Assessing interest and confidence as components of student motivation in informal STEM learning

Mark Blair
Computer and Information Science
Gannon University
Erie, PA USA
blair012@gannon.edu

Stephen Frezza
Computer and Information Science
Gannon University
Erie, PA USA
frezza001@gannon.edu

***Abstract*— This Work In Progress submission focuses on the deployment of assessment metrics of informal Science, Technology, Engineering and Mathematics (STEM) education conducted in various pre-college settings. Informal STEM Learning (ISL) is "learning [that] occurs across the lifespan and in places and spaces beyond schools or the school day" -NSF. NSF guidance recommends --among others-- the measure of the impact category of interest or engagement. The authors ask what is interest? What are the measures of its success? And how do we measure them? The authors theorize that interest as an emotional experience alone is not enough to compel learning behaviors and compose a composite model of interest and confidence. The authors suggest the psychological concepts of "motivation" and "self-efficacy" as a framework to devise measurements of interest/confidence. The authors propose an analytic model using QR codes to tie survey responses to a target behavior, "reciprocal engagement" as a new mechanism for verifying continued learning behavior. Initial data collection was interrupted by the Covid-19 lock down; however, the partial data body collected showed indications that self-reported data about learners' experience in a lesson is more predictive of continued learning behavior than self-evaluation items such as "I understood the topic" and "I am familiar with the topic."**

***Keywords*— STEM Education Assessment; Student Engagement; Confidence; Interest; Robotics Activity; Teacher Engagement; STEM Resources**

I. Introduction

On a warm May afternoon Mark Blair stood in front of the bleachers on the Wilson Middle School Soccer field. High pitch screams periodically drowned out his words as racing drones zipped behind him at speeds in excess of one hundred miles an hour. The only thing louder than the drones are the excited shrieks of children watching a cool STEM technology demonstration. Afterwards a couple of students approached with excitement and asked, "How can we get some of these drones?" When they learned of the $500 price they exclaimed, "Man, screw all that mess! [edited]." Had we just damaged these students' motivation toward STEM? While we had observed an increase in short-term interest, had we done so at the expense of a learner's long-term motivation?

Currently, few proclamations are as ubiquitous as "the world needs more engineers." This assertion is not without its detractors [1]. On the current trajectory, STEM skills will have needed no more justification than did typing skills in the 1980s and 90s. Informal STEM Learning (ISL) activities aim to generate interest in STEM [2], but are we asking what is interest? Are we asking what is the result of generating interest in young learners or are we assuming its end? It is easy to state we need to increase interest in STEM learning, but "Interest" is conceivably only valuable if it results in positive behavioral or attitudinal changes. From a research perspective, this opens the question "Is interest a discrete entity or is it part of a more generalized psychological component?"

The first challenge is to collect data anonymously while aligning a single respondent with multiple responses and behavioral interactions over time. The authors theorized that confidence also plays a critical antecedental role in learning behavior. If competence is the consequence of ISL, interest and confidence--or one's belief about one's competence--are antecedents of STEM learning. Our goal is to assess self-directed STEM learning behaviors in our students during and following the event(s). The research question is: H1: Confidence and interest in novice STEM learners combine to form a threshold for engagement in learning.

II. Interest and Confidence

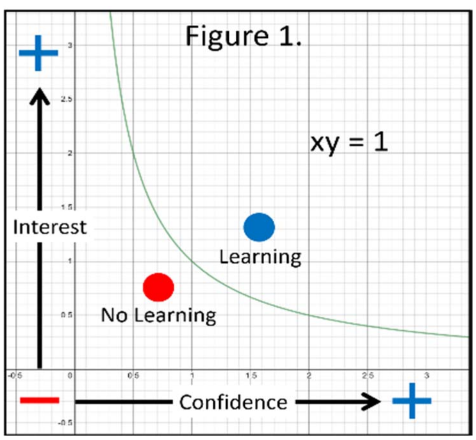
As identified above if interest and confidence are antecedents of STEM learning, these two antecedents must be discerned from each other. Additionally, each of these antecedents must be defined based on an empirically founded body of knowledge in order to build an effective research model. The problem is, interest does not have a discrete psychological definition. Psychological studies may use the term interest when affinity is the characteristic quality, or when attention/focus is the measured quantity, when curiosity is the characteristic quality, or how something affects someone is the measured quantity (as in: "in one's best interest"). "Theories of interest split into two fields: (1) interest as a part of emotional experience, curiosity, and momentary motivation; and (2) interest as a part of personality, individual differences, and people's idiosyncratic hobbies, goals, and avocations" [3]. Paul Silivia describes this dichotomy as the difference between "interest" and "interests" [3]. In psychology the body of research regarding motivation best matches the desired ends of "interest" as it applies to ISL. Motivation is defined as a need, desire, [or aversion] that serves to energize behavior and to direct it toward a goal [4].

Our working definition for confidence is "one's beliefs about one's competency." The difficulty is that there is no discrete psychological definition for confidence. Psychological studies use the term confidence to refer to the amount of certainty a person feels regarding a judgement, or as it relates to self-esteem. Albert Bandura's 1977 cognitive-behavioral theory of self-efficacy is the belief that one can master and perform needed behaviors when necessary. "An efficacy expectation is the conviction that one can successfully execute the behavior required to produce the outcomes" [5]. Research of self-efficacy has directly addressed STEM educational issues. "Researchers have reported that the mathematics self-efficacy of college undergraduates is more predictive of their mathematics interest and choice of math-related courses and majors than either their prior math achievement or math outcome expectations" [6] [7] [8] [9] [10] [11] [12]. Our study views this as a diathesis: assessing confidence and its relationship to student interest with regard to engagement.

A. *Two-Factor Diathesis Threshold Model*

Diathesis thresholds are a way of modeling the interaction between two factors that define a person's predisposition to a specific behavior. This is useful for modeling two factors defined by a function. For instance, imagine data revealed the relationship between interest and confidence, represented in Figure 1 by xy = 1 where x and y are greater than 0. With this model it is easy to see that if someone has low confidence, they must typically be very interested in order to be motivated enough to engage in learning. Whereas someone who is very confident needs much less interest to be motivated to learn.

Figure 1. Finding Learning from Interest and Confidence



These interactions imply that there are better approaches to developing ISL objectives; that in fact for effective learning to occur that neither confidence nor interest can be built independently, and that too much emphasis on one will/may have a detrimental effect on another such as the "overjustification effect" [13] [14]. This remains a theory but suggests strongly that assessment of motivational components such as interest and confidence would be justifiably valuable to ISE educators.

III. Preliminary Study

Within the proposed framework the authors have designed a parsimonious and pragmatic first attempt to collect data in vivo in a learning environment. The curriculum development team developed a survey to test and measure interest and confidence in students. The design incorporated self-reported surveys and a mechanism to detect reciprocal engagement. Data were not collected in an informal STEM learning environment; instead it was collected from high school students taking curricula designed for informal STEM learners. This allowed for the collection of daily surveys. The curriculum covered several areas of STEM including programming, electrical engineering, Computer Aided Design (CAD), and 3D printing. The data collection framework was designed to capture data impressions from at least a few examples of respondents who are interested and/or competent in one subject while not interested and/or competent in another.

A. *Preliminary Survey*

Student respondents in ISL settings were given anonymized QR codes which allowed them to access the daily survey and other resources via the web. Respondents using their QR code to access further educational materials are considered to be exhibiting reciprocal engagement. Reciprocal engagement is a key target behavior identified by the authors as a positive STEM learning behavior.

All questions used a common 7-point Likert scale: *Strongly disagree, Disagree, Somewhat disagree, Neither agree nor disagree, Somewhat agree, Agree*, and *Strongly agree*. Survey questions were broken down into two sections self-evaluation (self-eval) and experiential. Self-eval characteristics include two items (each item is given a short title in parentheses):

S1: (understand) I understood the content of today's topic.

S2: (familiar) I was familiar with the topic that was taught beforehand.

Experiential questions asked students about their experience during the lesson that was taught.

E1: (enough time) It felt like we did not have enough time to go over today's topic.

E2: (confused) I was confused and could not follow through the assessment.

Finally, the survey included two target questions.

T1: (interest) Today's topic was really interesting.

T2: (learn more) I am interested in learning about this topic more in the future.

Questions were presented in the following order: S1, T1, S3, T2, E1, E2. Experiential questions were inverted. E.g., *strongly agree* was associated with confusion in E2, where *strongly agree* was associated with understanding in S1. This was done to ensure thoughtful response.

B. *The Experiential Dichotomy*

The survey questions aimed to elicit the respondent's beliefs about themselves in relations to the periods of instruction. The two items (S1 and S3) were phrased in the first person: "I understood the content…" and "I was familiar with the topic…" Experiential items (E1 and E2) elicit the respondent's experience. Where self-eval items ask a respondent what they think of themselves, experiential items ask, "how does this make you feel?" Experiential items are phrased as emotional expressions: “I was confused…” or “It felt like we did not have enough time…” Target items cover an emotion-behavior dichotomy. T1 asks for an emotional response: the “topic was really interesting.” T2 asks for behavioral planning: “I am interested in learning… more in the future.” The analytic model is designed to determine how target items align with self-eval and experiential data.

C. *Event Data for Tracking Interest*

A key innovation we are piloting is the notion of providing students with an anonymized identifier to allow event tracking, e.g., that as students leverage on-line resources, the unique code allows tracking that some ISL participant accessed a particular ISL resource. The T2 (learn more) item on the survey indicates whether a respondent intends to engage in further learning behavior but cannot indicate whether the respondent truly carries out this intent. The access, provided by the QR code was intended to be identified as independent or not, but this did not work as intended. The design allows for no personal or identifying information besides gender to be tracked but allows each access of a resource to be associated with daily surveys.

IV. DATA

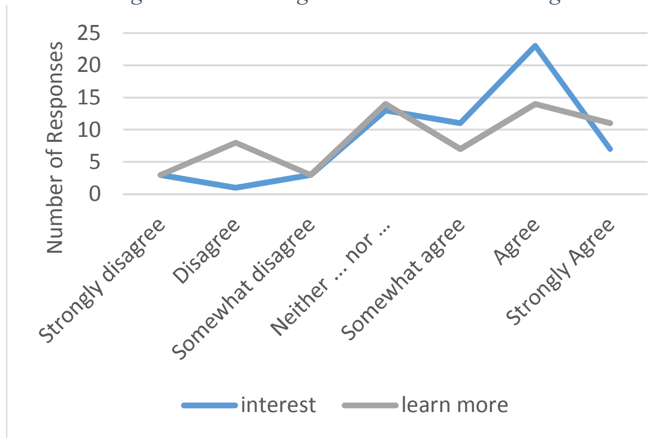
Experiential data provided the respondent an opportunity to give an honest evaluation even if it may be threatening to the respondent’s self-esteem. The 65 responses provided two types of data (self-eval and experiential) collected showed two different data shapes. However, the trends and characteristics appeared interesting. The goal of generating interest in STEM training is assumedly to encourage students to learn more STEM skills. Given that the two types of data demonstrated two different data shapes the question becomes which data shape best predicts the learners desire to learn more STEM.

Consistent effects across all responses to E1(not enough time) and E2(confused) will be considered an accurate evaluation of the period of instructions. In other words, consistent agreement with E1 and E2 could indicate a flaw in the instruction. If no consistent effect is present in the data, self-eval responses should align with experiential responses. For instance, if a respondent states they “agree” with S1(understand) the respondent should state they “disagree” with E2 (confused). If a respondent’s S1 and E2 do not align, it could mean that there is a difference between a respondent’s beliefs about their competency and their actual competencies. While, this last dynamic may be useful for determining the effectiveness of the curriculum and instruction methods, the real objective is to see how the target items align with self-eval and experiential items.

A. *“Interesting” and “Interested”*

What does it mean to be “interesting?” How do “interesting” responses differ from “interested in learning more?” Figure 2 shows the initial findings. Survey item T1: “Today’s topic was really interesting” elicits an emotional response to the subject matter, where T2: “I am interested in learning about this topic more in the future” indicates behavioral planning. While T1 is an easy positive response and conceivably does not require a positive self-efficacy, a positive response to T2 does. T1 showed a much stronger left-bias (toward Strongly agree). The difference in in these data shapes is due to the number of respondents who reported they “disagree” (T1: 1, T2: 8) and “agree” (T1: 23, T2: 14).

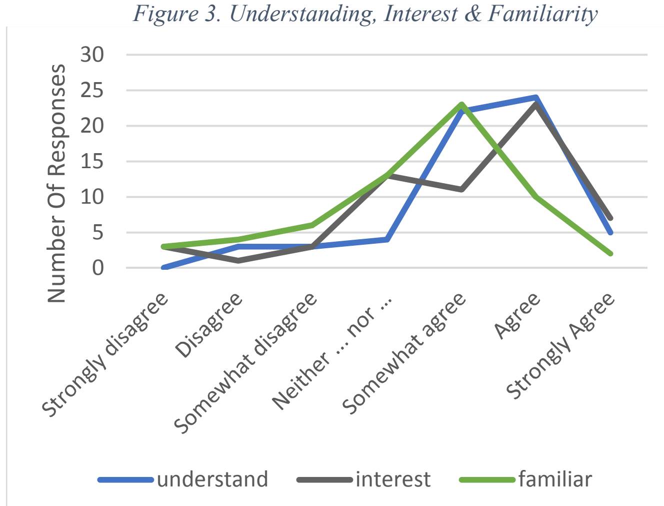
Figure 2. Interesting vs. Interested in Learning More



B. *Self-Evaluation Data*

The self-reported data is presented in Figure 3 and demonstrates a left-biased (toward “Strongly Agree”). The self-eval item “I understood the content of today’s topic” seemed to be characterized by T1 and S3, which indicate “interesting” and a familiarity with the subject matter of the lesson.

Figure 3. Understanding, Interest & Familiarity

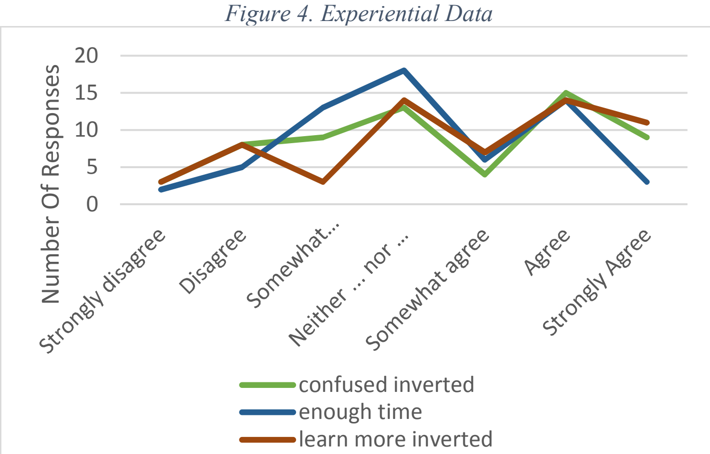


It is not surprising that familiarity predicts understanding. However, the data may indicate that interest and understand are related. These results may also suggest that respondents must understand the topic in order to find it “interesting” (T1).

C. *Experiential Data*

The experiential data (Figure 4.) showed a dual peak with a strong representation in the “Neither agree nor disagree” and “agree” responses. Experiential data was inverted to enhance data fidelity, exposing responses that have the same option for each item, and also describe an experience of academic frustration (graph shows responses inverted for comparison). Experiential responses directly address the learners experience during the period of instruction. Item E2 inquires whether learners were confused and could follow the lesson. The phrase “…could not follow” directly addresses a feeling of negative self-efficacy. Item E1 inquires if the entire class ("we") "did not have enough time" to cover the topic. Again, the item E1 shows experience over respondents' beliefs about themselves. In a perfect world data indicating understanding and the instruction tempo should align. Meaning, students should have the proper prerequisite knowledge and curriculum developers and instructors should teach at a proper pace. Whether a defect in instruction is due to lack of prerequisite knowledge or failure of instruction is not important at this stage. What is important is that experiential data is a better predictor of whether a person intends to learn more, than self-eval data.

Figure 4. Experiential Data



D. Data Interpretation

These preliminary results suggest that experiential data is more predictive of continued learning behavior than self-eval data. The data may indicate that interest as an emotion may not be an effective method to encourage STEM learning behavior. The failures and short comings of the initial study does not provide scientific certainty but does indicate a research model that may produce an extremely useful body of knowledge for developing, executing, and evaluating ISL events.

V. Mishaps and Failures

The data collected in the initial study had significant issues. While significant, these are issues we hope to overcome in the repeat attempts planned for Fall 2020. For example, four responses were eliminated due to the fact that respondents gave the same option for each question, e.g. they responded with *Strongly Agree* to every question.

A. QR Code Assignment

QR codes were printed on business cards and distributed to students. The class' teacher –worried students would lose them– collected the cards at the end of class in one section without marking the cards so they could be redistributed. This disjointed the entry surveys from later surveys in one class.

B. Survey Timing

Students were immediately released at the bell that ended class. The teacher found it difficult to complete a lesson and save enough time for students to take their daily surveys. He decided to begin each class by having the students complete the survey for the lesson given the day before. The effect of this delay before respondents take the survey is unknown.

C. Covid-19 Lockdown

Data collection began in mid-February 2020. In mid-March the Covid-19 national lockdown began, interrupting data collection efforts. Some lessons including programming, CAD, and 3D printing were conducted, but no lessons in physics, electrical or mechanical engineering were given. The interruption in data collection inhibited the ability of the studied to capture examples of respondents who are interested and/or competent in one subject while not interested and/or competent in another.

These issues affected the initial study as the data was not complete or robust. Due to the interruption in data collection and several mishaps during data collection, the data was not capable of providing scientific certainty.

VI. Summary and Future Work

The key innovation we are piloting is the notion of providing students with an anonymized identifier to allow event tracking, e.g., that as students leverage on-line resources, the unique code allows tracking that some ISL participant accessed a particular ISL resource. This plan is for this access to be identified as independent or not, where independent resource access indicates a continued or increase in student interest possibly linked to the ISL event where they received their resource tracker. These codes will be used to track respondents throughout the curriculum to provide discrete antecedents to learning behavior connecting both confidence and interest.

Future work will build on existing empirically founded research frameworks including research in motivation and self-efficacy. These two bodies of research cover behaviorism, cognitive-behavioral psychology, and social-cognitive theory giving the derive research a robust foundation with multiple faceted perspectives for analysis. Future research will include behavior tracking data to confirm self-reported survey items. To help mitigate other shortfalls experienced in the initial study, future work will incorporate gamification to provide other motivations for respondents to engage in surveys and other data collection efforts. Additionally, surveys will be delivered in the form of quizzes serving three purposes: 1) augment the lessons in the curriculum; 2) evaluate the knowledge learned in the lesson and 3) collect the same self-eval and experiential data.

Prototypical ISL sessions may not have the same set of students again. The authors proposed modeling takes a first step to detected desired behavioral objectives and capture it in data. The hope is that better-implemented data collection techniques will paint a more granular, diathesis picture of the interest and confidence generated during ISL. In the future, more longitudinal data may be gathered by utilizing open-source communities of developers and/or the potential use of federated identity systems within the open-source community to observe learning behaviors over a learner's development.

VII. Aknowledgements

The authors would like to thank Tejita Rajbahandari, Sinza Shrestha, Carley Sawyer, Katelyn Gourley and Sydney Oyatta for their work on the STEM Bot and assistance in data collection as well as Monaliben Patel for her help in background research.

References

[1] A. Hacker, The Math Myth: and Other STEM Delusions, York: The New Press, 2016.

[2] A. Friedman, "Framework for evaluating informal scienc education projetcs," 12 March 2008. [Online]. Available: http://informalscience.org/documents/Eval_Framework.p [Accessed 13 January 2020].

[3] P. J. Silvia, Exploring the Psychology of Interest, New Y Oxford University Press, Inc., 2006.

[4] D. G. Myers, Psychology, New York: Worth Publishers,

[5] A. Bandura, "Self-Efficacy: Toward a Unifying Theory o Behavioral Change," *Psychology Review,* vol. 84, no. 2, p 191-215, 1977.

[6] F. Pajares, "Self-Efficacy Beliefs in Academic Settings," *Review of Educational Research,* vol. 66, no. 4, pp. 543-5 1996.

[7] G. Hackett, "The role of mathematics self-efficacy in the choice of math-related majors of college women and men path analysis," *Journal of Counciling Psychology,* vol. 32 47-56, 1985.

[8] G. Hackett and N. E. Betz, "An exploration of the mather self-efficacy/ mathematics performance correspondence," *Journal for Research in Mathematics Educations,* vol. 20 261-273, 1989.

[9] R. W. Lent, F. G. Lopez and K. J. Bieschke, "Mathematic self-efficacy: Sources and relation to science-based caree choice," *Journal of Counseling Psychology,* vol. 38, pp. 4 430, 1991.

[10] R. W. Lent, F. G. Lopez and K. J. Bieschke, "Predicting mathematics-relate choice and success behaviors: Test of expanded social cognitive model," *Journal of Vocational Behavior,* vol. 42, pp. 223-236, 1993.

[11] F. Pajares and M. D. Miller, "The role of self-efficacy an concept beliefs in mathematical problem-solving: A path analysis," *Journal of Educational Psychology,* vol. 86, pp 193-203, 1994.

[12] F. Pajares and M. D. Miller, "Mathematics self-efficacy a mathematics outcomes: The need for specificity of assessment," *ournal of Counseling Psychology,* vol. 42, p 190-198, 1995.

[13] S. H. Tang and V. C. Hall, "The overjustification effect: meta-analysis," *Applied Cognitive Psychology,* vol. 9, pp. 404, 1995.

[14] M. R. Lepper, D. Greene and R. E. Nisbett, "Underminin chilfren's intrinsic interest with extrinsic reward: A test o overjustification hypothesis," *Journal of Personality and social Psychology,* vol. 28, no. 1, pp. 129-137, 1973.