A Mixed Methods Approach Towards Defining A Student's Ranges of Self-Efficacy

Carissa Myers (she/her),[1] Vashti Sawtelle (she/her),[1,2] and Rachel Henderson (she/her)[1,3]

[1]*Department of Physics & Astronomy, Michigan State University, East Lansing, MI, 48824*
[2]*Lyman Briggs College, Michigan State University, East Lansing, MI, 48824*
[3]*CREATE for STEM Institute, Michigan State University, East Lansing, MI, 48824*

Traditionally, self-efficacy (SE), or the confidence in one's capability to execute a task, is measured using pre/post-surveys to demonstrate shifts in students' SE. In this work, we present a preliminary analysis of a single student drawing on a mixed methods approach to examine how their SE fluctuates over time. This novel design employs the Experience Sampling Method, a quantitative technique using surveys of domain-specific self-efficacy, and daily reflections, a qualitative technique investigating threats and supports towards students' SE. The preliminary analysis was broken into two strands: (1) using interquartile range (IQR) to define low, normal, and high SE for a student based on their survey scores, and (2) using the student's daily journal reflection responses as proof of concept for defining the student's SE as low, normal, or high from the IQR analysis of survey responses. Findings indicate the boundaries of a student's IQR can define high, normal, and low SE and the student's responses to the daily journal prompts corroborates these definitions.

I. INTRODUCTION

Self-efficacy (SE) is one's confidence in their ability to successfully perform a task [1] . Researchers have shown SE is predictive of students' achievement in science courses [2–4], persistence in science majors [5, 6] , and an important factor in science career choices [7, 8] . Historically, researchers have studied students' science SE using pre- and post-surveys to demonstrate a shift in students' SE from the beginning to the end of a course [9].

Researchers commonly define low and high SE through a comparison among groups [10, 11] or with a comparison across time [12]. However, these papers do not explore how to define low and high SE for an individual student based on their own scores.

While these techniques for comparing the growth or differences in students SE are useful, they do not provide information unique to the individual. Identifying a marker of low or high SE to an individual provides access to understanding moments that threaten or support SE development and, in turn, could potentially lead to their persistence in science. In this paper, we present utilizing the statistical tool, interquartile range (IQR), for defining the ranges of SE (e.g. low, normal, and high) and address the research question: what does low, normal, and high SE look like for an individual student?

While we only utilize data from one student in this work, we do not intend to make any generalizations about students' SE. Rather, by analyzing one case, we can begin to operationalize our definitions to then explore how these definitions may change across students. Further validation of this idea will be conducted in future studies.

II. CONTEXT & METHODS

We conducted a mixed methods study at a large research-intensive university in the Fall 2020 academic semester. Students were selected to participate in this study if they were a physics major enrolled in upper-division physics course(s) and had transferred at least one credit from a two-year college. Six students chose to participate in the Fall 2020 semester. The mixed method study employed an explanatory sequential design [13]. We collected quantitative data through the Experience Sampling Method (ESM) [14] followed by collecting qualitative data through Daily Journal Reflections.

The ESM is a method for investigating individuals' experiences in their day-to-day lives. For our context, we utilized this method to quantitatively capture changes in students' individual SE throughout their day. To capture their SE, we adopted the SE scale from a survey developed by Nissen [15]. Questions on the survey included what task a student is currently engaged in, if the task is related to a course, and a series of task-specific SE items. In this work, we focus on the analysis of a student's responses to the following items: (1) How skilled are you in the activity? ("skilled"), (2) Do you feel in control of the situation? ("feeling in control"), and (3) Are
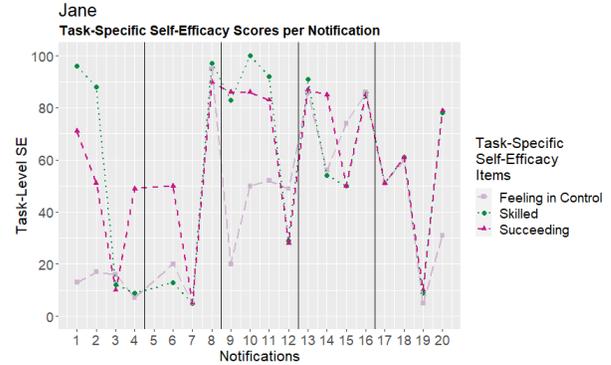

FIG. 1. The plot shows Jane's scores to the task-specific self-efficacy items from each survey she completed in Week 1. A notification represents a single survey response. For example, Notification 1 represents the first survey for Jane, and she scored 13 for the feeling in control item, 96 for the skill item, and 71 for the succeeding item. The vertical lines symbolize the notifications within a single day of data collection. For example, Day 1 of data collection are Notifications 1-4.

you succeeding at what you are doing? ("succeeding"). For information regarding the mixed methods design see Henderson's and Sawtelle's work [16] and for a description of the validity of the survey, see Nissen's work [15].

For two weeks with a week break in between, students were randomly notified to fill out the ESM survey four times throughout the work day (9:00am to 6:00pm, Monday-Friday). Thus, this design resulted in a total of 20 possible notifications per week for each individual. Participants received the ESM survey notifications through a smartphone application called LIFEDATA and received an incentive of $50 dollars for completing 80% of the possible notifications.

After the students completed the ESM survey notifications each day, researchers visually analyzed student responses in search for highs and lows in a student's SE (see Figure 1 for an example of one week of student responses to the ESM survey). From this analysis, individualized journal prompts were designed to probe further into a student's SE around such highs or lows while allowing students to reflect on their experiences throughout their day. The journal prompts were delivered and completed in an individualized Microsoft OneNote notebook. This paper intends to present the usefulness of the IQR to define high and low SE in the ESM survey data and then showcase a few entries from the journal reflections as evidence for corroborating those definitions.

A. Interquartile Range Boundary Definitions

Using a student's responses from the task-specific SE survey items on the ESM, we used R, an open-source statistical programming tool, to aggregate the raw data in the form of a box-and-whisker plot. A box-and-whisker plot is a means of visualizing the distribution of a student's responses to the

survey items. The "box" part of the plot defines the interquartile range (IQR), which is the difference between the first and third quartiles, Q_1 and Q_3, respectively. The IQR represents the spread of the middle half of the data [17] and we use this to define a normal SE response for an individual student. R uses the Tukey Method [18, 19] which defines Q_1 as the median of the lower half of the data and Q_3 as the median of the upper half of the data [17], while including the overall median in each of those definitions.

The "whisker" parts of the plot defines the upper and lower 25% of the data and are bounded by Q_1 and Q_3. In this work, we utilize these bounds to define low and high SE for an individual student. We define a moment of low SE when a student responds with a score below Q_1 for all three SE items on the ESM survey and a moment of high SE when a student responds with a score above Q_3 for all three items.

We acknowledge that our definitions of low, normal and high SE do not utilize statistical significance. We chose more liberal definitions to ensure we would not miss a threat or support to a student's SE toward the given task.

B. Jane

For this analysis, we will focus on one individual student who participated in the Fall 2020 study - Jane (pseudonym). At the time of the study Jane was a white, female physics major, and identified as a community college transfer student. At the time of the study, Jane was actively parenting young children. We focus on Jane because she completed 95% of the notifications, had a lot of variance in her quantitative data, and was enrolled in various STEM courses. Her STEM courses in the Fall 2020 semester consisted of two physics courses, which we will refer to as Physics Course and Physics Lab, a math course, and a computational course. It is important to note in Fall 2020, the institution Jane was enrolled in was still heavily impacted by the COVID-19 pandemic, and the majority of classes were ran virtually.

III. RESULTS

In this section, we will demonstrate how we use Jane's responses to the task-specific SE survey items to define low, normal, and high SE, and corroborate the credibility of these definitions using her daily journal reflections. Figure 2 presents Jane's aggregated responses to each of the SE items for Week 1 and Week 2 in the form of a box-and-whisker plot. The boundaries of Jane's IQR (i.e. Q_1 and Q_3) were used to define when Jane had low and high SE towards a task she was performing in-the-moment. Table I presents her first and third quartiles for each SE item across each week of data collection. We used these boundaries to identify notifications indicative of low and high SE. For example, during Week 1, a notification is a moment of low SE for Jane if she responds with a SE score below 16.5 on the "feeling in control" item,
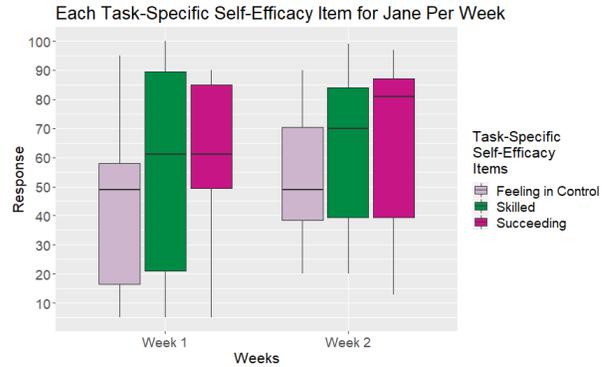

FIG. 2. The box plot visually shows the distribution of Jane's responses to the three, task-specific SE items on the survey across the two weeks of data. Week 1 represents the first week of data and Week 2 represents the second week of data.

TABLE I. Q_1 and Q_3 per week, calculated using Jane's survey responses, are provided below. The difference between Q_1 and Q_3 is the IQR visualized in Figure 2.

	Q_1		Q_3	
	Week 1	Week 2	Week 1	Week 2
Control	16.5	38.5	58.0	70.5
Skill	21.0	39.5	89.5	84.0
Success	49.5	39.5	85.0	81.0

21 on the "skilled" item, and 49.5 on the "succeeding" item. A moment of high SE during Week 1, is defined when Jane responds with a SE score above 58 on the "feeling in control" item, 89.5 on the "skilled" item, and 85 on the "succeeding" item. Below, we present the notifications that align with our definitions and the qualitative evidence that corroborates those findings.

A. Qualitative evidence for defining low SE for Jane

Using the definition of low SE for Jane established by Q_1, we identified the individual notifications that were moments of low SE for her within each week. These notifications and the task Jane reported for each notification are shown in Table II disaggregated by week. Within Week 1 and Week 2, there were four notifications per week identified as moments of low SE for Jane. Two themes appear from analyzing Table II: (1) Jane commonly experiences low SE when performing a task associated with either her computational or math course, and (2) Jane commonly experiences low SE when performing an academic task associated with one of those courses and balancing her personal life. We will first provide corroborating evidence for the lower boundary of Jane's IQR (Q_1) identifying low moments of SE in Jane's computation or math classes, and then show evidence supporting low moments of SE when Jane is balancing academics with her personal life.

TABLE II. Jane's reported task associated with each notification, identified as low, dissaggregated by week.

	Notification Number	Reported Task
Week 1	3	Nursing a baby in [Math Course]
	4	Nursing my baby just finished [Computational Course] homework
	7	Nursing my baby and [Math Course] homework
	19	[Math Course] class
Week 2	3	Feeding my baby and [Physics Lab] homework
	4	[Computational] homework
	7	[Computational Course] class
	13	[Math Course] recitation

There is qualitative data corroborating that Jane would score tasks located within her computational and math courses as moments of low SE. When Jane was asked to reflect about her feelings about classes and academics, she shared her experiences toward tasks within the course:

> *"The math classes are very difficult to ask questions in and are overwhelming in the way they teach/treat students...[Computational Course] is too fast paced and jumps from concept to concept too fast they have help rooms but the pace of the content leaps seem too overwhelming."*
> - Jane, Week 1, Reflection 1

Jane summarizes her experiences with tasks located in her math courses. She first describes the task of asking questions and communicating with the course instructors, located within her math courses, as difficult. Jane having low SE towards completing tasks within her math course suggests the lower boundary of her IQR (Q_1) reasonably identified moments of low SE in her math course. In this same quote, Jane then describes experiences with tasks such as the presentation of materials located in her computational course as too fast paced and disorganized. These qualitative data also corroborate the lower boundary of her IQR (Q_1) identifying moments of low SE for her.

Jane further discussed her experiences with tasks located within her courses when she was asked to reflect on moments where she felt very unskilled and not successful:

> *"The content in [Math Course] is just hard. I got frustrated and behind in the beginning of the semester and just felt like I couldn't catch up or grasp the content... The [Computational] class I feel the same way. I was doing good then I started to get lost. Now I feel like I am missing something key and I don't know what it is. I wish I could talk to the teachers...I don't have the confidence to speak out and get the help I need in the programming and [Math] classes...I went to their help rooms I felt dismissed and like a prob-*
> *lem rather than welcomed."*
> - Jane, Week 2, Reflection 4

Jane summarizes the challenges she experienced in doing tasks like understanding the material and engaging in the math and computational courses. She gave further details about her communication with her math and computational course instructors because, in this quote, she shares not having the confidence "to speak out and get help". This explicitly demonstrates Jane's SE towards tasks - communicating her work and seeking help - in her math and computational courses. She provides an example where she tried seeking help, but left feeling dismissed and not welcomed. Jane's description of performing tasks in her math and computational courses support her reporting tasks associated with these courses as low, confirming the notifications the lower boundary of her IQR (Q_1) identified as moments of low SE.

We now return to the second theme the IQR identified as low SE moments for Jane, which occurs when she is performing an academic task associated with her math or computational courses and balancing her personal life. There is qualitative data corroborating Jane scores these tasks as low SE. When Jane was asked to imagine answering questions on the survey differently, she reflected over her experience balancing her personal life and academic life:

> *"If I had been in a real classroom with access to real people, I would feel much more in control and capable of learning and focusing on the material. ... The struggles of my personal life interferes much more with online learning than with in person classes and learning."*
> - Jane, Week 1, Reflection 2

In this quote, Jane gives a hypothetical where she suggests she would feel more capable of learning and focusing on the material if she had access to in-person classes. These tasks are linked to Jane's SE because she describes what she feels she is "capable" of doing in relation to these tasks. She then discusses how the lack of attending classes in-person impacts her SE. Jane having low SE towards completing tasks when they require her to simultaneously perform an academic and personal life task is reasonable from this qualitative data confirming the notifications the lower boundary of her IQR (Q_1) identified as moments of low SE for her.

B. Qualitative evidence for defining high SE for Jane

Using the definition of high SE for Jane established by Q_3, we then identified the notifications that were moments of high SE. These notifications and the task Jane reported for each notification are shown in Table III dissaggregated by week. Within Week 1, there were two notifications identified as moments of high SE for Jane; within Week 2, there were three. The theme that Jane experiences high SE when performing a task associated with her Physics Course emerges from analyzing Table III.

TABLE III. Jane's reported task associated with each notification, identified as high, dissaggregated by week.

	Notification Number	Reported Task
Week 1	8	[Physics Course] class
	13	Feeding baby getting ready to math class
Week 2	2	[Physics Course] homework
	6	Lunch
	16	[Physics Course]

There is qualitative data corroborating that Jane scores tasks located within her Physics Course as moments of high SE. In the journal prompts, Jane was asked to reflect about a moment where she was really confident in her performance in a particular course, and she chose to describe an experience in Physics Course:

> "Yesterday in [Physics Course] I felt confident in my knowledge and experience and it was exciting and felt useful and I know it was meaningful to my career path."
> - Jane, Week 1, Reflection 3

We can connect this qualitative data to Notification 8 within her quantitative data in Figure 1. Jane's use of "Yesterday" in the Reflection on Day 3 indicates that she experienced that moment in Day 2, which corresponds precisely with Notification 8 occurring on Day 2. This quote indicates Jane has high SE towards completing a task within Physics Course which corroborates Notification 8 as a high moment of SE for Jane. We can expand this to Jane having high SE towards completing multiple tasks within Physics Course when taking into consideration Jane's reflection about her feelings towards classes and academic experiences:

> "[Physics Course]...they take time to ensure all students understand fully and are invested in their students..."
> - Jane, Week 1, Reflection 1

Jane described the task of communicating with her instructors located in Physics Course stating, they "ensure all student understand fully." This suggests Jane feels capable of receiving support in understanding the material. Understanding material is a key factor in performing other tasks within the course (e.g homeworks, work on in-class activities, using help room hours, etc.). This corroborates the analysis of the upper boundary of her IQR (Q_3) indicating Jane has a high SE towards tasks within Physics Course.

There is no confirming qualitative evidence that identifies Notification 13 from Week 1 and Notification 6 from Week 2 as moments of high SE for Jane. Notification 13 is a task requiring Jane to balance an academic task associated with her math course and a personal life task. We might expect Notification 13 to be a moment of low SE instead. We posit that the reason this Notification is high SE is because Jane

is "getting ready to Math class" in Notification 13, which is different than performing the task in-the-moment.

IV. DISCUSSION

In this section, we return to our original research question about what low, normal and high SE look like for an individual student. We then describe some of the limitations to this study and how we addressed these limitations in a recent, Fall 2021, iteration of our work. From the case of Jane, we established IQR as a useful tool for identifying Jane's ranges of SE, in turn, supporting the long-term goal of designing interventions for SE development. Low SE is defined as when Jane scores below Q_1 for all three task-specific SE items; high SE is defined as when Jane scores above Q_3 for all three task-specific SE items. Implicitly, we claim that "normal" SE for Jane is when her SE scores are within the IQR.

A limitation to this work was the amount of quantitative and qualitative data we collected from Jane. For example, Notification 6 in Week 2 was identified as a moment of high SE for Jane, but we did not have enough quantitative data to investigate how commonly Jane has high SE towards the task of eating. We addressed this limitation in a study in Fall 2021 by increasing the length of the study to four weeks. We will investigate the impact of this change in future work as well as exploring if the data should be aggregated or disaggregated by week.

A second limitation in this study was the corroborating qualitative evidence. For Notification 13 from Week 1 and Notification 6 from Week 2, we did not have the qualitative data to support these as moments of high SE for Jane. We addressed this limitation in our Fall 2021 study by crafting daily journal prompts to investigate specific tasks reported in the ESM survey response(s). As a part of future work, we plan to investigate how specific tasks and contexts impacts a student's SE [10, 20–22].

Finally, we only utilized the boundaries of IQR to define a single student's ranges of SE and Jane's ranges of SE may not be representative of all students. We have started the preliminary analysis with another student from this study, who has a ceiling effect (e.g. the IQR includes the maximum value within their scores) in their box plot. Utilizing criteria from the IQR to distinguish between high and normal for this individual may be less informative. In future work, we hope to utilize Jane's Fall 2021 data and other students' data to explore the effectiveness of using IQR to define ranges of SE and address these limitations.

ACKNOWLEDGMENTS

[1] A. Bandura, Self-efficacy: Toward a unifying theory of behavioral change, Psychological Review , 191 (1977).

[2] S. L. Britner, Motivation in high school science students: A comparison of gender differences in life, physical, and earth science classes, Journal of Research in Science Teaching **45**, 955 (2008), _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/tea.20249.

[3] L. M. Larson, K. M. Pesch, S. Surapaneni, V. S. Bonitz, T.-F. Wu, and J. D. Werbel, Predicting Graduation: The Role of Mathematics/Science Self-Efficacy, Journal of Career Assessment **23**, 399 (2015), publisher: SAGE Publications Inc.

[4] S. Lau and R. W. Roeser, Cognitive Abilities and Motivational Processes in High School Students' Situational Engagement and Achievement in Science, Educational Assessment **8**, 139 (2002).

[5] S. M. Glynn, G. Taasoobshirazi, and P. Brickman, Science Motivation Questionnaire: Construct validation with nonscience majors, Journal of Research in Science Teaching **46**, 127 (2009), _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/tea.20267.

[6] R. W. Lent and A. Others, Comparison of Three Theoretically Derived Variables in Predicting Career and Academic Behavior: Self-Efficacy, Interest Congruence, and Consequence Thinking, Journal of Counseling Psychology **34**, 293 (1987).

[7] P. R. Aschbacher, E. Li, and E. J. Roth, Is science me? High school students' identities, participation and aspirations in science, engineering, and medicine, Journal of Research in Science Teaching **47**, 564 (2010), _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/tea.20353.

[8] D. A. Luzzo, P. Hasper, K. A. Albert, M. A. Bibby, and E. A. Martinelli Jr., Effects of self-efficacy-enhancing interventions on the math/science self-efficacy and career interests, goals, and actions of career undecided college students, Journal of Counseling Psychology **46**, 233 (1999), place: US Publisher: American Psychological Association.

[9] R. Henderson, V. Sawtelle, and J. M. Nissen, Gender & Self-Efficacy: A Call to Physics Educators, The Physics Teacher **58**, 345 (2020).

[10] V. Sawtelle, E. Brewe, L. H. Kramer, C. Singh, M. Sabella, and S. Rebello, Positive Impacts of Modeling Instruction on Self-Efficacy (Portland, (Oregon), 2010) pp. 289–292.

[11] E. M. Marshman, Z. Y. Kalender, T. Nokes-Malach, C. Schunn, and C. Singh, Female students with A's have similar physics self-efficacy as male students with C's in introductory courses: A cause for alarm?, Phys. Rev. Phys. Educ. Res. **14**, 020123 (2018).

[12] C. Lindstrøm and M. D. Sharma, Self-Efficacy of First Year University Physics Students: Do Gender and Prior Formal Instruction in Physics Matter?, International Journal of Innovation in Science and Mathematics Education **19** (2011), number: 2.

[13] N. V. Ivankova, J. W. Creswell, and S. L. Stick, Using Mixed-Methods Sequential Explanatory Design: From Theory to Practice, Field Methods **18**, 3 (2006), publisher: SAGE Publications Inc.

[14] J. M. Hektner, J. A. Schmidt, and M. Csikszentmihalyi, *Experience Sampling Method: Measuring the Quality of Everyday Life* (SAGE, 2007) google-Books-ID: 05e5d_KBYY0C.

[15] J. M. Nissen and J. T. Shemwell, Gender, experience, and self-efficacy in introductory physics, Phys. Rev. Phys. Educ. Res. **12**, 020105 (2016).

[16] R. Henderson and V. Sawtelle, Implementing a mixed-methods approach to understand studentsâ self-efficacy: A pilot study (2020) pp. 204–209, iSSN: 2377-2379.

[17] D. L. Hahs-Vaughn and R. G. Lomax, *An Introduction to Statistical Concepts* (Routledge, 2020) google-Books-ID: X1PODwAAQBAJ.

[18] E. Langford, enQuartiles in Elementary Statistics, Journal of Statistics Education **14**, 3 (2006).

[19] How to Find the Interquartile Range in R.

[20] J. Stewart, R. Henderson, L. Michaluk, J. Deshler, E. Fuller, and K. Rambo-Hernandez, Using the Social Cognitive Theory Framework to Chart Gender Differences in the Developmental Trajectory of STEM Self-Efficacy in Science and Engineering Students, J Sci Educ Technol **29**, 758 (2020).

[21] E. L. Usher and F. Pajares, Sources of Self-Efficacy in School: Critical Review of the Literature and Future Directions, Review of Educational Research **78**, 751 (2008), publisher: American Educational Research Association.

[22] X. Wang, N. Sun, S. Y. Lee, and B. Wagner, Does Active Learning Contribute to Transfer Intent Among 2-Year College Students Beginning in STEM?, The Journal of Higher Education **88**, 593 (2017).