Three critical issues that shape and complicate STEM self-efficacy intervention research: Reflections and analysis from an interdisciplinary research team

Stephanie J. Sedberry, Ian D. Beatty, and William J. Gerace
Department of Physics and Astronomy, UNC Greensboro, PO Box 26170, Greensboro, NC, 27402-6170

Maha A. Elobeid
Department of Educational Leadership and Cultural Foundations, UNC Greensboro, PO Box 26170, Greensboro, NC, 27402-6170

Jason E. Strickhouser
Department of Behavioral Sciences and Social Medicine, Florida State University, 1115 West Call Street Tallahassee, FL 32306-4300

A student’s academic self-efficacy is a variable that predicts student achievement and persistence in STEM, and substantial research has focused on developing and testing interventions to increase STEM self-efficacy. Results have been inconsistent: Some efforts produced desired outcomes while others show weak or null effects. What factors affect whether a self-efficacy intervention is successful? Based on our experiences with an NSF-funded project that developed and tested a classroom-based self-efficacy intervention in university Physics courses, we identify three critical research issues that shape and complicate STEM student self-efficacy research, ground them in the literature, and illustrate them in practice. They are: (1) defining and measuring self-efficacy, (2) accounting for context, and (3) understanding related psychosocial factors. We conclude with implications for future research.
I. INTRODUCTION

A strong STEM workforce is essential to economic growth and offers workers higher-than-average salaries. However, the demand for STEM professionals in the US is outpacing the supply of STEM-capable candidates we produce, and this shortage is expected to get worse in the coming years. Additionally, both racial and gender disparities in STEM employment are high: Hispanics, African Americans, and Native Americans make up 27% of the workforce, but only 11% of STEM workers, and men are employed in STEM occupations at twice the rate of women [1].

A. Psychosocial factors and self-efficacy

One strategy to address the shortage of qualified STEM candidates in general, and underrepresented groups specifically, is to attend to psychosocial factors (also referred to as social-psychological, non-cognitive, or affective factors) associated with student persistence and academic success in STEM fields. A convincing body of research indicates that psychosocial factors such as motivation, perceived control, self-efficacy, grit, and growth mindset can predict student academic performance, persistence, and STEM career success. A focus on psychosocial interventions for promoting student success in STEM classrooms has proven particularly effective for helping women and minorities [2, 3].

Self-efficacy, the focus of our research, is one of the most widely studied psychosocial constructs. Self-efficacy is the set of beliefs we have about our ability to successfully complete a particular task or goal, and it is among the most robust predictors of academic achievement. A significant body of research demonstrates that university students’ self-efficacy beliefs predict their selection of STEM as a major, persistence and academic performance in science, and aspirations for STEM-related careers [4, 5], and substantial research has focused on developing and testing interventions to increase student STEM self-efficacy [2, 6–8]. Increasing student self-efficacy to improve STEM learning is promising because self-efficacy predicts high academic achievement [9, 10].

B. The SIISP project

Concerned about high rates of STEM student attrition in our classrooms and across the US and motivated by this compelling literature, we developed and tested a 30-minute classroom intervention to increase student STEM self-efficacy, designed to be easily adopted in other university STEM classrooms. Our NSF-funded research project, Self-Efficacy Intervention to Improve STEM Performance (SIISP), is an in-progress quantitative study using a quasi-experimental design with a control treatment. We have completed:

1. Intervention Development: a 30-minute, portable and interactive classroom-based intervention with animated videos, participant booklets, and facilitator discussion prompts, based on both attributional retraining (locus of control) and growth mindset instruction;
2. Instrumentation Development: a 34-item Likert-scale questionnaire to measure STEM self-efficacy, growth mindset, and perceived academic control, validated through exploratory factor analysis, Rasch model analysis, and multi-method multi-trait analysis; and
3. Data Collection: three consecutive semesters of data collection including pre-testing, conducting intervention and control workshops, post-testing, and collecting final grades for 853 students enrolled in either algebra-or calculus-based Physics courses at three universities in the southeastern US.

We are presently in the data analysis and dissemination phase of the project. Preliminary results of Hierarchical Linear Modeling show that our intervention statistically significantly increases students’ growth mindset (medium effect) but has no detectable effect on STEM self-efficacy, perceived academic control, or course grades. Details of the project’s motivation, design, and findings can be found in Ref. [11].

C. Reflective practice

The SIISP research project is quantitative and utilizes a quasi-experimental research design, and the research findings report [11] is typical for such studies. This paper, however, uses reflective practice to take a qualitative approach to understanding the practice of self-efficacy research. Reflective practice is deliberate reflection and retrospective analysis to advance understanding of process, experience, or outcome [12]. It is employed in many disciplines and common in the field of educational studies.

As a research team, we engaged in collective reflective practice by: (1) reflecting on our experiences conducting SIISP and identifying areas where we encountered challenges, disagreements, or struggles while designing our scalable intervention, securing earnest student participation, measuring psychosocial variables, and accounting for the results; (2) analyzing the issues in conversation with extant literature on self-efficacy research in STEM, and (3) providing illustrations of how these issues manifested in the practice of research.

Albert Einstein said, “If I were given one hour to save the planet, I would spend 59 minutes defining the problem and one minute resolving it.” The purpose of this paper is to define problem areas within the practice of STEM self-efficacy research, in order to promote conversation around these issues and advance our collective understanding and our ability to affect student self-efficacy.
D. Critical issues

While many of the published studies testing interventions to increase student self-efficacy show statistically significant results, the literature as a whole is inconsistent. Some research efforts achieve desired outcomes, while others show weak or null effects [2, 13]. Our study failed to support our primary hypothesis that a brief intervention about growth mindset and academic control in the STEM context would increase student STEM self-efficacy. Why have some research efforts produced desired outcomes while others haven’t? What factors might contribute to these inconsistencies?

In this paper, we discuss three critical issues that either shaped or complicated our research efforts. We ground our discussion in both the STEM education literature and in practice. Sections II, III, and IV each analyze one of these critical issues. We conclude with implications for future research.

II. DEFINING AND MEASURING SELF-EFFICACY

Though self-efficacy has been a popular construct for over 40 years and has been widely employed as a research variable in multiple disciplines, its definition and measurement remain ambiguous [14, 15]. While most definitions incorporate the notion that self-efficacy is a cognitive self-evaluation of capability, other aspects of the definition are less clear.

A. Outcome expectancy

One ambiguity is whether self-efficacy is a belief about capability or an outcome expectancy [16]. This is a complicated aspect of the definition because Bandura says that self-efficacy includes beliefs about both current and future competence, and that these beliefs affect an individual’s choice of activities and behaviors. He frames self-efficacy as a determinant of how people feel, think, and behave [16, 17]. For example, he says that people with low self-efficacy tend to avoid difficult tasks while those who feel capable are willing to embrace a challenge [18]. This implies that behavior, or intended behavior, is an aspect of self-efficacy.

However, he also says that self-efficacy does not include the intention to behave in a specific way or to attain a specific goal; self-efficacy is not a prediction of what one will do. Simply put, it is not what you believe you will do, but what you believe you can do [19]. Thus, someone might believe that they can earn an A in a STEM course but have no intention of putting in the effort to do so at this point in time. Research projects assessing students’ self-efficacy may need to consider fine distinctions among its various dimensions.

As we developed the SIISP project’s measure of STEM self-efficacy, we had to make decisions about the aspects of self-efficacy we wanted to capture. These informed the types of items we included on our survey. In the final version, we have items reflecting both beliefs and behaviors, which is consistent with existing self-efficacy scales (e.g., the MSLQ, SOSESC-Physics, and GSES/GSE) [20–22]. Examples of our items targeting self-efficacy beliefs include “I can correctly solve typical homework problems in my STEM courses,” and “I can do well on quizzes and exams in my STEM courses.” Items targeting behavior intentions include “I study enough to do well on STEM quizzes and exams,” and “If I get something wrong in a STEM course, I try again and/or try harder instead of giving up.”

In exploratory factor analysis, the items targeting beliefs loaded as distinct sub-factors from those targeting behaviors, but they also correlated with each other. This supports the idea that beliefs about self-efficacy and outcome expectancies are distinct but related components of overall self-efficacy.

In addition to the question of whether to target self-efficacy beliefs and/or outcome expectancies with our questionnaire items, we were challenged by another ambiguity relating to the beliefs vs. behaviors dichotomy. As part of our instrument validation process, we conducted eighteen in-depth interviews with students from the target population who were not in classes included in the primary research study. Below are three excerpts from the transcripts.

Excerpt 1:
Q: How well do you think you are doing in your STEM classes in general?
A: I would say this semester not so well but previous to this semester I think I did good... I did all my homework and like all the assignments... this semester is a little different–
Q: So tell me more about this semester...
A: I’m already accepted to grad school so I don’t have motivation to come to classes anymore and do my homework so I keep putting that off... I know that sounds crazy... I should be but I just feel like I worked so hard to get accepted and now that I got accepted I feel like I should... I’m free to rest.

Excerpt 2:
Q: How well do you think you are doing in your STEM classes in general?
A: Not so well... not as well as I can. I think if I put more work in...
Q: And what’s limiting you? Why aren’t you putting in the work?
A: ... it’s more about motivation right now... school is not the priority for me right now... I feel like I can spend my time doing something else that is important in the short term. I know that school is really important for later on... um but it’s just about like prioritizing my time. (Later in the same interview) As I said right now I’m not as motivated to study but I know that if I put the work in I would definitely have like A’s... I know it’s possible and I think it’s possible for everybody.

Excerpt 3:
Q: How well do you understand difficult concepts in your STEM classes?
A: It depends on the class, honestly... if I feel like... if that’s
something that I can use in the future then I’ll put in work and I will be more motivated but if I feel like I don’t really need this for the future and it doesn’t impact my grades that much, I’ll probably just give up.

These statements suggest that a student’s interest in and/or motivation to succeed in a course or program interacts with or mediates their behaviors and/or behavior intentions related to self-efficacy. It is possible that at least some students believe they are capable of succeeding, but choose not to enact that capability for certain reasons. Thus, high or low self-efficacy beliefs may not always relate to the behaviors we associate with high or low self-efficacy. It may also be the case that outcome expectancy (intended behavior) is a separate construct, related to utility value and/or interest, and not a dimension of self-efficacy. We need to understand these distinctions and their implications if we are to design measurement instruments and interventions that work.

B. General or specific?

A second ambiguity in the definition of self-efficacy relates to the question of domain specificity. Although self-efficacy has been operationalized as a general construct (general self-efficacy or academic self-efficacy), most researchers treat self-efficacy as content-specific: mathematics self-efficacy, physics self-efficacy, etc. Bandura argues that self-efficacy beliefs are contextual, but he also says that some are more generalizable than others. The more similar two situations are and the more similar the tasks between two contexts, the more likely that self-efficacy beliefs generalize across the contexts [23].

The idea that STEM disciplines share significant commonalities and should be viewed as interdisciplinary and integrated rather than as isolated subjects has growing support. The educational goals, lecture-and-lab course structure, teaching methodologies, learning outcomes, evaluation practices, and applications are similar and require similar cognitive skills [13]. However, some research findings disagree, pointing to the importance of distinctions between STEM domains [24]. The degree to which self-efficacy is generalizable across STEM subjects remains an open research question.

For the SIISP project, we attempted to measure and target STEM self-efficacy, although all measurements and interventions were performed in the context of an algebra- or calculus-based introductory Physics course. We are not sure of the degree to which students were generalizing across STEM disciplines, as opposed to focusing exclusively on their beliefs about the Physics course they were taking. If we had developed and administered self-efficacy measures more narrowly focused on single STEM courses (like Chemistry, Math, Engineering, etc.) and assessed students in multiple courses, we could have gauged the degree to which they are the same or different constructs. Such research would contribute to our understanding of the generalizability of self-efficacy.

III. ACCOUNTING FOR CONTEXT

Many characteristics of an educational setting surrounding instruction have been found to play an important role in improving undergraduate success in STEM fields [3, 13]. These are collectively called the context of the instruction. Just as context can promote or constrain student learning, it can also influence the efficacy of interventions. We can consider educational context to be a moderator in self-efficacy research. Moderators affect the relationship between an independent and dependent variable by impacting the direction or strength of the relationship, like a dial that increases or reduces the effect [25].

Some researchers have sought to identify the factors affecting the scalability of interventions across diverse contexts and to explain how these factors limit or expand the effects of a specific intervention [3, 26, 27]. Other research into the effects of context can be found in the literature on classroom learning environments as well as on problem-based, inquiry-based, engagement-based, and active teaching and learning classrooms in STEM [28–30]. However, more work is needed to identify, measure, and track the specific contextual factors that influence the portability of interventions across settings. Contextual factors include:

- **Institution characteristics**: Structural aspects such as student body demographics, prominence of STEM programs, policies (e.g., faculty reward systems), resources/opportunities, campus environment/climate, and the academic selectivity of the institution. For example, Chang et al. found that students who participated in an undergraduate research program or a pre-professional club had significantly higher rates of persistence than those who did not [13, 31, 32].

- **Teacher and Classroom characteristics**: Student perceptions of the teacher including immediacy, credibility, and supportiveness as well as characteristics of the classroom climate including task-orientation, competition levels, and social belonging. For example, Deemer and Smith found that professorial concern was related to both intrinsic motivation and goal adoption [28, 29, 33–37]; and

- **National/regional cultural characteristics**: Cultural value systems and traditions like individualism vs. collectivism, power distance, sex roles, and feedback directness and transparency. Evans says we must proceed with caution when employing s-e measures and comparing ratings across cultures, because sources of self-efficacy are highly culturally dependent [38–40].

It may be the case that these factors are interrelated such that parsing the individual impacts will prove difficult. Still, we should strive to assess these contextual variables in educational environments and account for them as moderators, rather than leaving them as unmeasured factors potentially affecting outcomes.

The three institutions involved in the SIISP study are similar in that they are all large universities in the University
of North Carolina public higher education system, located within 80 miles of one another. They differ in population and focus, however: One is a former women’s college with a majority female enrollment and a racially diverse population; another is an 81% African-American HBCU (historically black college or university), awarding more Baccalaureate engineering degrees to African-Americans than any other US institution [41]; and the third is a top-tier land grant university recognized for its STEM programs, especially in engineering, statistics, and veterinary medicine [42].

HLM analysis enabled us to control for differences in institutional context as a level 2 variable. However, on-site observations by our project staff identified differences in the classroom, teacher, and relational characteristics of the institutions, courses, and lab sections in which our project activity occurred. We had no way of measuring these context variables, and in HLM analysis were not able to fully account for their potential moderating effects on our results.

IV. UNDERSTANDING RELATED PSYCHOSOCIAL FACTORS

Many psychosocial factors relate to STEM student persistence and academic performance. Some of these factors correlate with self-efficacy, and others predict or mediate it [5, 43–47]. Mediating factors can have a stronger effect on the dependent variable than the independent variable does, which makes them particularly important to identify [25].

Numerous psychosocial variables are of interest to the STEM education research community, including motivation, grit, performance ambiguity, self-regulation, perceived academic control (locus of control), mindset, belonging, interest, identity, meta-cognition, anxiety, goal-orientation, and self-reflection. Although they have been the focus of many studies across multiple disciplines, we know little about how they relate to self-efficacy and to each other. A number of models, both theoretical and empirical, attempt to explain their interrelationships. For example: Judge et al. hypothesize that self-esteem, neuroticism, locus of control, and generalized self-efficacy represent a common construct [48]; Maddux says that self-reflection and self-regulation may be prerequisites for self-efficacy [19]; De Feyter et al. propose that many psychosocial factors relate to the “big five personality traits” widely understood in psychology [49]; van Aalderen-Sveets et al. found self-efficacy, stereotypical thinking, and motivational beliefs to be mediators of the effect of growth-mindset on academic performance and career choices [50]; and Simon et al. found that students’ achievement goals, self-efficacy, and perceived autonomy support intrinsic motivation, emotions, and achievement which, in turn, predicts persistence in the science domain [51].

The SIISP project tested the hypothesis that increasing students’ growth mindset and perceived academic control (success-failure attributions, a.k.a. locus of control)—two central components of self-efficacy [16, 52]—would increase their STEM self-efficacy. We found that although our intervention did successfully increase students’ growth mindset, it did not increase their STEM self-efficacy. Why? Are we missing crucial mediating factors? How are these constructs related? We lack a comprehensive conceptual framework for understanding self-efficacy and its connections to growth mindset, perceived academic control, and other psychosocial factors. How do they interrelate? Under what conditions, in what contexts, and for whom?

V. SUMMARY AND DISCUSSION

We have identified three critical issues that shape and complicate research efforts seeking to measure self-efficacy and develop self-efficacy interventions: ambiguities in the definition of “self-efficacy,” accounting for contextual factors, and understanding the interrelationships between it and other psychosocial factors. We connected these issues to ongoing discussions in the literature and illustrated them and their implications with examples from the SIISP research project.

During that project, we also encountered several other challenging questions and obstacles, which we were able to navigate with guidance from extant literature and our collective experience. We are still investigating these critical issues, student engagement and students’ perceptions of personal relevance. We are also finalizing the statistical data analysis for our primary investigation, which we will report in the near future.

In this paper, our primary goal has been to promote discussion across the STEM education community about these three issues in self-efficacy research in order to increase our common understanding of the factors that might contribute to inconsistent research findings. We recommend the following avenues for future research efforts:

1. Research into the definition(s) and dimensions of self-efficacy and the rigorous testing of measurement instruments in light of those definitions and dimensions;
2. Development and testing of methods for identifying and assessing important factors in educational context;
3. Empirically and theoretically exploring the relationships among psychosocial factors, particularly mediation effects; and
4. Working both theoretically and empirically toward understanding the role of self-efficacy in STEM student academic persistence and performance.

We hope that these reflections help expose the complexity of STEM self-efficacy interventions, encourage conversation on these issues, and promote further research into measuring, modeling, and increasing student self-efficacy in STEM.

ACKNOWLEDGMENTS

This material is based upon work supported by the US National Science Foundation under Grant No. DUE-1612053.


[38] G. Hofstede, Dimensionalizing cultures: The Hofstede model in context, Online Readings in Psychology and Culture 2 (2011).


[41] Top 100 Producers of Bachelor’s Degrees (Diverse Issues in Higher Education, 2018).


