We use this phrase to sidestep any confusion with either a subset of learning strands commonly used in the informal science learning community (see Fenechel & Schweingruber, 2010) or traditional definitions of learning as knowledge accumulation (e.g., Zimmerman, 2007).

and *when* (and how often) are the activities taking place? Unfortunately, easy-to-use tools are not available for measuring learning outside of formal settings, so it is not surprising that the National Science Teachers' Association recently included a specific declaration outlining the need for improving measurement of learning in informal contexts (NSTA, 2012).

We developed the STEMwhere app to begin meeting these measurement needs (what, when, where, and why of informal STEM learning) while attempting to meet the following criteria. The measurement app should be (1) relatively unobtrusive to use, (2) embedded in STEM experiences, (3) relatively easy to use across ecologies, and (4) used to collect multiple kinds of data (self-report and behavioral). The STEMwhere app allows for real-time measurement of STEM activity by asking users to "check-in" when engaged with STEM content (by the user's definition) and to answer a small number of questions regarding their experience. Thus, the app allows researchers to collect data during the experience, rather than only retrospectively (see Table 1 for measurement details).

The approach we have outlined uses a small number of measures in a brief period of time that may limit the psychometric validity of the constructs being investigated. However, we chose this approach for two reasons. First, instruments that record immediate data (termed experience sampling) often produce results that are more accurate than retrospective instruments (Alliger & Williams, 1993). Retrospective reports tend to under-report negative experiences (Piasecki, Huffor, Solhan, Trull, 2007) and overreport positive experiences compared to data from experience sampling (Stone et al., 1998). A good example is the "beeper studies" of adolescent moods that demonstrated that moods were much more consistent than had been reported in retrospective surveys (Larson, 1989). Second, approaches that require considerable effort and time on the part of participants might not be well suited for particular topics. In the present study, we are interested in the types of STEM activities in which families are engaged. Although previous research has used retrospective surveys (Fredricks et al., 2016), such methods incur time and effort costs on the part of participants that may reduce response validity. In addition, time-consuming surveys highlight response biases (e.g., availability biases) that may reduce the accuracy of what is reported (Bradburn, Rips, & Shevell, 1987). For these reasons, we chose to use experience sampling, while we acknowledge its potential limitations. Finally, we also measured constructs that have been investigated in larger-scale studies, so as to evaluate whether the current instrument yields data to support similar conclusions. Perhaps most notably, research has established that students' interest is predictive of subsequent engagement (Thoman et al., 2017), so we expect to find the same relationship in the present study (unless, of course, the small number of observations undermine the validity or sensitivity of our measure to appropriately characterize these relationships).

Any time participants were engaged in a STEM activity, they were instructed to open the app. When participants opened the app, they entered their participant number, age, and gender. Participants next selected their current location, which prompted the GPS coordinates for that location to be recorded. Participants were next asked to select a description of their location from a drop-down menu (my home, museum, library, park/arboretum/nature, science fair/exhibit, maker space, camp, or other). The next question asked participants to select the closest match to the question "What are you currently doing?" from a drop-down menu (watching TV/movie, watching web-based content, doing a web-based activity, listening to radio/streaming/podcast, listening to a speaker, hands-on activity, other). These options were provided because they are frequently discussed in informal STEM research (Thoman et al., 2017). After both questions, participants could enter additional information about their location or activity in a text box. The next eight screens each displayed one question (e.g., How interested are you in this activity?) with a 10-point sliding response bar with emoticons associated with text descriptions (e.g., interest: not at all, somewhat, partially, interested, extremely interested). Six questions were for children and two questions were for their parents. Child questions were relevant to interest, engagement, fun goal, learning goal, social goal, and child identity. Parent questions measured parent ratings related to their child's STEM identity (e.g., the belief that their child could become a scientist) and ratings of parent interest/engagement. Full questions and scales are provided in the Appendix.

Overview of research questions

A main goal of the present investigation was to create a tool for collecting information about informal STEM activities, to measure how learning supports are related, and how they influence activity across time and contexts. As foreshadowed by our prior discussion of STEM activities, we focused on answering the following key questions:

1. *What types of experiences do children consider STEM-related?*
2. *Where do STEM experiences occur and what kinds of experiences occur?*
3. *To what extent are self-reported ratings for STEM learning supports related to each other?* For example, is interest in a specific event related to future engagement in that event? And is interest related to learning goals or STEM identity?
4. *Are self-reported ratings related to behavioral measures (i.e., check-in rates)?* For example, as predicted above, do relatively high self-reported ratings of interest predict behavioral changes in engagement (i.e., an increase in the frequency of check-ins)?

Method

Participants

Participants were recruited by posting flyers in locations in Northeastern Ohio frequented by families with children including libraries, parks, bookstores, and churches and by posting to family-related listservs in this region (e.g., activities for families and children). Twenty-one families participated in the study. These families included 26 children, who ranged in age from 7 to 14 years old (mean age 11.5 years), were 58% female, and were 60% White, 20% Asian-American, 10% African-American, and 10% Indian-American. Participation occurred between June and August so that families would have more time to engage in “out of school” activities. We acknowledge that there is likely to be variation in the amount of engagement in informal STEM experiences across the calendar year and that summer is likely to represent a period in which there is higher family engagement levels compared to times when school is in session.

Procedure

Families were given an iPad on which the STEMwhere app was loaded and families were instructed to use the app anytime they were engaged with STEM content². Importantly, we did not define STEM activities so that we could measure what participants considered STEM content. Each family was compensated in two ways, they were given a family membership to a local science center and they received cash compensation (\$145) to cover their expenses related to participation (e.g., transportation costs and food while attending the science center) during the data collection period. The science center membership was provided before data collection to make sure that all participating families would have access to this setting, regardless of their financial situation.

Measures

The STEMwhere app measured STEM learning supports in two ways: self-reports and behavioral measures (Table 1). Self-report measures used Likert scales to measure each construct. We provided several categories of activities based on most likely settings for informal STEM activities and an “other” option for settings not covered by these options (Fenechel & Schweingruber, 2010; NASEM, 2016a). Because self-report measures are often subjective, behavioral measures were included in an attempt to provide more objective data from which to triangulate the measurement of learning. In addition, the behavioral measures provided novel data that could be used to uncover new elements to informal learning, specifically to map the learning ecology and to measure increases in interest and engagement.

Coding

The following section provides details regarding the coding criteria for the behavioral data (see Table 2). The first

coding was to classify the number of check-ins for each participant by location and type of activity. The app allowed participants to decide whether an activity was considered STEM. Obviously, a limitation to this approach is that participants might include activities that scientists and educators would not consider to be STEM activities. Accordingly, all activities were coded independently as to whether the activity should be categorized as a STEM activity, following the National Science Foundation definition. Those not coded as fitting the definition were eliminated from further consideration. Also, for those activities that were considered STEM activities, we classified which type of STEM content best described it (e.g., math). Although it is possible that an activity was interdisciplinary, it is quite difficult to identify activities as such. We used a conservative coding system that did not allow for multiple categories in order to minimize type I errors in coding. Check-in rates were calculated to provide a mean and standard deviation for each participant expressed as check-ins per day. Participant-generated descriptions were classified by location and activity and coded for the constructs outlined in Table 2. Two independent raters coded the descriptions. The initial reliability was .96 (Cronbach’s kappa) and 100% after discussion. A total of 334 activities were collected and 54 (16%) were eliminated after coding, leaving a total of 280 activities.

Topical runs

One goal was to measure increases in interest and engagement in STEM activities (hereafter runs). Topical runs were coded by identifying a series of check-ins that share a topic or event type, occur close together in time, and are unlikely to occur by chance. These runs were identified by comparing the frequencies of check-in runs to the probability of clusters generated from a Monte Carlo simulation (Heth & Cornell, 1987). Once identified, an analysis was conducted to identify whether the conditions immediately before a run predicted its occurrence (hereafter a trigger; Renninger & Bachrach, 2015).

Results

The analyses relevant to each research question are presented in the order that the questions were presented at the end of the “Introduction” section. Comparisons for age and gender yielded non-significant results, so age and gender were not considered a factor in the following analyses.

²The app was also used to record measurements before and after visiting three specific contexts: a science center, reading an online newspaper article about science, and watching an online video about science. Unfortunately, due to a technical issue with the app, nearly half of participants were missing questions related to these locations. Therefore, these data were not included in our analyses.

Table 2 Coding categories for behavioral constructs

Learning support	Operational definition	Coding
Interest	Number of and duration between check-ins	Code number of check-ins associated with specific events or locations
Engagement	User-initiated check-ins, user-entered descriptions	Identify runs related to the same events or location. Runs are defined as check-in patterns that occur in less than 1% of Monte Carlo simulations.
Identity	User-initiated check-ins, user-entered descriptions	Count specific key words related to STEM identity: 1. positive/negative evaluation terms (e.g., like, fun, enjoy) vs. (hate, etc.). related to topics, performance, and abilities (evaluations after activities as suggested by Hughes, Nzekwe, & Molyneaux, 2013) 2. Activities that influence identity (e.g., reading about scientists from under-represented groups)
Goal setting	Counting the number of user-initiated check-ins related to previous learning goals	Count specific key words related to learning goals (e.g., learn, tried to, wanted to, I did it!), social goals (friends, family, dad, mom, etc.), and fun goals (e.g., wanted to have fun, play)

What types of experiences do children consider STEM-related?

There were a total of 334 check-ins by participants. Fifty-four activities (16%) were eliminated after coding, leaving a total of 280 activities. An analysis of the data before and after removing these events showed no significant differences, suggesting that removing these activities did not change the overall result pattern. The results below are based on analyses after non-STEM activities were removed. Nearly all of the check-ins were related to science (52%) or mathematics (44%), with a small number related to technology or engineering (10% and 2%, respectively). Most science check-ins were classified as reading books and doing hands-on activities whereas most mathematics check-ins were related to hands-on measurement (e.g., building/cooking) or math activities (e.g., worksheets).

Where do STEM experiences occur and what kinds of experiences occur?

The majority of check-ins occurred in the family home (72%), followed by camps (10%), parks/nature (8%), library (6%), and museums (4%). The majority of descriptions were hands-on activities (64%), followed by web-based activities (15%), and watching TV/movies (10%). The next question was to what extent do self-reported ratings differ by location and type of activity? A one-way ANOVA demonstrated significant differences in ratings by location, $F(5, 280) = 8.3, p = .001$ and type of event, $F(5, 280) = 12.2, p = .001$. Least squared difference post-hoc tests were conducted to estimate significant differences between rating categories. Library check-ins were associated with the highest ratings for setting learning goals, parent engagement, and parent identity, whereas camp check-ins had the highest ratings for friend/fun goals and child identity (see Table 3 for detailed results). For activities, web-based content was given the highest ratings for interest, engagement, and learning goals, and TV/movies were given the highest rating for fun goals.

To what extent are self-reported ratings for learning supports related to each other?

A critical question is the extent to which individual learning supports were related to each other. Thus, the next set of analyses comprise a series of exploratory regressions measuring the unique variance for each specific support explained by other supports.

Interest as a predictor

Approximately 48% of the unique variance in the engagement rating was explained by the interest score ($R^2 = .48, F(1, 280) = 330.1, p < .001$), suggesting that higher interest ratings were related to higher engagement ratings ($\beta = .88$). Critically, this outcome is consistent with outcomes from previous research (e.g., Thoman et al., 2017). The second regression indicated that interest accounted for approximately 45% of the unique variance in setting fun goals ($R^2 = .45, F(1, 280) = 245, p < .001$), 11% of the variance in setting learning goals ($R^2 = .11, F(1, 280) = 42, p < .001$), and only 1% of the variance in setting social goals ($R^2 = .01, F(1, 280) = 24, p < .08$), suggesting that increases in interest were related to increases in fun goal setting but little change in learning and social goal setting ($\beta = .92, .51, \text{ and } .10$, respectively). Interest ratings were moderately related to changes in identity ratings ($R^2 = .10, F(1, 280) = 44, p < .001, \beta = .58$).

Engagement as a predictor

Engagement accounted for approximately 19% of the unique variance in setting fun goals ($R^2 = .19, F(1, 280) = 72, p < .001$), 29% of the variance in setting learning goals ($R^2 = .29, F(1, 280) = 122, p < .001$), and 0% of the variance in setting social goals ($R^2 = .00, F(1, 280) = 0, p > .993$). Higher engagement ratings were related to higher ratings for learning goal setting and fun goal setting but were unrelated to social goal setting ($\beta = .572, .458, \text{ and } .001$ respectively), consistent with previous research suggesting that engagement is associated with goal setting and achievement (Greene & Miller, 1996). Finally, higher engagement ratings were related to higher identity ratings ($R^2 = .27, F(1, 280) = 96, p < .001, \beta = .62$).

Table 3 Self-reported event ratings by location and activity

Activity	Location					Activity				
	Home	Camp	Park	Library	Museum	Web video	Web-based	TV/movie	Hands-on	Podcast/speaker
Interest	7.9 (2.6)	7.7 (2.2)	7.6 (1.4)	8.1 (3)	7 (5)	9.5 (2)	8.1 (2.1)	9.1 (1.2)	7.4 (2.7)	7.4 (1.2)
Child engagement	7 (2.6)	6.2 (2.3)	6.6 (1.9)	7.4 (3.3)	7.2 (1.2)	8.7 (1.6)	7.2 (2)	7.6 (2.7)	6.6 (2.6)	8.3 (1.9)
Parent engagement	6.6 (2.9)	7 (2)	7.8 (1.7)	8.8 (2.4)	7 (2.73)	7 (2.5)	7.7 (2.1)	4.1 (3.3)	7.2 (2.5)	8.1 (1.4)
Goal-friend	4.4 (3.2)	8.5 (1.9)	7.6 (3.2)	7 (3.6)	7.7 (5)	1.8 (2.3)	4.4 (3.1)	6.2 (3.4)	5.8 (3.3)	3 (2.8)
Goal-learning	6.6 (3)	7 (1.7)	5.1 (2)	7.4 (2.4)	6.5 (1)	8.2 (2.4)	6.5 (2.8)	7.2 (3)	6.4 (2.9)	8 (2.8)
Goal-fun	7.3 (3)	7.4 (3.6)	7.7 (2.5)	7.6 (3.3)	7.5 (1)	6.7 (2.8)	8.4 (2.6)	8.9 (1.7)	7 (1.4)	6.9 (3.2)
Child identity	5.3 (3.3)	6.8 (2.2)	5 (3)	6.3 (3.5)	4.7 (2.5)	5.5 (2.5)	6.4 (3.3)	6.2 (3.2)	5.2 (3.2)	6.7 (2.6)
Parent identity	5.8 (3.2)	7 (2.3)	6.1 (2.1)	7.2 (1.9)	3 (.5)	7.7 (2)	7.2 (2.3)	3.8 (3.1)	6 (3.1)	7 (4.2)

Note: Table displays means with standard deviations in parentheses. Cells in boldface indicate a significant difference at the $p > .05$ level based on the results of LSD post hoc tests

Parent-child ratings as predictors

Recall that parents provided two ratings: plans for engagement and parent ratings of STEM identity. Higher parent engagement ratings were related to higher child identity ratings, $R^2 = .27$, $F(1, 280) = 104.7$, $p < .001$, $\beta = .69$, and higher parent engagement ratings were associated with higher child engagement ratings, $R^2 = .41$, $F(1, 280) = 52.89$, $p < .001$, $\beta = .44$. Higher parent identity ratings were related to higher child identity ratings, $R^2 = .31$, $F(1, 280) = 84.98$, $p < .001$, $\beta = .73$, and moderately related to child engagement ratings $R^2 = .08$, $F(1, 280) = 38.6$, $p < .001$, $\beta = .39$. These results provide more evidence that parents play an important role in supporting engagement with middle-school age children and in supporting their children’s STEM identity (Barton & Tan, 2010).

Are self-reported ratings related to behavioral measures?

The behavioral data from the STEMwhere app can augment the self-report data. One possibility that we investigated was that the frequency of check-ins would also be a reasonable measure of interest and engagement. Presumably, the amount of interest would correspond to increases in the frequency of check-ins. Imagine that a child checks in 10 times over the course of 30 days for an average of one check-in every 3 days. If check-ins were not random, then the patterns might demonstrate something about the interest and engagement that is driving them. For example, if this child checks in five times in 2 days, this cluster may suggest a rapid increase in interest and engagement, assuming the events share similar topics (i.e., content or activities).

To evaluate such possibilities, the following analysis identified *topical runs*, defined as event clusters that shared common topics and were unlikely to occur by chance. Runs were identified by comparing the frequencies of check-in runs to the probability of clusters generated from a Monte Carlo simulation (Heth & Cornell, 1987). Specifically, we first calculated the total number of check-ins and the mean and standard deviation of the durations between check-ins (in days) for each participant. From these data, we created

simulations of check-ins to calculate the probability of different runs (i.e., sequences of check-ins) that occurred in less than 3% of simulations (see Fig. 1).

We then identified the activities that took place 1–2 days before and during each run to determine if the location or type of activity were similar within each run (shaded in Fig. 1). Ten runs were identified because they occurred in less than 3% of simulations. Of those seven runs, six shared the same topic (see Table 4). Topic themes were coded by two independent raters with an initial reliability of .97 (Cronbach’s kappa) and 100% agreement after discussion.

The topical run from Participant 26 began with the purchase of an inexpensive home weather station (e.g., displayed temperature, humidity). After this event, check-ins significantly increased, all of which were related to weather (e.g., reading books about tornadoes and storms, watching web-based videos on weather, watching the Weather Channel). The run that did not share a topic, was the result of a participant “catching up” on check-ins that had occurred at other times. Specifically, the participant noted that her family had been away on a trip, so she entered many check-ins upon their return home. In this case, the pattern was an artifact of access to the app. Although these patterns indicated an increase in STEM engagement, the locations and activities did not share common features.

An additional detail that clarifies these data is that four runs occurred within the same families. Participants 1 and 26 are siblings and both showed similar increases in interest and engagement in weather after the purchase of a home weather system. The pattern of check-ins from participant 25 were related to an interest in cooking and baking that emerged on this vacation. This suggests that family engagement can be viewed as a contributing factor but not a sufficient condition for increasing STEM engagement. In summary, check-in rates provided a novel measure of interest and engagement.

One issue still remaining is whether self-report data were consistent with behavioral data. For example, if a participant rated an event as being highly interesting and planned on engaging again in this activity, did he or she engage in this

activity again? To answer such questions, we compared self-report data associated with runs (as identified above) to mean self-report data using paired *t* tests. The results indicated no significant differences between the self-reported rating of events and run means (see Table 5), suggesting that the self-report data were not closely aligned with runs. A second analysis investigated whether self-report ratings predicted the number of check-ins for a specific topic. For example, if a child provided high interest ratings for events related to chemistry, would that child show higher levels of check-ins related to chemistry compared to a topic with lower interest ratings? A series of bivariate correlations compared the number of check-ins by topic to interest, engagement, goal, and identity ratings from parents and children. No significant correlations were found. A third analysis compared ratings to check-in rates regardless of topic. Increases in interest and engagement ratings were positively correlated with increases in the number of check-ins, $r = .17, n = 280$, and $p = .003$ and $r = .19, n = 280$, and $p = .002$, respectively. No other correlations were significant.

Post-study survey

We conducted a brief survey of parents after the 2-month study period had ended. In this survey, we asked four questions in total, with three evaluating specific features of the app (results not reported here) and one about user experience (“Tell us three features about the app that you liked”). For the user experience question, seventy-five percent of parents indicated that the app helped them remember STEM experiences (e.g., “It was fun to see all the fun activities we did this summer!”) and 90% indicated that using the app encouraged them to seek more STEM experiences (e.g., “The kids wanted to do science things so they could report it in the app”) than

usual and increased family communication about STEM (e.g., “[the app] encouraged brain-storming and helped my son realize some of his interests fell in the stem categories”; “We started talking about science on our family hikes”).

Discussion

Proof-of-concept and notable outcomes

The goal of the present study was to develop and field test a new instrument, the STEMwhere app, for measuring informal STEM activity and learning supports. In this field test, several outcomes were noteworthy and provide a proof-of-concept that the app holds promise for investigating informal STEM learning across contexts and time. First, most STEM activities occurred at home (72%), a finding consistent with other research on STEM engagement (e.g., Maltese et al., 2014). One possibility is that the family home provides more opportunity for engagement than other locations, which is fairly unsurprising because the family home is the default location for family learning opportunities. The majority of activities were hands-on (64%), with most of these activities occurring in the home (72%). Thus, as demonstrated here, the STEMwhere app can be used to map ecologies where informal STEM learning takes place; by sampling a larger number of participants in future research, the app could be used to identify hotspots of STEM activity across a larger geographical region.

Second, child interest and engagement ratings were high in all settings and activities (no significant differences across either), suggesting that high situational interest was relatively common during these activities. Parent engagement varied across activities, though not by settings. Many parent and child ratings were similar; for example, camps, libraries, podcasts, and web content were given high ratings. The results

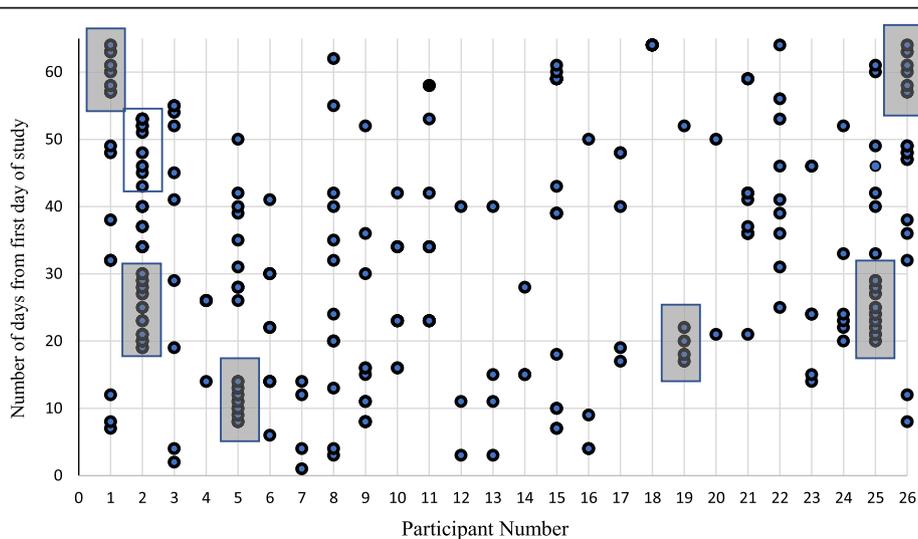


Fig. 1 Check-in rates for each participant. Shaded boxes indicate topical runs (check-ins shared topics) and unshaded boxes indicate runs (check-ins that did not share topics)

Table 4 Descriptions of topical run and run check-in patterns

Participant	Precursor of run	Theme of run
1	Family purchase of weather station	Weather-related activities including books, videos, and recording weather phenomena
2	Attended nature camp	Investigating animals and their habitats
2	Family vacation	None
5	Purchase of drone	Building and modifying drones
19	Medical internship	Health and medicine
25	Family vacation	Science of cooking and baking
26	Family purchase of weather station	Weather-related activities including books, videos, and recording weather phenomena

also demonstrated that parents and their children had some different ideas about STEM experiences. For example, children rated TV/movies as likely activities for future engagement and valuable for developing identity whereas their parents rated them as unlikely for future engagement and of limited utility for developing STEM identity. Although some activities might not fit traditional definitions of STEM by teachers or researchers, their designation as such by participants was meaningful and may help to illustrate how a variety of activities help to ignite STEM engagement.

Third, because STEMwhere can be used to collect data on multiple learning supports, relationships among them can be revealed. Our results are consistent with several findings from previous research. One, increases in interest predicted increased engagement (Thoman et al., 2017) and “fun” goal ratings but weakly related to identity ratings. This suggests that interest may drive later engagement and that a goal of having fun with friends is more closely related to interest than learning or other social goals. Engagement ratings were related to learning goals and identity, suggesting that moving beyond situational interest is important for developing content knowledge and positive STEM identity (Hidi & Renniger, 2006). The comparison between child and parent identity ratings is of interest because support from another person has been suggested as a critical factor in the formation of positive STEM identity (e.g., Calabrese Barton & Berchini, 2013). Increases in parent identity and engagement ratings were related to increases in child identity and engagement ratings, providing evidence that social support is critical as

children form positive STEM identities. This pattern of results suggests reasonable validity for the brief measures due to their fidelity with existing results.

Finally, the behavioral data from the app allowed us to create a novel measure—a topical run. One long-standing issue in education is how to measure increasing activity from its inception. By collecting participant-generated check-ins, we identified periods of increasing activity and their likely triggers. We operationally defined a run as a pattern of check-ins that were unlikely to occur by chance and shared a topic or location. We identified six runs that fit these criteria. These topical runs illustrated an increase in activity that differs from baseline behavior. The runs included a variety of locations and activities, for example, participant 26 checked in at a library, home, parks, and nature centers to engage with weather-related content. The pattern not categorized as a topical run provides a useful contrast in which there were increases in the frequency of check-ins that were statistically unlikely, yet the check-ins did not share a common topic. In this case, the family was on vacation and child was engaging in STEM-related activities that did not share a common topic. It is possible that this pattern was sustained (at least primarily) by the family, whereas those coded as topical runs might have been sustained by the child and the family. The analysis of run precursors and thematic patterns during the topical runs (Table 4) provide some evidence for this suggestion. Specifically, all runs were related to family activities; however, those classified as topical runs might demonstrate child-initiated activities, such as seeking books or videos about drones. In this way, family engagement is perhaps a

Table 5 Paired *t* test results comparing self-report and run data

Factor	Non-run means		Run means		95% CI	<i>t</i> score (<i>p</i> value)
	Mean	SD	Mean	SD		
Interest	8.33	1.99	8.53	2.3	.89, 1.29	-.88 (.18)
Engagement	7.23	2.34	7.77	2.8	.46, 1.25	1.24 (.10)
Goal fun	7.8	2.62	7.64	3.02	0.55, 0.83	1.03 (.15)
Learning goal	6.95	2.69	7.11	3.54	-.73, .16	.78 (.21)
Goal social	5.81	3.17	5.39	4.3	-.56, .31	1.18 (.11)
Identity	6.10	3.03	6.41	3.92	.23, 1.11	-.55 (.29)

necessary (though not sufficient) factor in long-term learning and engagement. Families who provide many opportunities are more likely to find topics or activities that engage their children, though it is difficult at the outset to predict which of these activities will ignite a child's interest.

Limitations and future directions

Although the aforementioned outcomes were based on a large number of observations per participant, there are several limitations. As noted in the introduction, the choice of a relatively brief measure of experience sampling has inherent limitations, most notably a reduction in psychometric validity as would be found in extensive surveys. The STEMwhere app allowed us to track participation over time due to the low amount of effort to complete measures at each observation. Our approach yielded a large number of observations from a series of low-cost measures rather than from one, high-cost instrument. Although we used a single item to measure each construct, the results were highly consistent with previous findings. For example, interest predicted unique variance for engagement as in previous research (Thoman et al., 2017). Perhaps most important, the STEMwhere app can be adapted by individual users, such as by adding more questions of a key construct or including questions to tap a construct not included in our original program.

Another limitation is allowing participants to define what constitutes a STEM activity. As noted above, approximately 16% of activities were excluded after coding because they were evaluated as being unlikely to constitute a true STEM experience. One notable example was one participant who checked in multiple times after playing Pokémon Go! On its own, playing this app has negligible STEM value. Nevertheless, defining STEM for participants would produce a different limitation, notably that researchers may miss activities that lead to later engagement (Maltese et al., 2014). This motivates a deeper question about the nature of what constitutes a STEM activity. Nearly any activity has the potential to be linked to some type of STEM activity. For example, the art of cooking itself might not be considered a STEM activity. However, a parent who asks her child why one salad dressing separates while another does not might set up experiments (e.g., does it separate after we add mayonnaise?) to discover the concept of an emulsifier. In this case, this cooking activity would be classified as a STEM activity. Thus, it is difficult to accurately measure the extent to which an activity is truly a STEM activity without obtaining more information about the nature of the activity and the opportunities for learning provided by it.

Another limitation of the present proof-of-concept study was that the observations were from a relatively small sample of participants. Though intriguing, the data should be generalized cautiously, particularly the topical runs, which are based on a small number of observations. An exciting avenue for future research would be to pool data collected using the app

so as to create larger data sets that can provide the basis for secondary analyses (analogous to the CHILDES database for language research, MacWhinney, 2008). We developed the STEMwhere app with this use in mind; in particular, the app itself has several hard-wired questions about demographics and the learning supports, so as to ensure that any investigator using the app will contribute outcomes relevant to core issues pertaining to informal STEM learning. And, as implied above, the most recent version of the STEMwhere app (which is available free from the first author) has also been developed so that an investigative team can include unique questions relevant to their specific research questions and goals.

The STEMwhere app was developed as a research tool to measure informal STEM learning, and, of course, any measure may influence the outcomes one is attempting to measure. Such reactive effects are typically viewed as a limitation. In formal and informal educational settings, any *positive* reactive effects may offset the limitation due to measurement reactivity per se. For example, in the present case, another encouraging result is the apparent effect of app use on family STEM engagement. In particular, the results of the survey suggest that the app encouraged parents to lead activities and to increase their communication about STEM content. These results suggest an opportunity to use the STEMwhere app to support parents as they guide their children's STEM activities outside of school. An example of such an opportunity is a recent study in which an app (Bedtime Math) was used to help parents increase their math and number talk with their preschool aged children (Berkowitz et al., 2015). The children of math-anxious parents who used Bedtime Math showed significant gains in math performance compared to children in a control group. In these cases, the fact that an app can promote informal STEM learning (as well as measure learning supports) is encouraging and should be investigated in future research.

Conclusions

The STEMwhere app was used to measure four supports for informal science learning (interest, engagement, identity, and learning goals) for 2 months by 26 children ages 7–14. During this time, the participants checked in during STEM activities and answered eight questions about each activity. Although preliminary, the results revealed that most STEM activities were hands-on activities that occurred in the family home and that there are specific relations between learning supports (e.g., increased interest was related to increased engagement). Finally, check-in rates yielded a new measure, a topical run, that provides new evidence for the conditions under which children become engaged and measured increasing activity over time. These results serve as a proof-of-concept for the STEMwhere app for measuring informal STEM learning in the wild by providing novel findings that contribute to our understanding of where and how informal STEM learning occurs.

Appendix

Table 6 STEMwhere questions and scales

Concept	Statement	Not at all interested	Somewhat interested	Partially interested	Interested	Extremely interested
Interest	How interested are you in this activity?	Not at all interested	Somewhat interested	Partially interested	Interested	Extremely interested
Engagement	I would like to learn more about this activity	Strongly disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree
Fun goals	I am doing this activity because I want to have some fun.	Strongly disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree
Learning goals	I am doing this activity because I want to learn something.	Strongly disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree
Social goals	I am doing this activity because I want to be with friends/family	Strongly disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree
Identity	I can imagine myself being a part of this field someday.	Strongly disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree
Parent interest	I am interested in finding more opportunities like this for my child.	Not at all interested	Somewhat interested	Partially interested	Interested	Extremely interested
Parent identity	I can imagine my child being a part of this field someday.	Strongly disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree

Abbreviation

STEM: Science, Technology, Education, and Mathematics

Acknowledgements

The authors thank the families who participated in this study and Ida Cellitti for her assistance in organizing the study.

Authors' contributions

BJM, WO, SE, KE, and JD designed the studies and drafted the manuscript. BJM and JD generated the idea and collected the data, and BJM conducted the analyses. All the authors approved the final manuscript for submission. Please address correspondence to BJM.

Funding

This research was supported by a grant from the US National Science Foundation (Award 1451284). The content is solely the responsibility of the authors and does not represent the views of the National Science Foundation.

Availability of data and materials

The datasets used in the current study and STEMwhere app are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Educational Psychology, Kent State University, 405 White Hall, Kent, OH 44242, USA. ²Cincinnati Museum Center, Cincinnati, USA. ³Great Lakes Science Center, Cleveland, USA. ⁴Department of Education, University of Oxford, Oxford, UK. ⁵Department of Psychological Sciences, Kent State University, Kent, USA.

Received: 14 May 2019 Accepted: 18 October 2019

Published online: 26 November 2019

References

- Alliger, G. M., & Williams, K. J. (1993). Using signal-contingent experience sampling methodology to study work in the field: A discussion and illustration examining task perceptions and mood. *Personnel Psychology*, 46(3), 525–549.
- Barton, A. C., & Tan, E. (2010). We be burnin'! Agency, identity, and science learning. *The Journal of the Learning Sciences*, 19(2), 187–229.
- Berkowitz, T., Schaeffer, M. W., Maloney, E. A., Peterson, L., Gregor, C., Levine, S. C., & Beilock, S. L. (2015). Math at home adds up to achievement in school. *Science*, 350, 196–198.
- Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, 64, 417–444.
- Bradburn, N. M., Rips, L. J., & Shevell, S. K. (1987). Answering autobiographical questions: The impact of memory and inference on surveys. *Science*, 236(4798), 157–161.
- Calabrese Barton, A., & Berchini, C. (2013). Becoming an insider: Teaching science in urban settings. *Theory Into Practice*, 52(1), 21–27.
- Dorph, R., Schunn, C. D., & Crowley, K. (2017). Crumpled molecules and edible plastic: Science learning activation in out-of-school time. *Afterschool Matters*, 25, 18–28.
- Dunlosky, J., & Metcalfe, J. (2008). *Metacognition*. Sage Publications.
- Eberbach, C., & Crowley, K. (2009). From everyday to scientific observation: How children learn to observe the biologist's world. *Review of Educational Research*, 79(1), 39–68.
- Eberbach, C., & Crowley, K. (2017). From seeing to observing: How parents and children learn to see science in a botanical garden. *Journal of the Learning Sciences*, 26(4), 608–642.
- Erdogan, N., & Stuessy, C. L. (2015). Modeling successful STEM high schools in the United States: An ecology framework. *International Journal of Education in Mathematics, Science and Technology*, 3(1), 77–92.
- Eshach, H. (2007). Bridging in-school and out-of-school learning: Formal, non-formal, and informal education. *Journal of Science Education and Technology*, 16(2), 171–190.
- Falk, J. H., & Dierking, L. D. (2010). The 95 percent solution. *American Scientist*, 98, 486–493.
- Falk, J. H., Storksdieck, M., & Dierking, L. D. (2007). Investigating public science interest and understanding: Evidence for the importance of free-choice learning. *Public Understanding of Science*, 16(4), 455–469.
- Farmer, V. L., & Shepherd-Wynn, E. (2012). *Voices of Historical and Contemporary Black American Pioneers*, 1. ABC-CLIO, 304.
- Fenechel, M., & Schweingruber, H. A. (2010). *Surrounded by Science: Learning Science in Informal Environments*. National Academies Press.
- Fredricks, J. A., Wang, M. T., Linn, J. S., Hofkens, T. L., Sung, H., Parr, A., & Allerton, J. (2016). Using qualitative methods to develop a survey measure of math and science engagement. *Learning and Instruction*, 43, 5–15.
- Gonzalez, H. B., & Kuenzi, J. J. (2012). *Science, Technology, Engineering, and Mathematics (STEM) Education: A Primer*. CRS Report No. R42642. Retrieved from Congressional Research Service website: <https://fas.org/spp/crs/misc/R42642.pdf>
- Greene, B. A., & Miller, R. B. (1996). Influences on achievement: Goals, perceived ability, and cognitive engagement. *Contemporary Educational Psychology*, 21(2), 181–192.

- Grønmo, L. S., Lindquist, M., Arora, A., & Mullis, I. V. (2015). TIMSS 2015 mathematics framework. TIMSS, 11–27.
- Heth, C. D., & Cornell, E. H. (1987). Monte Carlo Simulation as a Method of Identifying Properties of Behavioral Organization. In *Formal Methods in Developmental Psychology* (pp. 372–398). Springer New York.
- Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, 41(2), 111–127.
- Hughes, R. M., Nzekwe, B., & Molyneux, K. J. (2013). The single sex debate for girls in science: A comparison between two informal science programs on middle school students' STEM identity formation. *Research in Science Education*, 43(5), 1979–2007.
- Jones, A. L., & Stapleton, M. K. (2017). 1.2 million kids and counting—Mobile science laboratories drive student interest in STEM. *PLoS biology*, 15(5), e2001692.
- Kelley, T. R., & Knowles, J. G. (2016). A conceptual framework for integrated STEM education. *International Journal of STEM Education*, 3(1), 11.
- Larson, R. (1989). Beeping children and adolescents: A method for studying time use and daily experience. *Journal of Youth and Adolescence*, 18(6), 511–530.
- Lin-Siegler, X., Ahn, J. N., Chen, J., Fang, F. F. A., & Luna-Lucero, M. (2016). Even Einstein struggled: Effects of learning about great scientists' struggles on high school students' motivation to learn science. *Journal of Educational Psychology*, 108(3), 314–328.
- MacWhinney, B. (2008). Enriching CHILDES for morphosyntactic analysis. In H. Behrens (Ed.), *Trends in corpus research: Finding structure in data*. Amsterdam: Benjamins.
- Maltese, A. V., Melki, C. S., & Wiebke, H. L. (2014). The nature of experiences responsible for the generation and maintenance of interest in STEM. *Science Education*, 98(6), 937–962.
- Milne, C., & Otieno, T. (2007). Understanding engagement: Science demonstrations and emotional energy. *Science Education*, 91(4), 523–553.
- Morris, B. J., Croker, S., Masnick, A. M., & Zimmerman, C. (2012). The emergence of scientific reasoning. In H. Kloos, B. J. Morris, & J. L. Amaral (Eds.), *Current topics in children's learning and cognition* (pp. 61–82). Rijeka, Croatia: InTech.
- Morris, B. J., Masnick, A., Baker, K., & Junglen, A. (2015). Taking data to class: An analysis of data-related activities in middle-school science textbooks. *International Journal of Science Education*, 37(16), 2708–2720.
- Nasir, N. I. S. (2008). Everyday pedagogy: Lessons from basketball, track, and dominoes. *Phi Delta Kappan*, 89(7), 529–532.
- National Academies of Sciences, Engineering, and Medicine. (2016). *Parenting Matters: Supporting Parents of Children Ages 0-8*. Washington, DC: The National Academies Press.
- National Academies of Sciences, Engineering, and Medicine. (2016a). *Barriers and Opportunities for 2-Year and 4-Year STEM Degrees: Systemic Change to Support Students' Diverse Pathways*. Washington, DC: The National Academies Press.
- National Academies of Sciences, Engineering, and Medicine. (2016b). *Developing a National STEM Workforce Strategy: A Workshop Summary*. Washington, DC: The National Academies Press.
- National Academies of Sciences, Engineering, and Medicine. (2018). *Graduate STEM Education for the 21st Century*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/25038>.
- National Science Teachers Association. (2012). NSTA position statement: Learning science in informal environments. <http://www.nsta.org/about/positions/informal.aspx>. Retrieved December 11, 2017.
- Piasecki, T. M., Hufford, M. R., Solhan, M., & Trull, T. J. (2007). Assessing clients in their natural environments with electronic diaries: Rationale, benefits, limitations, and barriers. *Psychological assessment*, 19(1), 25.
- Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. *Handbook of self-regulation*, 451, 451–502.
- Provasnik, S., Malley, L., Stephens, M., Landeros, K., Perkins, R., & Tang, J. H. (2016). Highlights from TIMSS and TIMSS advanced 2015: Mathematics and science achievement of US students in grades 4 and 8 and in advanced courses at the end of high school in an international context (NCES 2017–002). US Department of Education. National Center for Education Statistics. Washington, DC. Retrieved from <http://nces.ed.gov/pubsearch>.
- Renninger, K. A., & Bachrach, J. E. (2015). Studying triggers for interest and engagement using observational methods. *Educational Psychologist*, 50(1), 58–69.
- Renninger, K. A., & Hidi, S. (2015). *The power of interest for motivation and engagement*. Routledge.
- Sacco, K., Falk, J. H., & Bell, J. (2014). Informal science education: Lifelong, life-wide, life-deep. *PLoS Biology*, 12(11), 1307.
- Saxe, G. B. (1988). Candy selling and math learning. *Educational Researcher*, 17(6), 14–21.
- Stevens, R., Satwicz, T., & McCarthy, L. (2008). In-game, in-room, in-world: Reconnecting video game play to the rest of kids' lives. *The ecology of games: Connecting youth, games, and learning*, 9, 41–66.
- Stone, A.A., J.E. Schwartz, J.M. Neale, S. Shiffman, C.A. Marco, M. Hickcox, J. Paty, L.S. Porter and L.J. Curise. (1998), A comparison of coping assessed by ecological momentary assessment and retrospective recall. *Journal of Personality and Social Psychology*, 74, pp. 1670–1680.
- Thoman, D. B., Sansone, C., & Geerling, D. (2017). The dynamic nature of interest: Embedding interest within self-regulation. In P. A. O'Keefe & J. M. Harackiewicz (Eds.), *The Science of Interest* (pp. 27–47). Springer.
- Wang, M. T., Fredricks, J. A., Ye, F., Hofkens, T. L., & Linn, J. S. (2016). The math and science engagement scales: Scale development, validation, and psychometric properties. *Learning and Instruction*, 43, 16–26.
- Winne, P. H. (2001). Self-regulated learning viewed from models of information processing. In B. J. Zimmerman & D. H. Schunk (Eds.), (2nd Edition). *Self-regulated learning and academic achievement: Theoretical perspectives* (pp. 153–189). Mahwah, NJ: Erlbaum.
- Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in Educational Theory and Practice* (pp. 277–304). Hillsdale, NJ: LEA.
- Winne, P. H., Hadwin, A., Schunk, D., & Zimmerman, B. (2008). Motivation and self-regulated learning: Theory, research, and applications.
- Zimmerman, B. J., & Schunk, D. H. (2001). (2nd Edition). *Self-Regulated Learning and Academic Achievement: Theoretical Perspectives* (pp. 289–308), Mahwah, NJ: Erlbaum.
- Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Developmental review*, 27(2), 172–223.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com