EXPLORING INTEGRATION WITHIN A MIXED METHODS APPROACH TO ILLUMINATE REAL-TIME ACADEMIC EXPERIENCES SHAPING STUDENTS' SELF-EFFICACY

By

Carissa Lynn Myers

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Physics—Doctor of Philosophy

2025

ABSTRACT

Research on self-efficacy (SE) - one's confidence in their ability to successfully perform a task - has been shown to predict students' achievement in science courses, persistence in science majors, and their science career choices. Traditionally, SE is assessed using quantitative surveys to explore changes in students' SE. However, this method limits the opportunity to capture the specific events shaping students' SE. To address this limitation, qualitative methods like interviews have been used, but they often involve significant time gaps between when specific events occur and when students report on them. My dissertation work aimed to design a mixed methods approach to leverage quantitative and qualitative methods to explore students' SE in real time.

For the design, we utilized an explanatory sequential mixed methods approach, which employed the Experience Sampling Method (ESM), a technique for capturing students' experiences throughout their daily lives combined with individualized daily journal prompts. Specifically, in alignment with the Experience Sampling Method (ESM), we randomly distributed four short surveys per day via a phone application to capture students' domain-specific SE (i.e., task-specific, course-level, and career-level). As the ESM data was being collected, I analyzed the survey responses to identify threads that would inform the writing of an individualized daily journal prompt, which were disseminated to each student at the end of the day.

Throughout the dissertation work, I leveraged our design to explore how to integrate the ESM survey responses with the individualized daily journal prompts to elicit SE information. Specifically, I utilized a nested case study approach, in which I compared the identified threads and the daily journal prompts written across two iterations of our mixed methods approach (i.e., Fall 2020 and Fall 2021) for six students. From this work, by adding a point of integration (i.e., examining both open- and close-ended ESM survey data to build threads to inform the writing of daily journal prompts), this lead to a more diverse range of threads to inform the daily journal prompts, as such, the daily journal prompts were able to explore the nuances of students' SE by including context-specific information from moments throughout an individual's day. Thus, adding points of integration in a mixed methods design allows for the quantitative and qualitative data to expand on

one another to add additional insights about the construct of interest.

As a result of the context-specificity that the additional point of integration added, we developed a codebook to explore the ways in which the individualized daily journal prompts were capturing the nuances of students' SE. From this work, we found that by including context-specific information, the prompts were able to further explore the domain-specific nature of SE and well-known sources of SE. This codebook will support STEM education researchers in considering the ways that their qualitative methods can capture the nuances of SE.

Finally, the last piece of this dissertation work included identifying the evidence of integration that supports that we are interpretting the ESM survey data with the individualized daily journal prompts to be able to explore students' SE within the data analysis stage. This work is critical for being able to confirm that the individualized daily journal prompt responses are exploring the situations and contexts that the ESM is reporting and measuring SE in relation to; thus, serving as a proof of concept that the individualized daily journal prompt responses and ESM surveys are exploring students' SE in relation to their real-time experiences. As a result of this dissertation work, STEM education researchers have been introduced to an innovative mixed methods approach to explore constructs in relation to students' real time experiences, and provides techniques for integration during the data collection and data analysis stage.

ACKNOWLEDGMENTS

[This will be heartfelt when I have a functioning frontal lobe again. Right now: thank you to the people who kept me alive, caffeinated, and convinced I could finish this. You know who you are. I still need to circle back to a few of you here—don't worry, I haven't forgotten.]

Thanks to my advisor, Dr. Rachel Henderson, for meeting with me every week and letting me take up extra time in those meetings - definitely more than the agreed hour, lol. You gave me space to grow as a researcher, provided the guidance I needed to strengthen my ideas, and helped me learn to advocate for myself and my work. It meant the world to me that you trusted my writing process and let me *cook*, so to speak, and then helped make sure I got it all out there for the world to see. I couldn't have asked for a better advisor - I know I'm a handful, always taking on more than I can handle - and your steady support kept me on the path through my dissertation. Thank you for sending me to as many conferences as possible, especially near the end. Those spaces, full of colleagues, peers, and friends, motivated me to keep going when things got tough. This dissertation is not just my victory, it is yours too. It is a celebration of your mentorship. So, if you see her, make sure to also congratulate her! We did it! Also, Rachel, if you read this, you are stuck with me forever. Hahaha.

Thanks to Dr. Vashti Sawtelle for being a co-mentor throughout my dissertation journey. Whenever I had a meltdown, felt imposter syndrome creeping in, or doubted myself, Vashti gave me space to feel and then lent me her strength and compassion to move forward. I can confidently say that I would not have got here without you. You helped me see value in myself and my work, and I hope to offer the same strength and wisdom to others someday that you offered to me.

Thank you to my committee: Dr. Daryl McPadden, Dr. Jennifer Schmidt, Dr. Brian O'Shea, and Dr. Gail Richmond. I appreciate the support and guidance you have provided throughout my research and career development. Your insights challenged me to grow and helped shape the final product into something that I am proud of. This also feels like the right place to apologize for the length of this dissertation. You deserve extra thanks (and maybe a medal).

A very large thanks to the students who participated in my studies. This work would not exist

me the space to vent and work through so many things that I was going through in a safe space. Your leadership inspired me in so many ways, including considering what a community means and how to support and develop one. You gave me the space to teach and care for others when I was buried in research, reminding me why I care about education.

Thanks to Laura Wood!!!

Thank you to Kirsten Tollefson. You have supported me in so many ways, from navigating the ins and outs of graduate school, to stepping into student leadership roles, to managing the National Science Foundation Graduate Research Fellowship within MSU's systems. Your guidance and steady support have made a big difference in my ability to show up and succeed.

Thank you to Kim Crosslan. You supported me to get out of my own head, gave me hugs when I needed them, and handled all the logistical chaos that comes with being a graduate student. I am so grateful for your constant presence and care throughout this process.

Thank you to my friends — Julia Hinds, Kelsey Bees, Joshua Wylie, Camila Monsalve Avendaño, Alyssa Waterson, Erin (Frank) Dachille, Andrea Wooley, David Greene, Hannah Berg, Jamie Kimble, Senora Blanco, Daniel Sharkey, Julia Willison, Rob Dalka, Liam McDermott, Danielle Seppala, Roy Salinas, and Bryan Stanley — for the many forms of support and community throughout the dissertation era. I know I missed some important names because the list of people who came to see me is long. If you are not listed here, please know that I still have a deep appreciation for you.

You all made this journey enjoyable—even fun at times—and I'm very grateful for that. A special thanks goes out to Julia Hinds, Joshua Wylie, Alyssa Waterson, Bryan Stanley, and Roy Salinas for dragging me through graduate coursework. I truly couldn't have done it without you.

Thank you to Dr. Jackie Chini at The Ohio State University for seeing potential in me and giving me the opportunity to join your lab after my dissertation work. You encouraged me early on to learn to advocate for myself in my career, and I've been striving to do that throughout my time in graduate school. I am so excited to be working with you, and collaborating during this transition has reminded me that this work does not end with a dissertation. Your mentorship and belief in my

potential have meant a lot to me as I enter the next stage of my career.

A special shout out to all the places that gave me space to work, think, and write: Your warm vibes, drinks, and food fueled this dissertation. Thank you to 517 Coffee Company, Koala Cafe & Bakery, Hooked, Blue Owl Coffee, Château Coffee Co., Strange Matter Coffee and so many others in the area. You all kept me going more times than I can count.

There are probably many more people to thank for their support, encouragement, and presence throughout this journey. If I did not name you directly, please know that your impact has not gone unnoticed. I am grateful for every conversation, check-in, shared laugh, and moment of solidarity that helped me carry through.

Thanks to [funding agency]

This dissertation is a reflection of not just my work, but the village that helped me carry it to the finish line. Thank you all.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1   Grounding My Dissertation

In starting this dissertation, I want to be transparent about the process that led me here. This opening section, titled "Grounding My Dissertation," is not just a prelude to the research but a space where I reflect on the epistemological questions that have shaped my work from the very beginning. As someone deeply involved in Physics Education Research (PER) for several years, I have come to realize that the researcher cannot be disentangled from the research. Our perspectives, experiences, and values inform the questions we ask, the methods we use, and the interpretations we draw. This section is, therefore, an opportunity to demonstrate how my own experiences, reflections, and positionality have influenced my research journey.

I am beginning with these reflections to not only offer insight into the origins of my work, but also to invite a dialogue about the assumptions that underpin our research practices. Stephen Secules wrote a piece inviting the engineering education community to think about "what academic expertise means epistemologically and positionally, how each of us knows what we know" [1]. This question has driven my dissertation work within PER: How do we come to know what we know? What is the role of the researcher's lens in shaping the inquiry? These are questions I have been grappling with since my early involvement in PER as an undergraduate, and one that I continue to explore throughout this dissertation and plan to pursue as a professional in PER. By grounding this work in my personal reflections, I aim to make my process visible and to foreground the ways in which I, as a researcher, am inherently woven into the fabric of this study.

1.2   Taking Pause: Why Explore a Methodology

In the early stages of my dissertation work, I intentionally chose a mixed methods research project, driven by the desire to engage with both quantitative and qualitative methodologies. This decision was not only about gaining experience with diverse research methods, but also about developing a deeper understanding of the role of methodology in designing a mixed methods study. In designing research, it is crucial to distinguish between methodology and methods, as they

1

represent two fundamentally different layers of the research process [2–4]. The methodology refers to the overall strategy, the rationale, and the philosophical framework that guides the entire study [5]. On the other hand, methods are the specific techniques or tools used to collect and analyze data [5]. These are the "how" of research, the concrete practices that allow one to operationalize the methodology. In this way, the methods are like plays in a playbook, but the methodology is the strategy that determines the game plan, the rules of engagement, and the long-term objectives. Although this distinction may seem subtle, it is crucial in framing how we approach research in a way that is both rigorous and deeply thoughtful. As such, I was curious how the combination of quantitative and qualitative approaches could enhance the richness of my findings and provide a more comprehensive understanding of my chosen topic. Moreover, I began to realize that the methods we choose are neutral - they shape the very narrative of the research itself. How a study is framed, whether through the lens of a particular methodology, perspective, or assumption, plays a critical role in shaping the findings and how they will be interpreted and applied in practice.

This realization was solidified in a conversation with Dr. Traxler, my undergraduate research advisor before I had started my dissertation work, about the ethical considerations of network analysis. Dr. Traxler emphasized the importance of making the process of constructing networks transparent in research publications because the way networks are built directly influences the conclusions drawn. This conversation left a lasting impression on me, reinforcing my commitment to conducting research that is not only methodologically rigorous but also transparent in the ways the research is performed. This includes being transparent about the decisions made and the assumptions and perspectives driving the work. I aimed for my research to be reflective, ensuring that I remained conscious of the framing and ethical implications behind every methodological choice. Thus, I did not want this dissertation to be implementing a research design and sharing my analysis and findings from student data. This dissertation is about taking a pause, sitting with the methodology, and exploring the methodological space before sharing narratives around students' experiences; as well as, sharing the significance of sitting with a methodology first.

As such, I am pausing to explore a methodology because I realized that the methodological

approach itself - the choice to blend the quantitative and qualitative approaches - is not simply about mixing techniques for variety or experience, but about intentionally shaping how the knowledge is produced [6]. The strategic combination of methodologies reflects my belief that the nature of knowledge is complex and multifaceted, and that no single method can fully capture the richness of some research questions. Thus, by integrating these approaches, I aim to uncover how these approaches can be intentionally used to explore the nuances of self-efficacy in real time.

In considering the intentional integration of both quantitative and qualitative approaches to capitalize on mixed methods approach, the research design itself also plays a pivotal role in how the study unfolds and ultimately how knowledge is produced [6]. Research design plays a crucial role in setting the direction and trajectory of a study, guiding it from conception to conclusion in an effort to answer specific research questions [7]. From the outset of my dissertation work, I recognized the importance of research design - not just as a methodological tool, but as a foundational element that shapes the entire study. The design of a research study influences how questions are framed, how data is collected, and ultimately how the findings are interpreted. This understanding deeply influenced my approach, as I sought to engage fully with the process of designing a study - from the selection of methodology to its implementation - so I could develop a comprehensive understanding of its significance.

When I embarked on my dissertation, my primary focus was to immerse myself in learning methods and increase my understanding of designing a study that would run the full gamut of research activities. While I have placed intentional thought into each stage of the study, my focus has been on specific aspects of the process, allowing me to engage deeply with the methodology and its impact on the study's outcomes. This was more than just a technical goal; it was an intentional effort to critically understand how research methodology themselves shape the outcome of educational studies.

In educational research, particularly when examining the experiences of marginalized students, the methodologies we choose have a profound impact on how we interpret and understand student experiences [8–12]. Yet, much of current research tends to prioritize results over the methodological

decisions that shape those results [13]. I've observed that the methods sections of research articles are often given less attention than the findings, despite the fact that these decisions directly impact what we conclude about students' experiences and outcomes. This gap troubled me, as it means that the influence of methodological decisions on the findings may be discussed only in limited ways or not given the room to be fully acknowledged. As such, this made me realize how often the methodological decisions that influence the findings are overlooked.

The findings we produce - often shaping classroom practices, department structures, and educational policies - are influenced by these behind-the-scenes decisions about research design. Without a critical examination of these methodologies, we risk reinforcing stereotypes or deficit models that harm minoritized groups [9]. For me, this personal realization underscored the importance of taking a deeper and more critical look at the processes behind research, not just the outcomes.

Additionally, I chose to prioritize a methods-heavy dissertation because I have observed that many research articles tend to emphasize the results, and due to the length of articles, researchers will give more space to the results than the methods sections, which can mean that the methodologies that shape those findings are not given as much space [13]. I have struggled with this tension because I think there is much work that researchers do behind the scenes in making decisions about their research methodology that impacts those findings that may not get published because the goal of the paper is to potentially publish the findings. In some sense, this feels not right to me because these findings, which influence classrooms, departments, and more broadly, the academic system itself, are influenced by these decisions regarding methodology that may be left in the background.

However, I have always believed that we must approach this work with a critical lens - recognizing that education research is not conducted in a vacuum. It has real-world consequences and can significantly impact people's lives. Without a critical perspective, there is a risk of reinforcing deficit frameworks, especially when dealing with minoritized, underrepresented, or marginalized students - those who are already navigating complex systems of oppression.

This is why I decided to pursue a methods-intensive dissertation with the aim of being critical of the methods, analysis, and interpretations that may potentially emerge from this research. In

doing so, I hope to encourage other researchers to engage thoughtfully and critically with their research design and implementation. This is important because the findings we produce do not just inform academic theory – they influence policies, practices, and the lives of the very individuals we study. Thus, we must be careful about how we design and interpret our studies, recognizing the real-world impact they can have on students and the systems they inhabit.

As I critically engage with the design of my research, I am reminded of the profound responsibility we bear at researchers. When investigating academic environments and the systems within which they operate, it is essential to acknowledge the complexities of students' experiences – particularly those from marginalized communities. These complexities may be hard to capture with certain methodologies, meaning they may be overlooked or oversimplified, and in doing so, research can inadvertently reinforce existing inequities [10, 11, 14]. This is why it is not enough to simply analyze the results; we must also critically examine the methodologies, analyses, and interpretations that shape our understanding of these groups. We must ask ourselves whether the methodologies we employ inadvertently reinforce dominant narratives that uphold systems of oppression, or whether they can be designed to actively challenge and disrupt these narratives, fostering more equitable and inclusive research outcomes.

In choosing to focus on methodology and mixed methods in this dissertation, my goal is to foster a more holistic, ethical approach to educational research. This approach is not merely about evaluating outcomes; it is about interrogating the processes through which those outcomes are determined. While the practical applications of these methodological decisions will be explored in later chapters, here I focus on the reflective process and the significance of engaging with research design and methodology as a living element that shapes the knowledge we create. A deeper examination of the design, data collection, and analysis processes allows us to reflect the lived realities of diverse student populations more accurately. Through this lens, the research becomes a tool not just for understanding the present but for crafting better, more equitable educational practices and policies for the future.

By highlighting the importance of design and methodology, this dissertation challenges an

academic culture that often privileges outcomes over processes. I advocate for a shift toward an ethical, reflective approach to research, one that acknowledges that research is never neutral. It has the power to shape public perceptions, influence policy, and, most importantly, affect the experiences of students—particularly those who are most often underrepresented or misrepresented. In this way, the work we do as researchers can either perpetuate or dismantle systems of oppression. By prioritizing a critical examination of our methods, we have an opportunity to make a meaningful difference. Thus, circling back to the grounding of my dissertation — when I consider how we, or maybe I, come to know what we know as a field that is in part through sitting with our research design and methodology and examining the "space".

As I reflect on the process of coming to know what we know as a field, I realize that part of this knowing happens not only through the data we collect, but through how we engage with the research design and methodology itself. The design is not a static framework or a passive blueprint for collecting information; it is an active, living space where questions, assumptions, and challenges intersect. In this space, we confront the boundaries of what we understand about self-efficacy and how we can most effectively explore it.

Research methodology, in my view, provides a lens not only for how we gather data, but for how we interact with the complexities of the phenomena we study. It is within this "space" that we examine our decisions—how we define variables, how we approach integration, and how we structure our inquiry—and it is here that our assumptions are tested. Through this process of self-reflection, we gain new insights not only into the phenomenon itself, but also into the methods that shape our understanding.

By sitting with the methodology, exploring its potential and limitations, we arrive at a deeper understanding of the knowledge we produce. This exploration invites us to reconsider the boundaries of what we know and how we know it. It becomes clear that research design is not a detached or purely technical element of the dissertation, but a fundamental part of how we construct meaning from the world around us. In this way, the methodology becomes just as important to the development of the knowledge itself as the data we gather and analyze. Through an ongoing dialogue with

our research design, we develop both our methodological approach and our theoretical grounding, each influencing and shaping the other.

1.3 Motivating Real Time Analysis of Self-Efficacy

In this section, I present a vignette of my experiences as I navigated from high school to being a PhD student at Michigan State University. This vignette serves as both a narrative and reflection of the complexities I faced as I navigated the academic system(s), highlighting the moments that supported me in developing my confidence towards successfully pursuing a bachelor's degree. My vignette is not intended to be a positionality statement but rather a reflection of my lived experience. I chose this form of storytelling because it allows me to move away from the more analytical focus of a positionality statement — one that typically explores the intersections of identities, power, and privilege — and instead center the narrative of how I developed my self-efficacy towards navigating academic systems that were unfamiliar and inaccessible to me.

In a positionality statement, the emphasis is often on discussing multiple identities, examining the ways in which they intersect, and analyzing the privileges or oppressions tied to those identities [15–17]. Although this approach is valuable, I am intentionally stepping away from that framework here. I am not seeking to dissect or categorize my experiences through the lens of power dynamics or privilege, but this does not mean that I won't share the challenges associated with being a first-generation college student. Other identities may be mentioned as well such as being a student at a rural high school. Instead, I aim to share a more personal, experiential account of what it was like to be a first-generation student in a space that didn't naturally align with my background; thus, meant I had to develop my self-efficacy towards pursuing a bachelor's degree in way non-first-generation students may not necessarily have.

Moreover, I was intentional about how I positioned myself in this vignette. One of the takeaways I wanted to avoid was that "if you just work hard enough, you can do it." Thus, when writing my story, I wanted to show that others around me did work to support me in not only pursuing my bachelor's degree but also developing my confidence in pursuing a bachelor's degree. By choosing this format and approach, I hope to provide a more holistic and authentic representation of how

7

certain moments in my academic journey built my self-efficacy. As such this vignette becomes an act of reclaiming my story without being limited by the structural analysis that may occur within a positionality statement. By focusing on storytelling, I can more authentically capture the process of transformation, from being unfamiliar with academic systems to becoming an active, thriving participant in them as a result of moments supporting my degree-level self-efficacy.

As a first-generation college student, my educational journey has been shaped by both the privileges and challenges of navigating higher education without a familial roadmap. I became aware of my first-generation status early in high school when peers spoke of attending bachelor's degree-granting institutions with ease – an opportunity that felt out of reach for me. Recognizing the financial challenges my family faced, along with the high cost of college, higher education seemed like a privilege reserved for those with the means to afford it. Conversations about college felt distant, almost laughable, as I realized many of my peers had access to resources I lacked – whether it was financial support from their families, guidance on how to apply, or the social capital of parents who understood the process. These were resources my family could not offer, as neither of my parents had attended college.

However, my parents, despite their own lack of experience with higher education, consistently emphasized the value of education. They viewed it as a critical path to escaping the blue-collar jobs they had worked for much of their lives and a tool for social mobility. They often reminded me that education was the key to improving my socioeconomic status, even though they had no direct experience in navigating the complex world of higher education. Their belief in the power of education became my driving force, even though I wasn't sure I had the necessary resources, or the confidence, to succeed in a bachelor's degree program. As a result, I often found myself in tension with the expectations placed on me – both the pressure to succeed academically, but also the nagging fear of inadequacy in navigating the unknown world of higher education.

Despite this motivation, it wasn't until the end of my sophomore year of high school that I realized attending college could be a real possibility. My guidance counselor, Mr. Woods, noticing my academic success and the potential I had, encouraged me to consider college seriously and

discussed how being in the running for valedictorian could potentially support me in earning scholarships to lessen the burden of my socioeconomic status. This conversation helped me recognize that I had more control over my educational trajectory than I had originally thought, even though I still grappled with self-doubt. No one in my family had pursued a bachelor's degree, so I questioned whether I was capable of performing at that level. This doubt was compounded by the pressure I felt to succeed - not just for myself, but for my family, who had high hopes for my future.

Encouragement from mentors and teachers also played a large role in my educational pursuit. For example, during junior year, attending AP Biology with Mr. Miller was critical to hearing and learning from senior students about the transition from high school to college. I feel Mr. Miller recognized my first-generation status, and in turn, at the beginning of the class, where only three of us – two seniors and myself – were taking it, he would ask them about their plans post-graduation. Both of the seniors were planning on attending college and as such at the beginning of class, I would learn from hearing them talk about selecting colleges, the application process, and visits. These conversations facilitated through Mr. Miller played a pivotal role in demystifying the application process and offering support and knowledge I otherwise would have had. Additionally, this taught me the value and power of learning the shared lived experiences of others as I was learning how to navigate college through the lens of my peers, and without this, I would have been lost in navigating the process of transitioning to college.

Although having small class sizes and being able to have these conversations were critical to my education journal, the academic system itself values courses and grades or performances. Attending a small rural school meant I had limited access to courses beyond what is required for a high school diploma or that I would need to take AP testing. AP testing sounded extremely anxiety inducing, especially considering I was still doubting whether I deserved to go or was capable of attending college. Mr. Woods suggested that community college would be the best next step for me, given the resources available to me and my confidence in my abilities. It was a place where I could continue my education including accessing college curriculum and a safe space to try on some college level coursework and navigating a college campus. It allowed me to take ownership of my

educational trajectory, but also still hold on to the familiarity of a more accessible and less daunting environment than a bachelor-degree granting institution. This experience, while still challenging, built my confidence as I was able to be successful in multiple courses. In the courses, I gained mentorship from a history professor, who also supported building my confidence in my abilities to be successful at a bachelor-degree granting institution. Having a community college professor instill their confidence in me encouraged me that I could be successful with a college environment. Through conversations with him and reflecting about the classes I enjoyed, I soon realized I wanted to become a STEM major.

Leveraging these experiences, I ended up at Wright State University, a bachelor-degree granting institution, in which I entered as a chemistry major, my favorite class in high school. During my first year, I was really struggling with the coursework; however, I ended up taking Introductory Physics I, with a faculty member, Dr. Adrienne Traxler, who would soon become a life changing mentor for me. In her class, I was really enjoying studying for physics but also found physics challenging for myself. In high school, this was the class that I struggle the most with, causing me to not be confident in my abilities to do physics at the collegiate level, much less to be a physics major. At the end of her class, I had found community but I had also found a love for the subject; however, I was still not sure about it so I approached the subject with Dr. Traxler. She encouraged me to pursue the major and that I was capable of being a physics major. From there I became a learning assistant with her as well as took classical mechanics with her.

One day in her office I was having a conversation with her about my interest in teaching at the high school level and she had asked me if I had considered teaching at the college level. When I had thought about college level teaching, I had thought all faculty did research and I could not imagine myself being a physics lab. I shared this with her, in which she informed me that physics education research existed and she shared her research with me and that she was looking for an undergraduate student to do research with her. This sparked my interest and I thought I could try but I was not sure I would be successful in being able to perform physics education research.

After this, I found I really enjoyed the work I did with Dr. Traxler and wanted more experience

in performing research. She encouraged me to pursue an NSF funded Research Experiences for Undergraduates (REU) experience, and supported me in applying for them. This research experience encouraged me to apply for the National Science Foundation Graduate Research Fellowship as well as to Michigan State University for graduate school. This is how I ended up at Michigan State University for my graduate education as a National Science Foundation Graduate Research Fellow.

From my time as an REU student, I worked with a graduate student building a survey to measure identity within physics lab environments. I specifically operationalized physics identity statements from student survey responses, which required using qualitative coding. However, performing identity work encouraged me to become aware of my identity and the ways my identity had impacted my journey. As part of this work, in which I was in group meeting spaces and a physics education research lab learning about physics education research more broadly, I learned not only about identity, but also about the construct of self-efficacy. Because my confidence in my abilities has been a large part of my trajectory, this motivated me to think about how self-efficacy plays a role in shaping other students' trajectories. Specifically, I felt like I was lucky to have so many amazing people – Mr. Woods, Mr. Miller and the two seniors, my community college history professor, Dr. Adrienne Traxler, and Vashti and Danny (my REU faculty mentors) – support me in my journey. There are many more mentioned in the acknowledgments section who have played critical roles in supporting me.

My story highlights how a bunch of distinct moments influenced my confidence in my abilities in pursuing an academic career. From the realization in high school that being valedictorian could help me secure scholarships to support funding college, to the conversation with Adrienne encouraging me to pursue a physics bachelor's degree, my journey has been defined by pivotal moments encouraging me that I could pursue a degree. These moments have been essential in building my self-efficacy towards pursuing a degree, but I recognize that there were many other, smaller moments that contributed to my self-efficacy as well — moments that, though seemingly subtle, provided the support and encouragement needed to navigate my academic journey.

For instance, in Mr. Miller's AP Biology course, the conversations I had with him and my peers contributed to my understanding of pursuing a degree at a bachelor's granting institution, enhancing my self-efficacy in being able to obtain such a degree. These smaller moments, though not as prominently featured in my narrative, were integral in shaping my self-efficacy towards pursuing a degree. They also helped me see that, as a first-generation college student, I was capable of obtaining the knowledge and mentorship I needed to navigate higher education.

However, by focusing solely on the most salient moments — the ones that I strongly associate with my confidence towards pursuing my degree — I risk omitting the richness of smaller, yet equally impactful, instances that supported my journey. These moments add nuance to the understanding of how my self-efficacy developed. By excluding them, we limit our ability to fully appreciate the complex, multifaceted nature of students' experiences. Capturing these smaller granular moments could add a layer of nuance and richness in understanding what information students need to feel confident in their capability to successfully perform a task; in turn, provides critical information about identifying and designing moments important for students' SE development.

For example, had I journaled about the conversations I had with Mr. Miller and my peers, I may be able to tell you what specific information the seniors were sharing from their lived experiences that made me feel more confident about my ability to pursue a bachelor's degree. I remember I felt like I had a timeline and process for what pursuing a bachelor's degree looked like (e.g., when do I apply for scholarships versus when do I apply to attend a school), which made me feel more confident about my ability to apply to a college successfully.

Overall, my story shows that developing SE towards pursuing a degree is a complex process. For example, I shared a lot of the salient moments, which were moments where I strongly weighted something about those moments as important for me in developing my self-efficacy towards my physics bachelor's degree and summarized other smaller moments as one large moment (e.g., rather than highlight each conversation from Mr. Miller's class, I mostly remember that those conversations took place and I found them invaluable).

Another important piece to this is that my identity as a first generation college student heavily

impacted what moments were crucial for me to feel like I can successfully pursue and obtain a degree. My story highlights that understanding how to build moments that support students' SE, accounting for their identity may be critical to the way in which students identify moments important for their SE development as well as play a role in what students' care about developing their SE towards. Through showcasing how identity and self-efficacy intersected, this only demonstrates further how complex developing one's SE is and the information that one weighs in perceiving their confidence in their abilities.

1.4 Structure of Dissertation

This dissertation is organized into eight chapters, each building toward a deeper understanding of how self-efficacy can be explored through a carefully integrated mixed methods approach. What follows is a roadmap of how the chapters develop this work across theoretical grounding, methodological innovation, and empirical insight.

Chapter 1 established the foundation for this dissertation by grounding the research in both my personal trajectory as a researcher and my broader goals in the field. It presents the argument for a methodology-forward dissertation—one that centers the design, development, and evolution of a research process as critical to the knowledge it produces. This chapter frames methodological rigor and reflexivity not as background to findings, but as core contributions.

Chapter 2 provides the necessary conceptual grounding in self-efficacy. It explores its historical development, key theoretical components, and the various ways it has been operationalized in prior research. This chapter ensures a shared vocabulary and conceptual orientation that frames the rest of the dissertation.

Chapter 3 offers background on the methodological choices made in this study, particularly the use of mixed methods. Provides an overview of the literature that informs this design and approaches used in the dissertation, focusing on how other researchers have worked across qualitative and quantitative paradigms, and introduces key methodological commitments and tensions in which this dissertation is concerned.

Chapter 4 presents a preliminary analysis that acted as a methodological pivot point for the

work. This chapter details early insights from integrating the Experience Sampling Method (ESM) data and individualized daily journal responses, which revealed important questions about data integration, prompting the more layered methodological moves made in the chapters that follow.

Chapter 5 explores the introduction of a point of integration between the quantitative and qualitative ESM data. This addition allowed the two forms of data to "speak" to one another more directly. The resulting interplay reshaped how emerging threads in the ESM data were identified, which, in turn, influenced the individualized daily journal prompt design. This chapter shows how methodological design directly shaped data generation.

Chapter 6 builds on the prior chapter by moving from methodological integration to theoretical exploration. It demonstrates how the individualized daily journal prompt design was refined to more intentionally investigate the theoretical underpinnings of self-efficacy. This chapter reflects on how design choices impacted not only data collection but also the depth and direction of theoretical engagement.

Chapter 7 deepens the methodological contribution by focusing on the development and application of the "identifying a thread" technique—a tool for establishing integration between the ESM survey responses and the individualized daily journal responses. This chapter articulates how this technique surfaced patterns of experience that were otherwise difficult to trace, and makes a case for how mixed methods integration can serve both analytical and conceptual aims.

Chapter 8 concludes the dissertation with a reflective synthesis of the study's contributions. It revisits the dissertation's central questions and offers insights into how this research informs future work—methodologically, theoretically, and practically. The chapter also identifies open questions and opportunities for extending this work in future research.

CHAPTER 2

SELF-EFFICACY LITERATURE REVIEW

Self-efficacy (SE), first coined by Bandura, is the confidence one has in their ability to successfully perform a task [18]. Researchers have identified SE as a critical construct to study because people with higher SE towards a task are more likely to choose to perform the same or similar task in the future, as well as persist and persevere in performing that task than people with lower SE [19–21]. The importance of SE is particularly evident in fields like science education, where studies have highlighted its role in predicting students' achievement in science courses, as well their choice to major in the sciences and pursue science-based careers. Additionally, scholars have explored how SE interacts with various instructional strategies [22] as well as academic activities [23].

Building on the definition of SE, it is essential to understand the theoretical foundations that underpin this concept. The foundational theory of SE stems from Albert Bandura's social cognitive theory (1986), which emphasizes the role of cognitive, behavioral, and environmental factors in shaping human behavior. According to Bandura, SE beliefs are central to how individuals approach goals, tasks, and challenges. These beliefs influence not only the choices people make, but also their effort, persistence, and emotional reactions to challenges and setbacks. In Bandura's model, SE beliefs affect the way people think, behave, and feel, influencing their motivation and overall performance.

SE is situated within a social cognitive framework, which asserts that individuals learn through the interaction of personal, behavioral, and environmental factors [18, 24]. Within this framework, individuals are not passive recipients of environmental influences but rather active participants in shaping their own experiences. Central to this framework is the idea that SE is developed and reinforced through various sources, including mastery experiences, vicarious learning, social persuasion and physiological state. In this work, we will refer to this as the source nature of SE.

Additionally, SE is inherently dynamic [18, 25] and domain-specific [18, 26–28], meaning that an individual's belief in their ability to succeed varies depending on the task or context and can change over time. According to Bandura's social cognitive theory, SE is not a fixed trait

15

but rather a dynamic construct that evolves through individual's experiences and interactions with their environment. This variability is essential to understanding how SE functions within specific domains. The domain-specific nature of SE means that individuals can experience fluctuations in their beliefs across different tasks or contexts. A person may feel high SE in one context (e.g., biology) and may have low SE in another (e.g., physics), each domain involves distinct skill sets and experiences. As such, SE can be highly contextual, changing as individuals navigate through different situations or acquire new competencies and capabilities.

Thus, the theoretical foundations that underpin SE include the dynamic nature, domain-specific nature, and source nature of SE. As such, we have divided this section into three parts: (1) Domain-Specific Nature of SE, (2) Source Nature of SE, and (3) Dynamic Nature of SE.

2.1 Domain-Specific Nature of Self-Efficacy

A key characteristic of SE is its domain-specific nature [18, 29], which stems from the definition of SE as one's confidence in their ability to perform specific tasks. The specificity of the task influences how one perceives their SE. For example, a student might feel confident in presenting their research orally but lack confidence in writing their findings for a conference proceedings paper. Therefore, when investigating students' SE, researchers must account for the specific tasks being assessed.

Researchers have used varying levels of specificity in defining tasks related to SE [27]. In the context of physics education research, for example, there are numerous assessments that measure SE, each tailored to the researchers' focus and the specific aspects of the field being examined [30–37]. One common survey used to assess SE within physics is the Source of Self-Efficacy in Science Courses - Physics (SOSESC-P) survey [34, 38, 39]. Although this survey was not designed to focus on a particular domain of physics beyond general course-related SE, it includes items that probe SE as varying levels of specificity. For example, one item - "In-class discussions and activities helped me to relax, understand, and enjoy my experience in the course" - addressed broader, course-level engagement with the material and course, reflecting students' general sense of comfort and enjoyment within the course. In contrast, another item - "When I came across a

tough physics question, I worked at it until I solved it" - assesses task-specific SE, focusing on students' persistence and problem-solving abilities in response to a specific academic challenge. These examples illustrate how the survey captures SE at different levels of specificity, from general course-related engagement to specific task-oriented behaviors.

The long-term goal of this project is to explore transfer students' SE as they transition from an associate degree-granting institution to a bachelor's degree-granting institution. Specifically, our objective is to investigate how SE evolves across and within these contexts. This evolution of SE is shaped not only by the specificity of the tasks, but also by the broader educational context in which those tasks occur. For example, students typically begin with confidence in their ability to complete specific tasks within individual courses, reflecting task-specific SE. As they progress, they develop confidence in their ability to navigate a course and meet the requirements for that course, which constitutes course-level SE. This confidence continues to evolve, and as students approach their degree completion, they gain confidence in their ability to meet all degree requirements, which constitutes degree-level SE. Finally, students' SE extends beyond their academic journey as they begin to consider their future careers, leading to career-level SE. By breaking SE down into these distinct levels, we can see how SE evolves from the specific (task) to the broad (career), shaped by both the complexity of the tasks students encounter and the broader educational context in which they are situated.

While the above paragraph implies that SE across these levels evolves in a linear fashion, this may not be true for students. For example, a student may have high SE for completing a modern physics homework, but may identify as having low SE towards passing their modern physics course at that moment, or vice versa. This could be because students might develop confidence in one domain without directly impacting others. Another explanation could be that some students may not see the connections between tasks and higher levels (such as their course, degree or career); hence, having high SE towards a task may not necessarily build their SE in another domain. This acknowledges that SE at each level may be influenced by different factors and may not always transfer between levels; thereby, treating each domain as distinct. As such, this mixed methods

approach sets up a methodology to be able to explore these domains, so we can start to explore what factors influence students' SE in these domains, and how these domains are related to one another. In addition, this leaves room for diverse student experiences, suggesting that students' domain-specific SE may be shaped differently depending on their individual contexts and perspectives. In the following passages, we will explore each of these domains in more detail.

2.1.1 Task-Specific Self-Efficacy

This domain of SE refers to an individual's confidence in their ability to successfully complete a particular task or activity [18, 25]. Academic SE refers to a student's overall confidence in their ability to succeed in academic tasks, such as studying versus note taking versus writing [40]. In contrast, task-specific SE focuses on confidence in completing particular, often more specialized, tasks. For example, the technical writing required for a chemistry lab report might differ from that in a physics lab report because of the specific content involved, or writing in STEM courses might be very different from the writing required in an English class. Research has shown that students' SE towards a particular discipline, chemistry, differs from their overall academic SE. In one study, students' chemistry SE increased across the semester, while their academic SE decreased [41]. The academic SE scale they used was modified from the *Motivated Strategies for Learning Questionnaire (MSLQ)* [42], in which the 8 items they used were related to the students' abilities to perform broad tasks within their courses (e.g., can understand the most complex material presented in my courses, can understand the basic concepts taught in my courses, can master the skills being taught in my courses). Notably, the way chemistry SE was measured in this study involved assessing students' understanding and comfort with specific chemistry concepts, such as interpreting chemical equations and understanding the properties of elements. Thus, this demonstrates how the broader academic tasks differ from other more specialized tasks.

This chemistry scale focused on tasks related to understanding particular concepts within the discipline of chemistry, highlighting one way in which the task can be specialized. Similarly, in physics education, researchers have examined SE through problem-solving tasks, considering how confident students feel in performing specific problem-solving tasks [32, 43]. For instance, Durk et

al. (2020) explored how students' SE was linked to their ability to calculate physical phenomena, while Hung and Wu (2020) investigate how SE varies depending on the type of representation used (e.g., numerical versus symbolic). These studies underscore how task-specific SE in physics is shaped by the kinds of tasks students are asked to perform and how they approach them.

Miller and colleagues [44–46] highlighted another way to explore task-specific SE by examining how peer discussions influence students' confidence in switching their responses when answering "ConceptTests," a type of conceptual multiple-choice question. This example underscores that when examining task-specific SE, researchers must consider the diversity of tasks students encounter, such as clicker questions, ConceptTests, homework, and projects, to name a few. These tasks each require distinct methods for assessing students' confidence in their abilities, as their SE can be shaped by the nature of the activity itself.

For example, Fencl and Scheel explored how various instructional strategies shaped students' SE in response to specific course tasks, illustrating the importance of adapting pedagogical approaches to the task at hand [34]. Additionally, both studies emphasize that instructional strategies—such as peer discussion—can directly influence students' task-specific SE, providing valuable insights into how teaching methods shape students' confidence in completing particular academic activities.

An alternative approach to exploring task-specific SE within physics is to recognize that depending on the course type, educators may be interested in how different skills influence students' SE. This approach involves selecting particular skills of interest and investigating how these skills are related to students' confidence in their physics abilities. One study explored students' physics lab SE, which was defined as one's confidence in their abilities to complete laboratory tasks [47]. The items they proposed to explore students' SE in relation to physics laboratory tasks are provided below.

- I feel I could critique a laboratory report written by another students.
- I am confident that I could read the procedures for an experiment and conduct the experiment on my own.
- I feel I am capable of helping my classmates with physics in the laboratory.

These items capture a range of skills critical for students' success in the laboratory, providing insight into how task-specific SE can be related to the development or use of particular skills. Through this passage, we have seen the different approaches used to explore task-specific SE, depending on the focus of the specific tasks and the aims of the research. In this work, we will discuss how tasks are defined in the context of the codebook later in the paper.

2.1.2 Course-Level Self-Efficacy

This domain of SE refers to one's confidence in their ability to successfully pass a course [42, 48, 49]. A key approach to assessing course-level SE involves using items specifically designed to assess students' beliefs about their abilities to pass the course, including their abilities to complete the course requirements. For example, the *Motivated Strategies for Learning Questionnaire* (MSLQ) was designed for use in college settings and includes items that focus on students' goals and value beliefs for a course, their beliefs about their abilities to succeed in a course, and their anxiety in relation to tests within a course [42]. These dimensions were critical as students' SE was positively related to greater cognitive engagement and actual academic performance (Garcia and Pintrich, 1995). These items were carefully tailored to explore students' course-level SE, shared below. As this allows researchers to assess students' course-level SE, physics education researchers have modified these for physics courses [49].

- I believe that I will receive an excellent grade in this class.
- I'm certain I can understand the most difficult material presented in the readings for this course.
- I'm confident I can understand the basic concepts taught in this course.
- I'm confident I can understand the most complex material presented by the instructor in this course.
- I'm confident I can do an excellent job on the assignments and tests in this course.
- I'm certain I can master the skills being taught in this class.
- Considering the difficulty of this course, the teacher, and my skills, I think I will do well in this class.

Although some researchers have developed course-level SE measures, some may choose to measure subject-based SE, which is more broadly focused on students' confidence in their ability to learn or master the content of a subject, rather than students' ability to complete a specific course. In some cases, researchers use subject-based SE measures to assess students' SE throughout the course (i.e., they measured their subject-based SE at the beginning and end of the semester) [50, 51]. For example, in one study, they used the Achievement Motivation Questionnaire, which includes a SE scale that assesses students' confidence in learning physics, not necessarily passing a specific course [48].

In determining what scale or whether to use course-level versus subject-based SE is dependent on the research questions and nature of the research, and is thus up to the researcher to determine. Our choice to use specific course-level items is in alignment with Bandura's (1997) argument that SE measures should be as closely aligned with the domain in question as possible [18]. He posited that the more directly a measure reflects the specific context or task being studied, the better it will predict individuals' behaviors and abilities within that domain. In the case of course-level SE, using items focused on students' confidence in their abilities to pass a course provides a more accurate measure of students' confidence in their ability to succeed in that course. By ensuring that the assessment is tightly aligned with the domain of academic coursework, we increase the likelihood that the measure will reflect students' true beliefs about their ability to succeed in the course, rather than their general SE or subject-specific SE.

While alternative measures, such as subject-based SE, can provide valuable insights into students' confidence in mastering broader subject material, these measures may not capture the nuances of students' confidence in successfully completing the specific tasks and challenges presented within a course. By focusing on course-specific SE, researchers can more accurately assess students' preparedness for, and beliefs about, the course requirements, leading to more meaningful conclusions about how SE influences students' engagement, performance, and persistence in the course.

2.1.3   Degree-Level Self-Efficacy

This domain refers to an individual's confidence in their ability to successfully complete all the requirements necessary to obtain their degree [52]. This domain can be tricky to find in the literature because there are other terms to discuss students' degrees such as their major (Lent et al., 1986) or educational requirements [53]. The *Self-Efficacy for Technical/Scientific Fields* assessed SE by asking participants to indicate if they believed in their ability to successfully complete educational requirements and work duties in the science and engineering fields. Informed through Betz and Hacket's study, Lent and colleagues asked their participants to rate their ability to perform specific tasks related to academic success in science and engineering majors (Lent et al., 1986). This particular measure of SE is referred to as *Self-Efficacy for Academic Milestones* and has been used in multiple other studies (Lent et al., 2007, 2008; Jones et al., 2010; Jaeger and Adair, 2018). Directions for this scale read as: "The following is a list of major steps along the way to completing the engineering degree. Please indicate how much confidence you have in your ability to complete each of these steps in relation to the engineering major that you are most likely to pursue" The items are shared below.

- How much confidence do you have in your ability to ...

    - complete all of the semester 1 degree units with grades of B or better?
    - excel in your engineering studies over the next semester?
    - excel in your engineering studies over the next two semesters?
    - complete your engineering studies with an overall GPA of 3.8 or better?

Although these items focus on academic milestones as the work intended, these items are limited in their scope, as they primarily reflect short-term academic success rather than broadly covering degree-level requirements. Edzie (2014) expanded on others ways to explore students' degree-level SE in her work of exploring factors that influence female students' pursuit and persistence within STEM degree programs. Through mixed methods research, Edzie conducted focus group

22

interviews to explore female students experiences. Some of the focus group questions used in this study are in line with broader exploration of degree-level SE:

- Was there a particular moment that stands out for you when you decided that this was the right major for you?
- What do you think educators can do to encourage more females in math and science - STEM?
- How do you feel if you get a B on a homework or exam after you worked really hard to prepare for it?
- What motivates you to persist in this major?

These questions touch on aspects of SE related to motivation, persistence, and emotional experiences - critical components of degree-level SE. Within her mixed methods research, Edzie also adapted SE scales, specifically, those from the *Motivated Strategies for Learning Questionnaire* [42], to explore students' confidence in their ability to succeed in STEM majors. By adapting these items to ask about "STEM major" rather than "class," Edzie's researcher aimed to better capture students' confidence in their ability to complete degree requirements within their chosen STEM discipline.

Importantly, recent research suggests that hands-on, immersive research experiences can potentially enhance students' degree-level SE (Amelink et al., 2015). These experiences can provide students with a tangible sense of what it is like to succeed in the field, increasing their confidence in completing the academic milestones necessary to achieve their degree. For example, a hands-on research experience not only helps students build skills but also offers direct evidence of their ability to solve complex problems and contribute meaningfully to research teams. This can potentially improve students' belief in their ability to complete key tasks in their studies required to obtain their degrees.

As students gain first-hand experience and develop their technical skills, their SE grows, not just in terms of academic performance but also in their ability to meet the broader demands of their degree program. This increased confidence in their skills and abilities fosters persistence and

motivation, ultimately contributing to their academic success and degree completion.

2.1.4   Career-Level Self-Efficacy

There are four ways in which SE has been linked to career within the literature. Betz and Hackett (1981) used social cognitive theory as a lens through which to explore the differences in how men and women perceive their confidence in their ability to succeed in various career-related tasks or fields [53]. The kind of career SE could encompass skills, tasks, and roles within the profession in addition to the decision-making process about which career to pursue. Then, others have conceptualized career SE in relation to the decision-making process specifically. For example, Gordon (1998) defined career decidedness as the level of confidence regarding a particular career-related decision whereas others refer to career decision SE as the confidence one has in their abilities to make career-related decisions and produce positive outcomes (Betz and Hackett, 1986; Kim et al., 2016). Then, Restubog and colleagues defined career SE as one's confidence in their ability to succeed in an academic program, which is part of the pursuit of a specific career (2010). Then, another way to conceptualize SE in relation to career is career-level SE, in which this is one's confidence in their abilities to complete the educational requirements and job duties for their intended future profession or career [49].

2.2   Source Nature of Self-Efficacy

An individual's SE is influenced by how they interpret information from various sources. Bandura identified four primary sources of SE: mastery experiences, vicarious learning experiences, social persuasion experiences, and physiological state experiences (REF). These four sources have been widely recognized as fundamental in shaping SE, which is why we have prioritized crafting individualized daily journal prompts that align specifically with these sources. Although the focus of this paper is on these four primary sources, we also recognize that other factors such as self-regulation (Webb-Williams, 2017) may influence students' SE in real-time experiences. However, for this present work, we will focus on designing prompts in alignment with the four primary sources of SE.

2.2.1 Mastery Experiences

Personal experiences can be identified as successful or failure, depending on the interpretation of the individual. They can select these successful or failed experiences to evaluate their confidence in performing a future task [18, 19]. When individuals succeed in completing a task, this success reinforces their belief in their ability to handle the same or similar tasks in the future; conversely, a failure can threaten one's SE [54]. The way an individual interprets this failure can impact whether the failure threatens their SE. For example, if the individual interprets that they put a lot of effort into the task and then failed, this could threaten their SE [20]. Furthermore, if the individual sees their failure as an indication of their lack of ability, this may also threaten their SE. However, if the individual interprets the failure as an opportunity to learn or grow, then the failure may not threaten their SE. This shows how the interpretation of the failure of the individuals can impact their SE, and this idea of interpretation also holds for successes. For example, when individuals overcome particularly difficult tasks, this can support building their SE [18, 20]. Mastery experiences are crucial for individuals as this means that they can use their own experiences when evaluating their SE toward doing a task. Given that mastery experiences allow people to leverage their own experiences, researchers have claimed that this is the most influential source for interpreting one's SE (Peker, 2016; Bandura,). Although this has been argued, research has shown that different populations of students may prioritize different sources of SE [19, 21, 55–57]. This suggests that while mastery experiences are undoubtedly important, they may not be the primary driver of SE for all student populations.

2.2.2 Vicarious Learning Experiences

When evaluating confidence in a task, people can observe others performing the same or similar task, so they use this information to evaluate their own likelihood of success in the same or similar task [18, 19]. For example, in academics, when determining how to use grades to assess their confidence, students may compare their own grades to the graders of their peers to get a baseline for how they did. In this case, a student may have received a B on their exam, and when they find out that their peers received lower grades, this could boost their SE. However, if that student finds

25

out that their peers received higher grades, this could threaten their SE. This example demonstrates how various learning experiences can threaten or increase students' SE depending on how they evaluate their performance compared to others. The "other" that students compare themselves to matters, as well. Using the previous example, a student that they see as normally scoring A's on the exam may not be the person they choose to compare themselves to, or they could decide to use this person as the comparison. In this case, because the student is comparing themselves to someone who performs better on exams than themselves, then they may be likely to feel less confident in their own abilities to perform that task (Bandura, 1993).

2.2.3 Social Persuasion Experiences

Social persuasion refers to verbal and nonverbal messages from others that influence one's belief in their ability to perform a specific task [18, 19]. These messages can be either positive negative in nature. When individuals receive supportive or encouraging messages, it can boost their SE, making them more confident in their capabilities. Conversely, when individuals receive messages that imply doubt or disbelief in their abilities, this can threaten their SE, leading them to feel less confident in their capabilities.

When conceptualizing SE, some researchers have also focused on the experience of receiving feedback from others (Bandura, ; Hutchinson-Green et al., 2008). When students are uncertain about their ability to perform a task, they often turn to others for feedback, particularly if they lack direct knowledge or experiences with the task at hand (Bandura, ; Sawtelle et al., 2012). Social persuasion is often studied by asking individuals to rate whether and how often they receive encouraging messages from those around them (Usher, 2005). These messages can come in many forms, ranging from verbal praise to subtle non-verbal cues.

A critical aspect of social persuasion is the interpretation of these messages. The impact of feedback depends on how individuals perceive and interpret these messages, which is influenced by factors such as the credibility of the person delivering the message [19]. For example, students may be more likely to accept and be positively influenced by feedback from individuals they trust, such as faculty members or mentors, than from less familiar or less credible sources. The credibility of

the messenger is shaped by factors such as their sincerity, status, and familiarity with the student (Bandura, 1997; Usher, 2005). For example, when students have a positive, trusting relationship with their instructor or a faculty member, they may be more likely to interpret the feedback as supportive and confidence building (Eva et al., 2012; Usher, 2005).

While much of the focus in social persuasion research is on feedback from others, it is also important to recognize that messages can be communicated more subtly, including through nonverbal cues and cultural narratives (Sawtelle et al., 2012; Usher, 2005). For example, body language, facial expressions, or tone of voice of the person delivering feedback can significantly influence how the message is perceived. In addition, students are often exposed to cultural narratives or societal expectations that can implicitly affect their SE, even if they do not consciously acknowledge them. These implicit forms of social persuasion can be challenging to study, as they may not be overtly recognized or may be dismissed by students in favor of more direct forms of feedback (Usher, 2005).

An ongoing challenge in the study of social persuasion is understanding which messages individuals choose to internalize and which ones they dismiss. Given the variety of social sources and contexts in which messages are delivered, it is still unclear which types of social persuasion are most likely to have a lasting impact on SE (Usher, 2005). Some feedback may have an immediate effect but little long-term influence, while other experiences may resonate with individuals and shape their beliefs about their abilities for much longer. Research is needed to explore how people prioritize and interpret different social influences, as well as how these messages interact with other sources of SE (e.g., mastery experiences or vicarious experiences).

2.2.4 Physiological State Experiences

Physiological State: People may use their physiological or affective states to evaluate their confidence in their abilities [18, 19]. Part of physiological or affective states includes accounting for the emotional states of individuals, in which people can use their emotions in a moment to evaluate their confidence in their abilities. For example, when people feel nervous about a particular task, they may see this nervousness as a sign that they are lacking in ability. Other factors

that can influence one's physiological and affective states are mood, physical strength, and levels of distress [58] . For example, when an individual is in a good mood, they can use their positive memories around the task, enhancing their SE, whereas when an individual is in a bad mood, this can have the opposite effect, threatening one's SE (Usher, 2005). Furthermore, the intensity of feelings or arousal can influence the degree to which the physiological state affects SE (Usher, 2005). The more intense a feeling is during a task, the more likely this physiological state may affect one's SE [18].

In terms of examining physiological state, student anxiety within academic contexts has been commonly assessed (Gainor and Lent, 1998; Lent et al., 1991, 1996; Lopez and Lent, 1992; Britner, 2008; Britner and Pajares, 2006; Smith, 2001; Usher and Pajares, 2009; Sawtelle et al., 2012). However, little attention has been given to how positive physiological states have affected individuals' confidence in their capabilities [19, 20]. Physiological state can be a difficult source to investigate in part because individuals can interpret their emotions differently in the sense of how they interpret what their emotions are telling them and in part because physiological states are a part of the human experience, so disentangling this source from the other sources of SE may be difficult [19].

2.3 The Dynamic Nature of Self-Efficacy

SE is inherently dynamic, fluctuating from moment to moment based on a variety of internal and external factors. Unlike fixed traits or stable abilities, SE reflects the changing beliefs individuals hold about their capabilities in specific situations. Bandura initially conceptualized SE as the belief in one's ability to execute actions required to achieve particular goals [59]. This belief, however, is not static; it is shaped by ongoing experiences, feedback, and situational contexts. For example, a person's sense of SE may rise after a small success, but may also diminish following a failure or when faced with a challenging task (Schunk, 1989). These fluctuations occur in real-time, highlighting how SE is continually shaped by momentary experiences, social feedback, and the challenges or successes encountered in a given context [28]. As such, SE is a fluid and context-dependent construct, constantly evolving as individuals navigate their day-to-day experiences.

2.4 Contextualizing the Self-Efficacy Literature to Within the Dissertation Framework

The overarching goal of this dissertation is to design a mixed methods approach that can explore the theoretical foundations of SE, specifically its dynamic, domain-specific, and source-specific nature. These foundational characteristics of SE form the core focus of our study and are examined through multiple methodological lenses. Our primary quantitative approach—Experience Sampling Method (ESM) surveys—addresses the domain-specific nature of SE by measuring task-specific, course-level, and career-level SE at different points in time. By capturing students' SE across various academic and career-related domains, the ESM surveys provide a detailed snapshot of how these beliefs fluctuate based on the context of the task at hand.

In addition, the individualized daily journal prompts complement this exploration by allowing for a deeper and more reflective examination of domain-specific SE. Through these journal entries, we can capture the nuances of SE as it unfolds across the day, adding richness to the quantitative data and making visible how these prompts expand our understanding of domain-specific SE. For example, students can reflect on their feelings of SE in relation to specific academic tasks, providing insight into how these beliefs are shaped by ongoing experiences and shifting challenges.

To explore the dynamic nature of SE, we leverage the ESM surveys' ability to measure SE at different moments throughout the day. This allows us to capture fluctuations in students' SE in real time—such as a student's confidence before and after studying for or taking an exam. This dynamic measurement approach enables us to trace how SE beliefs evolve in response to different experiences and challenges.

Finally, in exploring the source-specific nature of SE, we turn to the individualized daily journal reflections, which give students the opportunity to reflect on the sources influencing their beliefs in real time. Through these journal entries, we gain insight into how various sources—such as personal experiences, feedback from others, or task challenges—affect students' SE. A central aim of this dissertation is to make the ways in which these individualized journal prompts contribute to exploring the source-specific nature of SE visible and transparent, showcasing the methodological work being done to capture these influences.

In sum, this dissertation employs a mixed methods framework to investigate the dynamic, domain-specific, and source-specific nature of SE, making visible the ways in which different data sources interact to offer a comprehensive understanding of how SE is shaped by both internal and external factors over time.

CHAPTER 3

MIXED METHODS LITERATURE REVIEW

This study uses a mixed methods research design to explore students' SE as it unfolds in real time. While mixed methods is often broadly defined as the intentional combination of qualitative and quantitative approaches [60], such definitions risk oversimplifying the epistemological and interpretive complexity involved. As Sandelowski (2003) notes, the field often reduces these methods to "words versus numbers," overlooking how data are conceptualized, contextualized, and used in practice [61]. To understand the development of SE, particularly in moment-to-moment contexts, this study required a research design capable of both capturing and interpreting the *dynamic*, *context-sensitive*, and *multifaceted nature* of students' experiences — something mixed methods, when rigorously integrated, is uniquely suited to do.

3.1 The Evolving Conversation Around Integration

At the heart of mixed methods inquiry is the concept of integration, widely considered the defining feature of this approach [62–64]. Integration is not merely a procedural step but a methodological commitment to generate insights that neither strand could provide independently [63]. However, as noted in both foundational [65] and more recent work [66], integration is often less clearly articulated in practice. In particular, researchers point to a lack of practical guidance around how integration is enacted and how its quality can be assessed [65, 66].

In the early stages of mixed methods literature, integration was largely framed through the lens of triangulation—a strategy aimed at validating findings by demonstrating convergence across data sources [65] (Denzin, 1978). While this framing positioned triangulation as a core motivation for mixed methods, subsequent scholarship has questioned whether it fully captures the breadth and potential of integrated inquiry. Specifically, the term triangulation has become overly associated with validation, which can obscure the broader interpretive and generative purposes of integration [67–69].

In response, scholars have expanded the conceptual landscape of integration. Greene et al.'s (1989) influential typology—triangulation, complementarity, development, initiation, and expan-

31

sion—helped recast integration as a vehicle for deepening insights, provoking new questions, and surfacing contradictions, rather than simply confirming prior claims [65]. More recent definitions center integration as the dynamic interaction or "dialogue" between qualitative and quantitative strands [69, 70], with the shared goal of generating insights that could not be achieved through either approach alone [71].

To support this goal, scholars have offered a range of frameworks, models, and strategies for achieving integration across different stages of the research process [62–64, 69, 70]. For example, Fetters, Curry and Creswell (2013) describe three integration techniques [62]:

- Merging, where data are brought together during analysis;
- Connecting, where one strand informs the sampling or data collection of the other; and
- Embedding, where one type of data is nested within the logic of the larger design.

These strategies provide concrete ways to operationalize integration and offer flexible approaches for facilitating communication between the quantitative and qualitative strands. However, these strategies represent more general approaches to how the strands might be integrated, and they may need to be further specified or adapted depending on the nuances of the research context and the specific research questions. Other scholars have developed additional strategies that further specify how communication between the strands can occur more effectively at different stages of research.

For example, O'Cathain, Murphy, and Nicholl (2010) provide additional insights into Fetters et al.'s concept of "connecting" by highlighting strategies like "Following the Thread" [69]. This approach involves identifying a theme, concept, or question emerging from one data strand and using it to inform or guide the inquiry within the other strand [69, 72]. By tracing these emergent threads across data types, researchers can deepen their understanding and uncover new insights that might not have been possible through independent analysis of each strand.

Building on the strategies proposed by Fetters, Curry, and Creswell (2013) [62], Vicki Plano Clark (2019) has further advanced our understanding of how integration can be conceptualized

and implemented across different phases of the research process [70]. Plano Clark emphasizes the importance of intentional integration at both the methodological and practical levels, noting that integration is not just a technical step but also a critical design principle that shapes how research unfolds. Her work provides several strategies that can be incorporated into the framework of connecting, merging, and embedding. Plano Clark (2019) highlights four key strategies for achieving integration [70]:

1. Formulating Integrative Research Questions: This strategy emphasizes the importance of designing research questions that invite both qualitative and quantitative investigation. These questions act as a bridge between the two strands, ensuring that both data types are aligned and contribute meaningfully to addressing the overall research goals.

2. Aligning Quantitative and Qualitative Data Sources: This strategy calls for the careful selection of data sources that complement each other and provide a comprehensive view of the phenomenon under investigation. For example, qualitative data can be used to explore the nuances and underlying factors that quantitative data might not fully capture.

3. Identifying Points of Integration: This strategy involves identifying specific moments or locations in the study where the two strands of data intersect. Whether during data collection, analysis, or interpretation, recognizing these points allows researchers to leverage the strengths of both qualitative and quantitative methods to generate deeper insights.

4. Developing Joint Displays and Mixed Interpretations: This strategy focuses on how the results from each strand can be presented in an integrated way, such as through joint displays that allow for a visual representation of the data. This approach helps to demonstrate how the two strands complement each other and how the integration contributes to the overall findings.

Plano Clark's framework provides a more structured approach to integration, reinforcing the idea that integration is not just about combining data but about strategically designing research to

allow for meaningful interaction between data types.

These approaches provide language and structure for operationalizing integration across different phases of the research process. By offering clear, actionable strategies such as merging, connecting, and embedding, these frameworks allow researchers to systematically apply integration throughout the design, data collection, and analysis stages. This operationalization not only facilitates the practical execution of mixed methods research but also underscores the importance of intentionally aligning the two strands of data in ways that enhance the overall research process. Integration, as Zhou and Wu (2022) argue, should be understood as both a product and a process [66]. As a process, it unfolds over the course of the study, involving ongoing interactions between the qualitative and quantitative strands. As a product, integration results in tangible outcomes that are the visible evidence of how the strands have been meaningfully combined to generate deeper insights. Through these strategies, integration becomes more than just a procedural step; it is a core methodological commitment that shapes the interpretation of findings and contributes to a more comprehensive understanding of the phenomenon under investigation.

3.2 Operationalizing Integration Across the Dissertation

In this study, we build on this expanded view of integration by using what Fetters et al. (2013) term a "connecting" strategy, in which one data strand informs the design or collection of the other [62]. Specifically, our use of Experience Sampling Method (ESM) surveys provided both open- and closed-ended responses that were used to generate individualized reflection prompts for student journaling. This approach operationalizes the analytic technique of "following a thread" [69, 72, 73], a strategy wherein a theme or question that emerges in one data strand is traced and developed in another.

By embedding this practice into the data collection process itself, we extend the use of "following a thread" beyond its traditional application during analysis. In our design, analysis begins during data collection—as researchers interpret ESM responses to write individualized prompts—and continues after data collection has concluded. This dual-phase application not only highlights the iterative nature of integration but also establishes a deliberate point of communication between the

qualitative (open-ended) and quantitative (closed-ended) components of the ESM survey. In doing so, we create a novel point of integration that is both analytic and procedural.

This methodological choice responds to a growing call in the literature for greater transparency and intentionality in integration [62, 74]. Researchers have noted that integration is frequently described at the level of design but remains difficult to identify in published work. Zhou and Wu (2022) argue that integration should be understood as both a process and a product—something that unfolds across a study and must be made visible through evidence [66]. In our study, this meant making explicit how student survey responses informed individualized daily journal prompts and how these prompts created interpretive bridges across data types. Integration, then, became both an epistemological stance and a practical design principle.

In Chapter 5, we treat the act of "following a thread" as a proof of concept. By weaving together the different data sources—open-ended task survey responses, open-ended course survey responses, and task-specific self-efficacy items—we assume that this integration is happening because we have purposefully constructed prompts to facilitate that connection. At this stage, following a thread is the methodological design that enables integration to occur, but it remains somewhat abstract, based on the assumption that if we bring these data points together, integration will naturally follow.

Chapter 7, however, pushes this idea further by introducing the concept of evidence. Here, I demonstrate that integration doesn't simply happen because we try to follow a thread from ESM data to daily journal prompts. Instead, we can actually see the integration in the data itself. By tracing the movement of specific themes and language—such as task-specific SE items reappearing in students' journal responses—we provide concrete evidence that the integration has taken place. This chapter moves beyond the conceptual argument for integration, offering an empirical basis to show that the linkage between ESM and journal data is not only intentional but also verifiable.

Thus, while Chapter 5 assumes that integration occurs through the process of following a thread, Chapter 7 provides the proof: integration is visible in the evidence, substantiating the claim that the connection between data sources is not only a theoretical construct, but something that can be empirically traced and validated.

In sum, this mixed methods study contributes to the literature in two key ways:

1. by operationalizing "following a thread" during data collection to create a novel integration point, and

2. by making the logic and labor of integration visible, transparent, and methodologically grounded—responding to recent calls for clarity around what integration means in practice and how it supports the interpretive aims of mixed methods research.

CHAPTER 4

A REFLECTIVE APPROACH TO STUDYING SELF-EFFICACY IN REAL TIME

4.1 Introduction

This chapter outlines the research design, methodology, and methods employed in this dissertation. The goal of this chapter is to provide a top-level description of how the integration of quantitative and qualitative components of our mixed methods research design can support understanding a student's real-time academic experience, particularly focusing on how a particular moment in a student's course plays a role in their self-efficacy. This supports the larger goal of the dissertation, which is to demonstrate how the integration of quantitative and qualitative components provides a robust framework for exploring student experiences in real time within STEM education, particularly focusing on how these moments play a role in students' self-efficacy as they navigate their academic journey.

I begin by presenting an overview of the mixed methods approach, outlining the rationale for using an explanatory sequential design. Then, I describe the quantitative data source and method used, followed by a discussion of its design and implementation. Next, I highlight the qualitative data source and collection method, along with a discussion of its design and implementation. After presenting the general approach, I detail the Fall 2020 data collection, including the key decisions that shaped this first iteration of the design. This chapter also offers a preliminary analysis of the Fall 2020 data through the lens of Jane's (pseudonym) experience, emphasizing the value of a mixed methods approach in studying students' self-efficacy in real-time contexts. Finally, this analysis serves as a critical point of reflection, illustrating how the data from this phase directly informed the adjustments made to the second iteration of which happens in the Fall 2021 data collection.

Jane's experiences, characterized by fluctuations in her self-efficacy scores and her enrollment in a range of STEM courses, serve as a key case study throughout this chapter. These experiences not only illustrate the kinds of characteristics of self-efficacy we aim to capture, but also underscore the limitations of the research design implemented in the Fall 2020 iteration in doing so. The insights gained from Jane's preliminary data analysis in Fall 2020 led to three primary modifications to the

Fall 2021 data collection. These modifications were based on the limitations identified through the preliminary analysis and will be discussed in detail throughout this chapter. Furthermore, the preliminary analysis helped shape each of the research narratives explored in Chapters 5-7.

By setting the stage with this preliminary analysis, this chapter demonstrates the iterative process of this research investigation, revealing how the design, methodology, and data collection were refined across the two iterations (Fall 2020 and Fall 2021) to capture a comprehensive understanding of the complex experiences that influence the self-efficacy of students. Ultimately, the discussions in this chapter provide the foundational context for the findings and analysis presented in the subsequent chapters.

4.2   Overview of the Research Design

4.2.1   Rationale for Explanatory Sequential Mixed Methods Design

The decision to use an explanatory sequential mixed methods research design was driven by the way self-efficacy has traditionally been explored in the existing literature [35, 39, 50, 51, 75, 76]. This design emerged as a fitting approach to address the research goals of the broader project, which is to be able to explore how students' real-time experiences are interacting with their self-efficacy, as well as improving upon the more traditional data collection methods. In the following, I will explain how this design aligns with the study's objectives, offering deeper insights into students' domain-specific self-efficacy by combining both quantitative and qualitative data.

The purpose of this methodology is to explore how real-time experiences influence students' self-efficacy, and conversely, how self-efficacy impacts students' real-time experiences. To investigate how students develop their self-efficacy over time, a method is needed to capture fluctuations in self-efficacy from one lived experience to another. Specifically, this involves designing a methodology to examine how self-efficacy changes from one moment (e.g., studying for a quiz) to another (e.g., taking the quiz).

Quantitative methods, such as surveys, have been used to track shifts in self-efficacy over time [35, 39, 50, 51, 75, 76]. Although these methods can measure changes in self-efficacy, they often lack the depth necessary to understand the underlying reasons for these changes [39]. This

limitation suggests that a more nuanced exploration is required.

Qualitative methods are well suited to capture the complexity of individuals' experiences in response to specific events [77, 78]. Using qualitative methods, researchers can capture the complexities of students' thoughts, feelings, and behaviors in response to specific events, providing richer, more detailed data that sheds light on how and why self-efficacy fluctuates in relation to particular tasks or moments. However, qualitative methods alone cannot capture the real-time fluctuations in self-efficacy that quantitative methods can, highlighting the need for a more integrated approach.

The explanatory sequential mixed methods design, as shown in Figure 4.1, is particularly suitable for this investigation because the structure of this design first pinpoints self-efficacy signals through quantitative analysis before leveraging qualitative insights to uncover the underlying explanations or contextual information. A signal typically refers to a momentary indication of change or fluctuation in a measure construct - here, self-efficacy - captured through repeated surveys. These signals can be shifts in responses over time that suggest meaningful variations in an individual's self-efficacy. Thus, in the quantitative analysis we leveraged analytic observations of the students' survey data to examine for characteristic changes that could manifest as (1) a noticeable drop or increase in a student's self-efficacy rating in response to specific tasks, (2) a pattern where self-efficacy is consistently lower at certain times of the day or after specific academic experiences, or (3) deviation from student's typical response patterns that may suggest that a moment of their self-efficacy is lower or higher at that particular moment. Then, to leverage the structured approach of an explanatory sequential mixed methods design, the quantitative observations can then meaningfully inform a individualized daily journal prompt, allowing for a deeper understanding of the factors influencing students' self-efficacy as those specific moments.

Integrating both quantitative and qualitative data provides a more comprehensive understanding of the contextual factors influencing self-efficacy, insights that would be difficult to capture through a single method alone [64]. Although the quantitative phase identifies broad patterns and variations in self-efficacy, the qualitative phase offers deeper, individualized narratives that help explain the

Figure 4.1 This figure illustrates the phases of an explanatory sequential mixed methods research design. This process is structured in three key phases: (1) quantitative data source, (2) quantitative data observations, and (3) qualitative data source. Thereby, this design highlights that the quantitative data observations occurs between the quantitative data collection and qualitative data collection, with the results from the quantitative data observations informing the subsequent qualitative data collection and analysis.

emergence of these patterns. This combination allows for a more nuanced interpretation of how contextual variables shape self-efficacy experiences over time.

The explanatory sequential mixed methods design is an established approach in research and is particularly effective for studies aiming to combine the strengths of both qualitative and quantitative methods [64, 71]. This approach aligns with the study's objective of exploring events that impact students' self-efficacy in real time, identifying fluctuations through quantitative data, and then using individualized daily journal reflections to uncover the underlying factors driving these changes.

4.2.2   Overview of the Quantitative Source: Experience Sampling Method (ESM)

For the design, we used the ESM as the quantitative approach to collect data on domain-specific self-efficacy of students at various time intervals throughout the day. The ESM is broadly a technique used to capture real-time, in-the-moment experiences of individuals [79–81]. ESM studies typically involve signaling participants to answer a combination of open and closed-ended questions at random times throughout their day. These prompts are delivered through a signaling device, typically a smartphone application. The data collected provides a rich, context-specific snapshot of the ongoing thoughts, feelings, and behaviors of the participants, reflecting their experience as they navigate daily life [82].

A major advantage of ESM is its ability to reduce retrospective biases [80, 83, 84]. By collecting data in real time, we reduce the limitations of memory recall and provide an accurate and dynamic understanding of how self-efficacy fluctuates in everyday contexts. The proximity of ESM to the

lived experiences of the participants allows for a finer-grained analysis of how self-efficacy evolves throughout the day and with respect to different situations. Moreover, the repeated measures nature of ESM enables us to track intra- and inter-individual changes over time [80, 84].

Intra-individual analysis involves examining how self-efficacy fluctuates within an individual over time, providing information on the personal, situational, and contextual factors that influence these fluctuations. Inter-individual analysis, on the other hand, focuses on comparing these fluctuations across different students, allowing us to identify broader patterns and themes in self-efficacy within academic contexts. Together, these analyses provide a nuanced understanding of how self-efficacy fluctuations both on an individual level and across a larger population. In this dissertation, we will focus on examining fluctuations in self-efficacy within students rather than across students.

4.2.3   ESM Survey Design and Implementation

This section explains the specific aspects of the ESM survey used in the research design, including the domains measured, the structure of the survey, the survey items, and the methods used for data collection. To implement ESM to explore self-efficacy in real time, we used smartphone applications to signal participants and prompt them to complete brief surveys about their domain-specific self-efficacy. In Fall 2020, we used the LifeData application [85], and in Fall 2021, we used the Expiwell application [86]. These platforms enabled us to deliver the surveys in a timely manner, ensuring real-time data collection.

In the survey, students were asked open- and closed-ended items. The open-ended questions were used to support the position of the individual's experiences within an activity, course, or some kind of context. For example, the open-ended item of "What is the main thing you are doing right now" is there for the participants to describe the activity they are working on when completing the survey. The closed-ended items were then measures of domain-specific self-efficacy: task-specific, course-level, and career-level self-efficacy.

These domains were selected based on their relevance to the broader research goals and their established validity in prior studies on domain-specific self-efficacy [25, 26, 49]. Specifically, task-specific self-efficacy items have been shown to be valid and reliable measures in a high school and

41

introductory college level physics context [25, 26], while course-level and career-level self-efficacy items have been shown to be valid and reliable in the context of physics [49]. The survey items can be found in the appendix.

Focusing on task-specific, course-level, and career-level self-efficacy aligns with the broader aim of developing this mixed methods approach, which is to be able to capture transfer students' self-efficacy as they navigate their academic journey, beginning at a community college, through the transfer process, and ending at a bachelor's degree-granting institution. Given this larger intention, focusing on broader academic domains (task, course, and career) rather than domain-specific measures such as physics self-efficacy stems from the desire to reflect the wide-ranging academic contexts transfer students encounter. These students navigate various institutions, disciplines, and academic settings, each with its own challenges and expectations. By focusing on domains exploring their academic experiences, this approach can provide a more holistic understanding of self-efficacy that applies across diverse academic experiences, rather than limiting the scope to a particular subject area or field. This approach aligns with a larger institutional context within which students develop and adapt their self-efficacy, allowing a deeper exploration of the factors that influence their academic journey. This also supports this mixed methods approach in being able to be used within other contexts such as at the graduate education level, in which students must navigate moving from their bachelor-degree granting institution to their PhD-degree granting institution as well as the differences in expectations placed on PhD students. Furthermore, this approach can still explore specific disciplines such as physics self-efficacy, since a student's physics self-efficacy would be related to their self-efficacy to perform physics tasks, pass their physics courses, obtain their physics degree, and pursue a career in which they can leverage their physics experiences. As such, this mixed methods approach, with its emphasis on broad academic domains, allows for a flexible exploration of self-efficacy across different contexts.

To capture these varied experiences, the design of the survey was intentionally kept short to minimize the participant burden while still ensuring that the data collected would provide meaningful insights [87]. During data collection, students were prompted to answer either the

course-level self-efficacy questions or the career-level self-efficacy question depending on their current context and relevance. For example, if students were engaged in an activity related to a course, they would complete the self-efficacy items at the course level. If the activity was not related to the course, they would complete the career-level self-efficacy items.

A critical component of any ESM study is the sampling method used to collect data on the experiences of the participants. In this research, we used semi-random sampling. This approach allowed us to capture a diverse range of experiences, such as meetings with peers, engaging with faculty, or completing assignments, without overwhelming participants. Random sampling of experiences ensures that a variety of contexts are captured, but we also sought to minimize the participant burden. Therefore, we decided to restrict the signaling time to between 9 AM and 6 PM, with surveys delivered in four designated time windows. The application signaled participants to complete a survey during one of these windows, for a total of four surveys a day.

Overall, leveraging the ESM via surveys allows real-time data on students' self-efficacy to be measured across 3 different domains while minimizing participant burden and maximizing the relevance of the data to their academic experiences. With the ESM, we can gain a more accurate and dynamic understanding of how self-efficacy fluctuations in real-world academic contexts.

4.2.4 Overview of the Qualitative Source: Individualized Daily Journal Prompts

As part of the mixed methods approach, individualized daily journal prompts were designed to complement the real-time data collected through the ESM surveys. The purpose of these journal prompts was to capture participants' reflections on their daily academic experiences, allowing a deeper understanding of their self-efficacy as it is related to these specific moments. By asking participants to reflect on their day-to-day experiences, the journal prompts provided valuable qualitative data that could help contextualize and enrich the task-specific self-efficacy scores from the ESM surveys.

Asking for daily reflections makes sense in the context of this study, as the students completed four ESM surveys each day. This frequency of data collection allowed us to capture specific moments from each day, providing a detailed snapshot of students' experiences. The individualized

daily journal prompts gave students an opportunity to reflect on their self-efficacy fluctuations in real-time, either within a specific academic context or across different moments throughout the day. By linking the individualized daily journal prompts with the same daily experiences captured in the ESM surveys, we were able to gain a more nuanced understanding of how self-efficacy may fluctuate within a single day or across different academic tasks, improving the depth and relevance of the data. Reflections help to provide depth to the quantitative data by offering students the opportunity to confirm, disconfirm, or expand upon their survey responses, enriching our understanding of their self-efficacy experiences. The relevance of these individualized daily journal prompts is guided by the quantitative data itself. By using the results from the ESM surveys, we can tailor the daily journal prompts to address specific aspects of students' self-efficacy, thereby encouraging deeper reflection on areas where they may have shown fluctuations or challenges, and fostering a more focused exploration of their academic experiences. In some cases, students may also reflect on particular daily experiences that were not captured in the ESM surveys. These reflections can still serve as a point of triangulation around how a particular academic context, environment, or activity more broadly may be impacting student's self-efficacy; in other words, these reflections can serve as complementary data supporting our understanding of students' self-efficacy.

4.2.5 Individualized Daily Journal Prompts Design and Implementation

Students were signaled to complete their individualized daily journal prompts through the same application that was used to signal them to complete their ESM surveys. These journal prompts were delivered at the end of the day, after their final survey was completed. We delivered the journal prompts through Microsoft OneNote, in which each student had their own notebook to keep their data private and secure. These prompts allowed students to reflect on their experiences close to when they occurred, offering rich and more nuanced insights that complemented the quantitative data collection through the ESM surveys.

The individualized nature of these prompts was a critical aspect of the study design. Each prompt was tailored to the unique academic experiences of the participant, ensuring that their reflections were directly relevant to their current context. This customization allowed for more meaningful

insights into how students perceived and evaluated their academic capabilities in relation to their real-time experiences. For example, prompts might encourage participants to reflect on a particular class discussion, an upcoming assignment, or moments where they felt more or less confident in their abilities, either toward a specific task or within a broader academic context. By aligning the prompts with the students' immediate academic experiences, we were able to capture more precise and context-rich reflections, which in turn provided a clearer picture of their self-efficacy. An example of an individualized daily journal prompt is provided below.

> *Yesterday we asked how your experiences in the high energy physics lab affects your confidence in pursuing a career and you said, "They reassured me that I am on the right path...". Can you tell us a little bit more about how your experience in the high energy physics lab affects your confidence in pursuing a career?*
>
> - *Tell me a story from today that is an example of how working in the high energy physics lab affects your confidence in pursuing a career. What were you doing? What were others doing?*

- Jane's Example INDIVIDUALIZED Fall 2021 Daily Journal Prompt, Week 1, Day 4

The example above highlights some of the specific strategies used to individualize the daily journal prompts provided to students. One strategy involved incorporating the setting into the prompt design. This meant specifying whether the experience the participant was reflecting on occurred in the context of a course, a research lab, or another academic setting. In the example, the prompt encourages reflection on the student's experience in their high-energy physics lab and how that setting impacts their career-level self-efficacy.

Another strategy for individualization involved using the participant's own language, drawn from their previous journal entries. For instance, in the example above, the phrase "They reassured me that I was on the right path..." directly reflects the student's own words from their previous daily journal prompt response. This approach strikes a balance between incorporating the participant's language and maintaining the context of their experiences, while also allowing for a sense of agency

in the reflection process. This is demonstrated in the open-ended nature of the prompt: "Can you tell us a little bit more about how your experience in the high-energy physics lab affects your confidence in pursuing a career?"

A critical strategy used to individualize the daily journal prompts involved tailoring them based on the ESM data, a process that evolved across the two iterations of the mixed methods approach in Fall 2020 and Fall 2021. In both iterations, the prompts were individualized by incorporating analytical observations drawn from the ESM survey data, as highlighted in Figure 4.1. Specifically, students were asked to reflect on moments in which their self-efficacy appeared to fluctuate. However, as revealed through preliminary analysis discussed in this section, another strategy involved incorporating task and course-specific information from the open-ended ESM survey responses. This use of task and course data manifested itself in several ways. For example, it allowed us to focus on particular moments that were of interest to the researchers, such as meetings with a research or academic advisor. Furthermore, this information was used to provide context for the self-efficacy signals identified in the ESM surveys, allowing us to explore further the threads and patterns that emerged. These threads will be explored further in Chapter 5.

In some scenarios, students received generic prompts instead of individualized ones. This decision was based on the idea that offering a generic prompt was preferable to creating an individualized prompt when there was no clear data-driven rationale to do so. Specifically, if a student's ESM survey and/or previous daily journal responses did not yield new or conflicting insights, it would have been misleading or counterproductive to guide their reflection with an individualized prompt. In such cases, the use of a generic prompt was necessary to minimize researcher bias and ensure that students' responses were not unintentionally shaped by the researchers' expectations or experiences.

Although our goal was to provide each student with an individualized daily journal prompt, we apriori prepared a set of generic prompts: prompts that did not incorporate contextual or specific student experiences or language from their data. These generic prompts could be given to any student, regardless of their individual responses. An example of a generic prompt is given below.

Think of a time you felt really confident about your performance in a particular class -
either one you're taking now or one you've taken in the past. What about it made/makes
you feel confident?

- *Take me back to that moment; what were you feeling and experiencing?*

- Jane's Example GENERIC Fall 2021 Daily Journal Prompt, Week 1, Day 1

In examining the generic prompt above, although it specifies a temporal element (i.e. asking about the student's experiences in the past or present) and the environment (i.e., a course they are currently taking), it does not draw on any personalized information from the student's experiences. For instance, a more individualized prompt might specify a particular course, such as a physics course, which would be directly tied to the student's own academic context. However, this type of prompt would not be applicable to every student in the dataset, as not all students may be enrolled in a physics class. Additionally, while the generic prompt asks students to reflect on a course in which they felt particularly confident, it does not directly ask about their self-efficacy in relation to a particular course. If a student chooses to reflect on a particular course they are taking during the data collection period, this could still provide useful information about their self-efficacy, especially when analyzed alongside their ESM survey data.

Moreover, the student's response to this prompt might align with the ESM data, offering us a more comprehensive understanding of their self-efficacy in that course. For example, if a student indicated high self-efficacy on an ESM survey related to a particular class, their response to the generic prompt could confirm that they felt confident in that class, reinforcing the validity of the ESM data. Thus, while the generic prompt does not specifically leverage the ESM data, it can still provide valuable insights when paired with it.

The use of generic prompts proved invaluable, especially when time constraints made it difficult to analyze each student's data in depth and create tailored prompts. In these instances, pre-prepared generic prompts allowed data collection to continue without compromising the integrity

of the responses. More details about the individualized daily journal prompts will be discussed in Chapter 6.

4.3 The Story of Two Iterations

This subsection tells the story of the two iterations of the study, focusing on the key logistical decisions made during the preliminary analysis. It explores the design and implementation of the mixed methods approach, with particular attention to the sample, the data collection period, the technology used, and other logistical considerations. These details are integral to understanding the preliminary analysis.

4.3.1 Iteration One: Fall 2020 Data Collection

In Fall 2020, data were collected as part of the mixed method study conducted by the original research team. I was not directly involved in the Fall 2020 data collection process, but the following descriptions are based on the documentation and information provided by the original research team. The key features of the data collection process were as follows:

Sample: Data was collected from a sample of undergraduate physics students enrolled in upper division physics courses during the Fall 2020 academic semester. These students were contacted because they were physics majors who had taken at least one credit from a two-year college and were currently enrolled in upper division physics courses. The choice to include students with previous two-year college experience was intentional, as the goal of the study was to eventually implement this approach in two-year college settings. This limited sample was also a practical consideration, as this was the first implementation of the mixed methods approach. A smaller and more focused sample allows for the evaluation of the feasibility and effectiveness of the approach before scaling it up to include more students. In Fall 2020, a total of 6 students participated in the study. Additionally, because of the researchers' connection to the physics field and more specifically the department, my colleagues had access to the physics student body and recognized the opportunity to explore self-efficacy in a discipline where students often report lower self-efficacy compared to other STEM fields [25, 51].

Duration of Data Collection: In conducting the Fall 2020 data collection, one key consideration

48

was determining how long to ask students to participate and when to have completed the surveys and individualized prompts. Given the comprehensive nature of data collection, the length of participation posed an important challenge; especially, given that the data collection process requires a substantial commitment from the students.

The original research team recognized the need to strike a balance between several factors to determine the duration of participation. First, the burden on students was an important consideration. Asking students to complete four surveys and a daily journal entry required time commitment, and it was essential not to overburden them with too much data collection over an extended period. Given the logistical challenges and the potential for participant fatigue, the aim was to ensure that the duration of the study remained feasible and sustainable for the participants.

At the same time, the study design required sufficient data over a number of days to capture significant fluctuations in student self-efficacy. The surveys were sent semirandomly throughout the day, and to detect any shifts or changes in self-efficacy, we needed to ensure that students participated for a long enough period to produce variability in the data. This meant enough data points needed to be collected to observe trends or patterns over time, while also accounting for individual variations in self-reports.

Another consideration was the amount of data needed to capture a diverse range of academic experiences among participants. A shorter duration, such as one week, risked missing key academic tasks or variations in students' experiences, which is one of the primary goals of this mixed methods approach. Along with this variation, the collecting the data semi-randomly could result in encounter instances where students were not reflecting on particular academic experiences. To illustrate, students might have been eating breakfast or lunch or commuting on the bus, rather than participating in academic-related activities when completing the surveys.

As such, the original research teams chose to collect two weeks; however, they also considered when to collect those data during the semester. In alignment with midterms, they implemented a four-week break between the first and second weeks of data collection, in which the first week of data collection occurred in the first week of October and the second week of data collection occurred

49

in the second week of November. This break was intentional, allowing for a pause in data collection to reduce participant fatigue and ensure that responses were not influenced by the continuous nature of the surveys and individualized daily journal prompts. The purpose of implementing this gap was to maintain the quality of the responses by giving participants time to reset, preventing patterns of complacency or exhaustion from affecting the data. This approach aimed to balance the need for consistent data collection with the well-being of participants, ensuring that their participation remained high and that their reflections were as accurate and meaningful as possible.

Administration of ESM Surveys and Individualized Daily Journal Prompts: During the data collection, the ESM surveys and individualized daily journal prompts were administered only on weekdays, excluding weekends. This decision was based on the assumption that the primary academic experiences of the students occur during the workweek, making it the most relevant period to study self-efficacy in academic contexts. Weekends often involve different routines, such as social activities, work obligations, or personal time, which might not provide the same insights into students' academic experiences. Furthermore, the original research team recognized that having students complete surveys throughout the week and the weekend would burden the students more considering that students' weekends may involve different routines and activities. Although this approach helped focus data collection on academic contexts, it also introduced potential limitations. For example, some students may be involved in coursework or study sessions on weekends, and these experiences were not captured.

The original research team chose to administer the ESM survey four times a day, during four different time windows. During the time windows, students could be signaled at any time to complete the survey. Once the student was signaled they had 45 minutes to complete the survey, and if they did not complete it within this 45 minutes, the student missed responding to the survey. Recognizing that their audience of students were transfer students at a bachelor's degree-granting institution, the team set the survey time windows between 9 am and 6 pm, and students completed their journal reflections after 6 PM. Table 4.1 illustrates the specific time windows and the distribution of the survey for each week.

To distribute the ESM surveys, the research team used a smartphone application, which required students to own smartphones. This approach allowed for the semi-random, real-time distribution of surveys, ensuring students could be signaled throughout the day without needing to be physically present in a specific location. The use of a smartphone app also minimized logistical challenges associated with other survey distribution methods, such as email or paper responses. Other advantages of using a smartphone application included protecting their data and ensuring that students could complete surveys quickly and efficiently, regardless of their location, as long as they had access to their smartphones. Although the use of a smartphone app provided several benefits, including real-time distribution and location flexibility, it also presented some limitations, such as the need for all participants to own compatible devices or for the research team to provide devices, potential challenges in ensuring consistent engagement across diverse student schedules, and the application itself costs.

The application was also used to send reminders to students to complete their individualized daily journal prompts. For these prompts, OneNote notebooks were utilized to deliver the individualized daily journal entries to the students. This required students to work with an additional application in addition to the ESM survey smartphone app. However, since the students' home institution uses Microsoft applications, including OneNote, the intention was that they would likely be familiar with this platform. Each student had their own private notebook to ensure that no one outside the research team had access to their responses.

In summary, the Fall 2020 data collection process used a mixed method approach that combined

Week Number	Daily Time Window Number	Time
Week 1 and Week 2	Daily Time Window 1	9:00 AM - 11:15 AM
	Daily Time Window 2	11:15 AM - 1:30 PM
	Daily Time Window 3	1:30 PM - 3:45 PM
	Daily Time Window 4	3:45 PM - 6:00 PM

Table 4.1 ESM Survey Distribution for Fall 2020: Each student was randomly prompted to complete the survey four times throughout a single day, resulting in four distinct time windows. Within Daily Time Window 1, one student might be signaled at 9:13, while another might be signaled at 10:27, demonstrating the semi-random nature of the survey distribution.

ESM surveys with individualized daily journal prompts, as shown in Figure 4.2. The use of smartphone applications for survey distribution and OneNote for journal reflections was used with the intention of ensuring flexibility and ease of data collection for participants. Despite limitations such as the need for compatible devices and the potential for varied engagement among students, the chosen methodology allowed for a rich set of data that will be used for the preliminary analysis in this chapter and for the analysis in subsequent chapters.
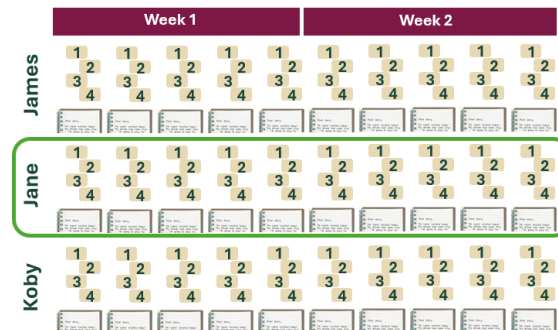
Figure 4.2 This figure illustrates the explanatory sequential mixed methods (ESM) approach used. Specifically, this shows three students' data across the two-week period (i.e., Jane, James, and Koby). Each tan box represents an ESM survey that students were required to complete, while the notebooks with writing symbolize the individualized daily journal prompts. For example, Jane's first ESM survey on Day 1 is shown by the tan box labeled "1'. Each student participating in the Fall 2020 data collection completed four ESM surveys per day across a 5-day "week" (i.e., Monday through Friday), with a total of 40 surveys completed per student over two weeks (20 surveys per week). While this figure depicts the data collection for three students from the Fall 2020 cohort, the full dataset includes responses from six students. Jane's data is highlighted in green, and a preliminary analysis of her responses is presented later in this section.

4.4 My Reflection through a Preliminary Analysis

As discussed earlier, I was not originally a part of the research team that designed the overall mixed methods approach nor the team that implemented it to collect the Fall 2020 data. My understanding when I came in was that the research team was interested in validating or evaluating whether the ESM surveys and the individualized daily journal prompts could be connected to explore the self-efficacy of students in relation to events as students were experiencing these events. My goal was to come in to explore how the ESM surveys and individualized daily journal prompts were leveraged to do this work, and to determine the ways in which we could build claims and what

claims we could build from using such a mixed methods approach.

In working toward this goal, I started this work by conducting a preliminary analysis to explore how the quantitative and qualitative elements of the ESM surveys and the individualized daily journal prompts could be connected to each other with the Fall 2020 data. A preliminary analysis felt right for this stage because I was becoming familiar with the ESM surveys and the individualized daily journal prompts, and because having data from students supports understanding how students are engaging with these data collection methods. Furthermore, the preliminary analysis was foundational for driving the work that I present in Chapters 5-7, which speaks to the validity and rigor behind the research design.

Thus, in the following subsection, I present the preliminary analysis. The preliminary analysis was conducted through the lens of Jane's data. As such, her experiences and engagement with the Fall 2020 iteration of the mixed methods approach will serve as a lens through which I explored the connections that can be built to link the ESM surveys and the individualized daily journal prompts and responses. The intention of connecting her quantitative and qualitative elements was to see how these components worked together to explore Jane's self-efficacy in relation to events of her day. From the preliminary analysis, we made three key modifications to the Fall 2021 iteration of the mixed methods approach to address the limitations identified by analyzing Jane's Fall 2020 student. Thus, for the remainder of this section, I will discuss the preliminary analysis through the lens of Jane's Fall 2020 data collection, the limitations identified from the preliminary analysis, and how those limitations translated into modifications made to the Fall 2021 iteration of the mixed methods approach. Additionally, I will discuss how the modifications made to the Fall 2021 iteration then motivated the studies produced in Chapters 5-7. This aforementioned process is illustrated in Figure 4.3. Specifically, Chapters 5-7 will expand on one of the key limitations/modifications. This modification laid the foundation for gathering information on how to integrate across the quantitative and qualitative components to explore the self-efficacy of students in relation to events as these events occur. I will conclude the dissertation with a discussion in which I share my insights into the iterative process of refining and enhancing this mixed methods approach.
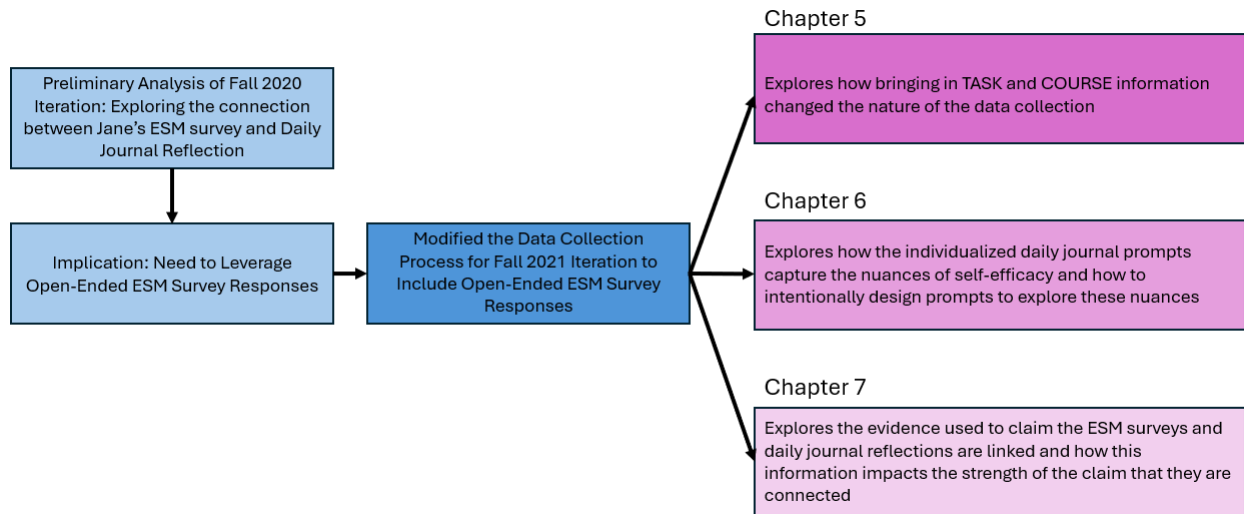
Figure 4.3 In the figure above, the light blue box marks the starting point of this process: the preliminary analysis. From there, the arrow leads to the next light blue box, where we built on the insights gained to rethink how to explore self-efficacy in relation to participants' real-time experiences. Based on these findings, we modified our data collection strategy by incorporating open-ended ESM survey responses, as shown in the next, darker blue box. These insights also informed the creation of individualized daily journal prompts. Chapters 5-7 branch from this work, as indicated by the arrows splitting into three directions toward purple boxes. Each chapter is represented by a different shade of purple to highlight the different aspects of the research methodology I explored.

4.5 Preliminary Analysis through the Lens of Jane

The purpose of this section is to present the preliminary analysis that I conducted in the early stages of my dissertation when I first engaged with the project. This analysis centers on the case study of a student, referred to by the pseudonym, 'Jane,' and examines how her qualitative and quantitative data interact. In exploring Jane's case, I began investigating how the ESM surveys and individualized daily journal reflections could be connected. This initial integration of these data sources led to a deeper exploration of their relationship, which is detailed in Chapters 5-7.

In particular, this section demonstrates how combining the ESM survey and individualized daily journal prompt data can create a more comprehensive understanding of a student's self-efficacy in real time. Through Jane's data, I will illustrate how these different data types complement each other and produce deeper insights into her self-efficacy at a specific moment.

This analysis was crucial not only for refining our research design, but also instrumental in shaping the mixed methods approach used in subsequent data collection during the Fall 2021

semester. This section aims to establish a clear connection between the initial findings of this analysis and the direction of the research that unfolded in the following chapters.

4.5.1 Contextualizing Jane's Data

In this section, I introduce Jane, a student whose data serves as the basis for this case study. Jane was her chosen pseudonym used to protect her privacy, and her data provided insights into how this mixed methods approach captures a student's self-efficacy fluctuations in real time. This analysis focuses on two key sources of data: ESM surveys and individualized daily journal prompts, both of which were collected as part of the larger project aimed at exploring student self-efficacy.

Jane's data was collected as part of the Fall 2020 data collection to support a preliminary study to implement this mixed-method approach. Since her data had been collected, my primary task was to analyze Jane's data. Then, the main goal of this analysis was to see how the quantitative and qualitative data informed each other.

The ESM surveys were used to collect Jane's quantitative data. These surveys were designed to capture Jane's domain-specific self-efficacy, including her task-specific self-efficacy at specific moments throughout her day, using notifications sent at random intervals. Each notification asked Jane to rate her confidence in completing a task at that time, offering a snapshot of her self-efficacy related to a specific activity she was engaged in at the time. For the purposes of this preliminary analysis, I focused on analyzing the task-specific self-efficacy scores, as they were directly related to Jane's immediate experiences and reflect her confidence in completing the task at hand at that time.

Each survey response, or "Notification," corresponds to a specific point in time in Jane's day. Figure 4.4 illustrates the data collection process through the lens of Jane, providing a visual representation of how these survey responses are captured throughout her day. For example, Notification 8 refers to the fourth survey completed on the second day of data collection, providing a measure of her task-specific self-efficacy as she engaged in a particular activity during that time. The self-efficacy scores from these surveys serve as the quantitative data in this case study.

In addition to the quantitative data from the ESM survey, Jane also completed individualized
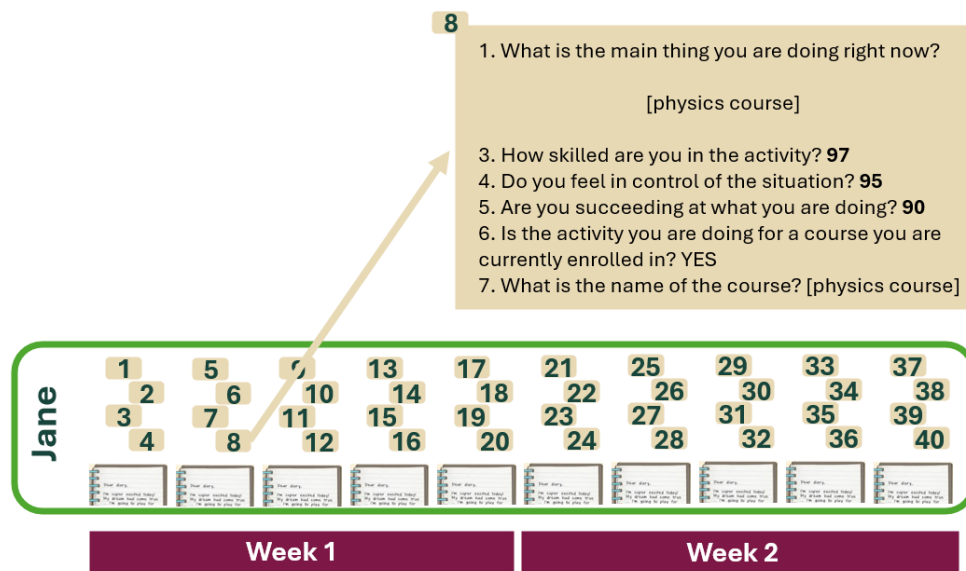
Figure 4.4 This figure illustrates the explanatory sequential mixed methods (ESM) approach used in this dissertation. Each tan box represents an ESM survey that students were required to complete, while the notebooks with writing symbolize the daily journal prompts. For example, Jane's first ESM survey on Day 1 is shown by the tan box labeled '1'. Each student participating in the Fall 2020 data collection completed four ESM surveys per day across a 5-day "week" (i.e., Monday through Friday), with a total of 40 surveys completed per student over two weeks (20 surveys per week). While this figure depicts the data collection for three students from the Fall 2020 cohort, the full dataset includes responses from six students. Jane's data is highlighted in green, and a preliminary analysis of her responses is presented in this section later.

daily journal prompts as part of the Fall 2020 data collection, providing qualitative insights into her experiences. These prompts were designed to capture Jane's reflections about her experiences from specific moments or overall experiences in relation to her self-efficacy. The journal responses are an essential component of the analysis, as they provide additional context to the ESM survey responses.

For example, while the ESM surveys offer a numerical representation of Jane's self-efficacy at specific moments, the journal prompts give her the opportunity to reflect on her experiences in greater depth, relating her experiences to how they interacted with her self-efficacy. These qualitative data offer valuable context to help interpret the fluctuations in Jane's self-efficacy scores. Figure 4.5 illustrates one of the daily journal prompts and Jane's corresponding responses. The analysis that follows will explore how this journal prompt and response enrich our understanding of her self-efficacy within this particular moment and context.
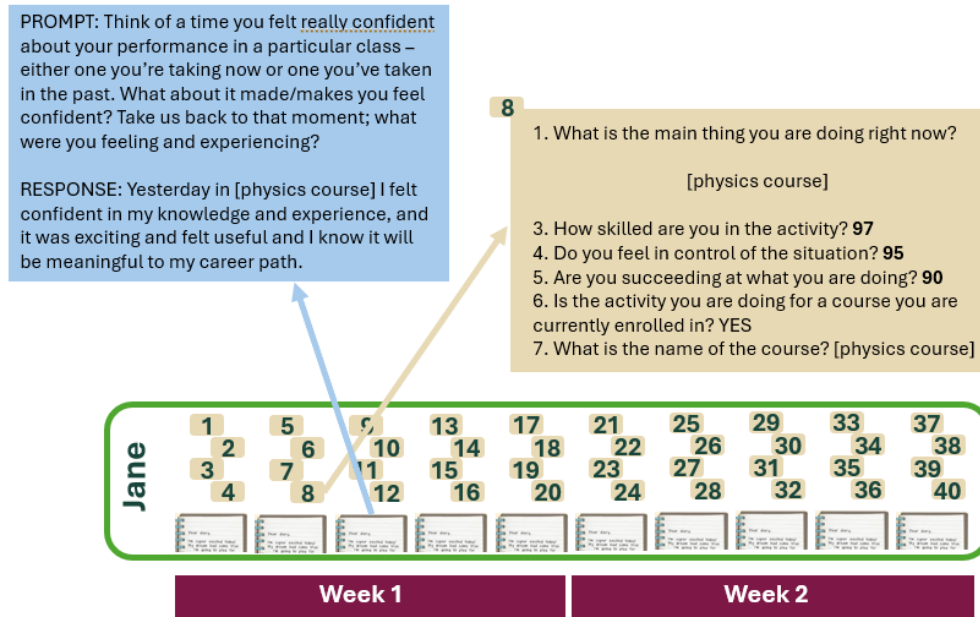
Figure 4.5 This figure illustrates one of Jane's responses to a daily journal prompt she completed over the course of the two-week study period. The journal reflection (i.e., prompt and Jane's response) is provided in the light blue box. This prompt was given to Jane on Day 3 of Week 1 of the data collection demonstrated, as indicated by the start of the arrow coming from the third notebook. The prompt given to her is identified by "PROMPT:", and Jane's corresponding response is provided after "RESPONSE:". Jane would have written her response on Day 3, after receiving the prompt.

By providing an example of an ESM survey responses with the corresponding daily journal prompt and response, this section will illustrate how both quantitative and qualitative data are essential for capturing a comprehensive representation of Jane's self-efficacy. On their own, the ESM data show fluctuations in task-specific self-efficacy, but this is limited without contextualizing the fluctuation. The individualized daily journal prompts provide this rich or additional context, helping us to represent Jane's self-efficacy at particular moments.

This case study serves as an initial exploration of how these two types of data can work together to offer a more holistic view of self-efficacy, particularly in the real-time context of Jane's academic experiences. The integration of these data types will be explored in greater detail in later sections, but it is important to first understand the separate components of Jane's data and their relevance to this study.

4.5.2 Analyzing Jane's Task Specific Self-Efficacy Scores: Identifying Peaks and Dips in Jane's Scores

In examining Jane's data from the Fall 2020 iteration of the mixed methods approach, I began with analyzing Jane's fluctuations in her task-specific self-efficacy scores. I started by making sense of the quantitative ESM survey scores, as this is in alignment with the research approach being an explanatory sequential mixed methods design. Jane's task-specific self-efficacy scores are directly linked to the activity she was doing at the given moment she completed the survey. These scores reflect Jane's self-reported confidence in her ability to perform a specific task successfully at the time she completed the survey, providing valuable insights into her task-specific self-efficacy related to a specific activity in a particular moment.

A plot was built to visually display Jane's raw task-specific self-efficacy scores for each notification in order to identify patterns, trends, or changes in her scores. Figure 4.6 provides Jane's task-specific self-efficacy per notification for week 1 of the data collection. To situate the plot within the context of the ESM survey data, a notification refers to a single survey response from a student. For example, Notification 8 on day 2 would be the fourth survey completed by Jane on the second day of the week 1 data collection. In each of these surveys, Jane answered a set of items, covering her domain-specific self-efficacy.

The advantage of plotting Jane's raw task-specific self-efficacy scores is that it allows us to examine fluctuations in her self-efficacy. These fluctuations are significant because they indicate moments when Jane's task-specific self-efficacy was either threatened (leading to a decrease) or supported (leading to an increase). These moments are valuable for researchers and educators, as they could provide insight into the factors that influence students' task-specific self-efficacy. This understanding can guide efforts to support students in developing their self-efficacy. For example, as shown in Figure 4.6, Notification 8 appears to reflect a high moment of self-efficacy. Later in the section, we will discuss using the individualized daily journal prompts to support further interpretation of Jane's task-specific self-efficacy score at Notification 8.

Identifying this as a peak is based on my interpretation of the plot, where a noticeable increase
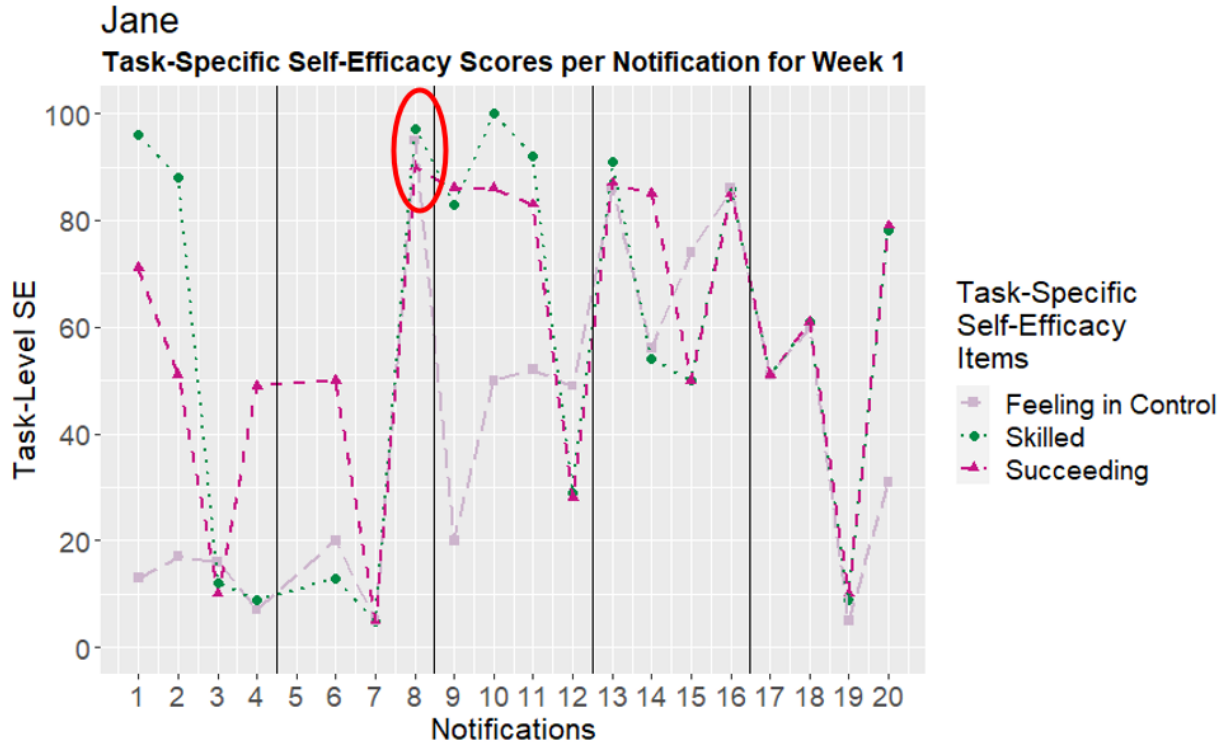
Figure 4.6 This figure illustrates Jane's task-specific self-efficacy scores per notification for Week 1. The y-axis represents Jane's task-specific self-efficacy ratings, ranging from 0 to 100, which corresponds to her confidence in completing tasks on a 0-100 scale. The x-axis shows the notifications, with each notification representing a specific survey completed on different days of the week. For instance, Notification 8 corresponds to the fourth survey completed on Day 2. Vertical lines are drawn to separate the surveys by day, with the first four notifications representing Day 1, and a vertical lined placed after Notification 4 to separate Day 2. The legend on the right-hand side indicates the different task-specific self-efficacy items, with each of the lines representing the fluctuations in Jane's ratings for a particular item. For example, the purple dotted line shows the changes in scores for the success item on the survey (i.e., Succeeding). Notification 8 is circled with a red, solid oval as an indicator of a peak that the research team identified and because throughout this sub-section, I will be examining Notification 8.

in Jane's self-efficacy stands out. In other words, no statistical tool was used to identify peaks or dips. Statistical tools may be limited in how they identify peaks or dips in self-efficacy based on the amount of data we have during data collection. For example, on Day 2 of Week 1 of data collection, we would not have enough data to potentially make a strong statistical inference about an increase or dip, during the data collection period we have relied on the researcher's interpretations of the plots to identify peaks or dips.

59

4.5.3   Analyzing Jane's Task-Specific Self-Efficacy Scores: Contextualizing Jane's Scores with Courses

Acknowledging that the activity in which Jane was engaged at the time might be related to her career or course self-efficacy became an important element of both her narrative and future analysis of her task-specific self-efficacy scores. This insight was encouraged by both the literature and Jane's qualitative data, allowing for an exploration of the relationships between different domains of self-efficacy. For example, research suggests that students often feel more efficacious in certain courses than in others [25, 34, 39]. Jane's qualitative responses also indicated that she experiences different courses in varying ways, indicating that her domain-specific self-efficacy may vary when situated within her course information. Below are two excerpts from her daily journal responses, illustrating her distinct experiences with different courses.

> *"Frustrated with the math department, proud and excited about the physics department, frustrated with the [computational modeling course]..."*
> - Jane's Daily Journal Prompt Response, Week 1, Day 1

> *"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's Daily Journal Prompt Responses, Week 1, Day 1

From these responses to daily journal prompts, it is clear that Jane experiences her courses in her Fall 2020 semester differently. Given this, I was interested in determining whether these differences might be reflected in her task-specific self-efficacy scores. I posited that her task-specific self-efficacy ratings could differ depending on the course they were associated with. To be able to explore such patterns, I wanted to incorporate course information into her task-specific self-efficacy plots.

To do so, I color-coded the notifications on the plot according to the specific course Jane reported that was related to the activity she reported in her survey response (i.e., notification). This feature enabled me to contextualize Jane's task-specific self-efficacy scores, providing a way to examine how her task-specific self-efficacy varied across different courses or even situate a peak or dip in a student's task-specific self-efficacy within a particular course. For example, Figure 4.7 illustrates how color-coded notifications help us recognize that Notification 8, which reflects a peak in Jane's task-specific self-efficacy, is specifically associated with her physics course. This visual context provides a clearer picture of how her self-efficacy aligns with her course experiences and provides additional contextual information about that notification.
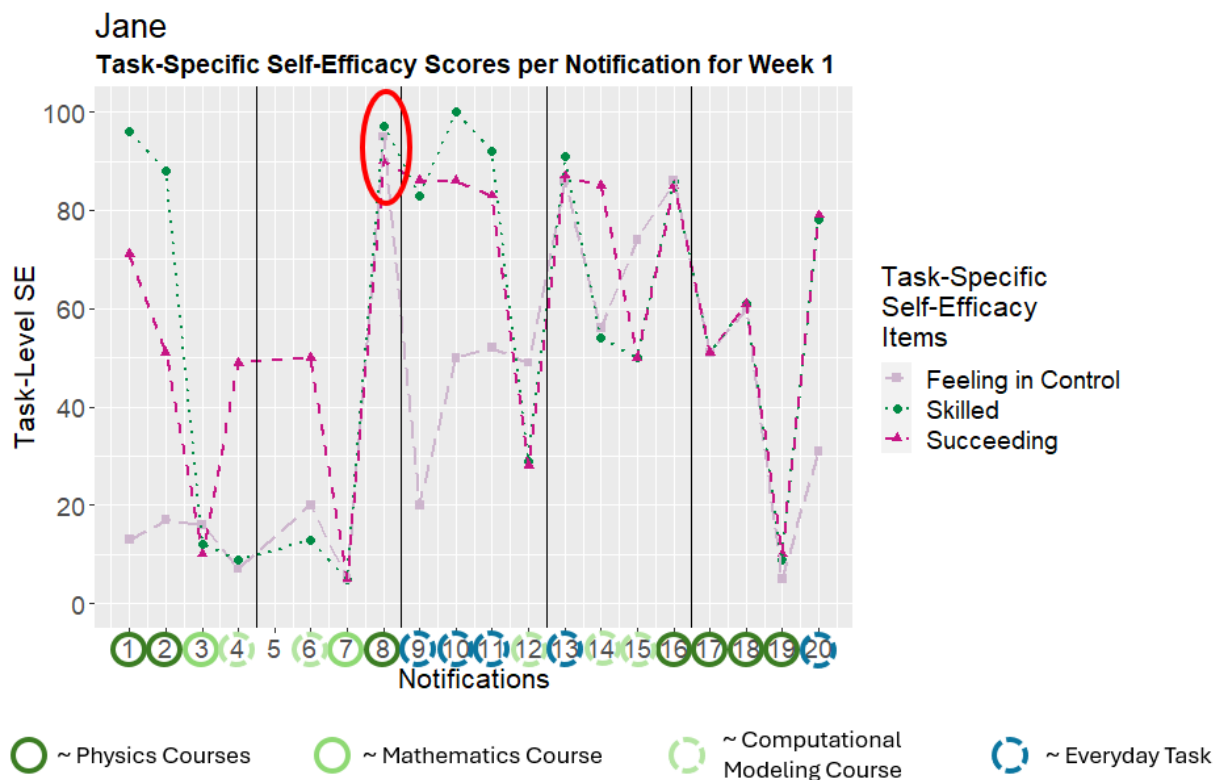
Figure 4.7 This figure illustrates the the circle analysis conducted on the task-specific self-efficacy scores plot. Jane was enrolled in three STEM courses: physics, math, and computational modeling. Each course is represented by a different colored circled to indicated which course a given notification is associated with. For example, Notification 8 is highlighted with a green circle, signifying that it is associated with Jane's physics course.

4.5.4 Interpreting Jane's Task-Specific Self-Efficacy Scores: Confirming My Interpretations of Jane's Scores

Upon analyzing the task-specific self-efficacy graph, I identified Notification 8 as a high moment of task-specific self-efficacy. In isolation, this observation could suggest a moment of increased confidence towards the activity, but to confirm the significance of this peak, I turned to Jane's journal responses.

To determine which journal response is related to the ESM survey, the timeline of when the data was collected was used. Jane's task-specific self-efficacy scores for Notification 8 corresponded to her fourth survey response on Day 2 of Week 1, so I reviewed Jane's journal responses from the second and third days of the data collection to see if she had reflected on this moment. Indeed, Jane's reflection from Day 3 read:

> *"Yesterday in [physics course] I felt confident in my knowledge and experience and it*
>
> *was exciting and felt useful and I know it will [be] meaningful to my career path"*
>
> - Jane's Daily Journal Prompt Responses, Week 1, Day 3

Given that Jane refers to the day prior (that is, Day 2) in the response and discusses her physics course, I can confidently align her journal response with the high task-specific self-efficacy scores from Notification 8 thanks to color-coding the notifications with the course information. Her reflection about feeling confident and engaged in her physics course matches the peak we observed in her task-specific self-efficacy plot. This alignment between Jane's ESM data and journal responses reinforces the conclusion that Jane felt high self-efficacy at the moment captured in Notification 8.

Given the prompt response, we have shown that Jane has confidence in her physics course at that time. This work demonstrated how combining Jane's quantitative task-specific self-efficacy scores with her qualitative journal responses provided the validity that Notification 8 was a high moment of task-specific self-efficacy for Jane. Overall, we also have evidence that Jane's quantitative and qualitative data can be integrated with one another to get a snapshot into Jane's self-efficacy towards that task for her physics course at the moment she is doing it. Figure 4.8 highlights the process

through which the quantitative and qualitative data informed each other and as such weaved together Jane's qualitative and quantitative data to come to this conclusion regarding Jane's self-efficacy towards this task within a course at this particular moment. This also serves as a proof of concept that the mixed methods approach can be used to build a narrative around a student's self-efficacy with a given domain in a particular moment through integrating the quantitative and qualitative components. For additional details about Jane's case see Myers's work (REF).
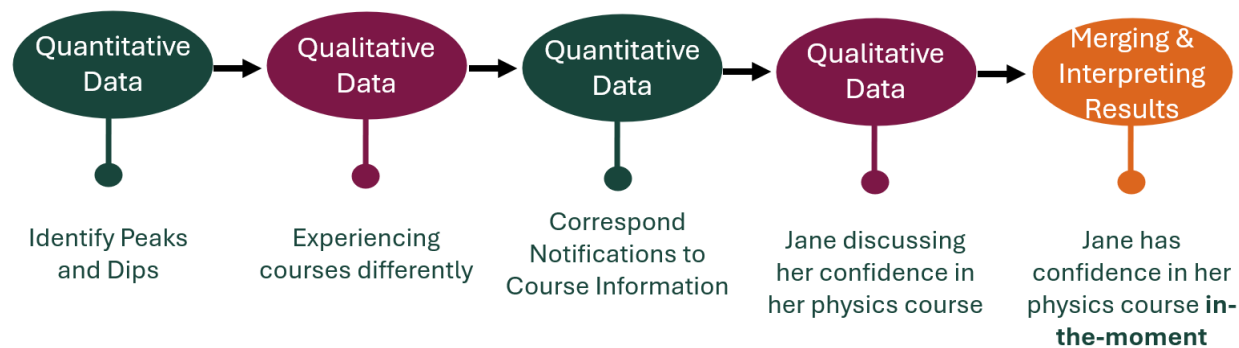
Figure 4.8 This figure illustrates the iterative process through which quantitative and qualitative data were used to inform and validate each other. The analysis begins with the task-specific self-efficacy plots, which highlight peaks and dips in Jane's scores. Next, the qualitative data - specifically, Jane's daily journal prompt responses - were used to contextualize and further inform the quantitative analysis. The course information then helped identify which notifications were associated with specific courses, when applicable. This cross-referencing process led to a merged conclusion: Jane demonstrated confidence in her physics course at that particular moment.

4.6   Modifying Fall 2021 Data Collection Based on Preliminary Analysis

After conducting the preliminary analysis, I used the findings and insights from this work to consider how to modify the Fall 2021 data collection to address some of the limitations of the current iteration of the mixed methods approach and to enhance its validity, trustworthiness and credibility. I identified three key areas of focus for these modifications: (1) the need to expand participation beyond physics transfer students. This broadening was necessary to capture a wider range of student experiences across different STEM disciplines, allowing for a more diverse and representative sample. (2) The need to increase the duration of the data collection period. A longer data collection time frame supports within-participant analysis, allowing for a deeper exploration of how participants' self-efficacy fluctuates over time, particularly in relation

to their real-time academic experiences. (3) The need to add open-ended responses to the ESM surveys. This addition was necessary to enable a richer interpretation of task-specific self-efficacy scores by allowing participants to explain and contextualize their experiences, providing a deeper understanding of the factors influencing their self-efficacy at specific moments.

4.6.1 The Need to Broaden Participation Beyond Physics Transfer Students: Collecting Diverse Experiences

One of the preliminary findings from Jane's data revealed that she reported feeling differently about each of the courses she was taking, suggesting that her self-efficacy in these courses varied. This difference in emotional and physiological responses across courses highlights that self-efficacy is not a uniform experience, but one that may fluctuate depending on contextual factors, such as the nature of the course or the specific challenges it presents. This suggests that students' experiences are highly individualized and context-dependent, making it crucial to consider how self-efficacy manifests differently across various academic contexts.

Given this finding, broadening our participant pool for the Fall 2021 iteration of our mixed methods approach would allow us to more effectively explore the diversity of student experiences. By including a broader range of participants with varied academic backgrounds and course loads, we could capture the nuanced ways in which self-efficacy is shaped by different experiences. This expanded sample would provide a more robust understanding of how self-efficacy operates across different contexts, enhancing the generalizability of the findings that could be produced from such an approach, and would offer deeper insights into the factors that interact with students' confidence in their perceived abilities in academic settings.

Furthermore, broadening participation would allow us to refine our data collection and analysis methods, ensuring that the mixed methods approach can adequately account for variability in student experiences. This would make our study more comprehensive and better equipped to uncover the complex relationship between self-efficacy and the real-time experiences of the participants.

4.6.2 The Need to Broaden Duration of Data Collection Period: Collecting Data for Within Individual Analysis

Upon analyzing Jane's data, we found that while the mixed methods approach allowed us to make claims about specific moments, there are limitations in capturing such moments. In Fall 2020, only two weeks of data were collected, which covered about 40 moments if participants completed all surveys. However, this does not mean that participants are experiencing a lot of fluctuations in their self-efficacy across a two-week time period, nor does it mean that we can capture a lot of fluctuations in their self-efficacy around a particular course or context, when we are semi-randomly collecting data. In addition, how many moments would we need, for example, to say that Jane commonly reports high task-specific self-efficacy within her physics course, or Jane commonly reports high task-specific self-efficacy within her physics course when working with her group on in-class problems. This raises the question of how much data do we need from a participant to be able to claim that there is a trend in the kinds of activity or context that are threatening or supporting Jane's task-specific self-efficacy.

Thus, rather than collecting two weeks of data like in Fall 2020, in Fall 2021 we decided to collect four weeks of data to see if we could capture enough moments to be able to make claims around a participant's self-efficacy. Intentionally, the semi-random nature of this data collection was not changed because the goal was to be able to explore their self-efficacy across a diverse range of environments rather than exploring their self-efficacy within a particular context (e.g., a course).

4.6.3 The Need to Add Open-Ended Survey Responses for the Integration: Enhancing the Interpretation of Task-Specific Self-Efficacy

An important aspect of integrating the quantitative and qualitative data here was the use of course information to interpret Jane's task-specific self-efficacy in context. Specifically, to link her daily journal prompt response with the ESM survey response for Notification 8, it was crucial to recognize that the task Jane referred to in the ESM survey was associated with her physics course. Without accounting for this connection through the "circle analysis", it would have been difficult to link the Notification 8 survey data to the journal prompt, as the journal response focused solely on her physics course. In particular, the journal prompt did not mention her task-specific self-efficacy

scores directly or describe a particular task, so the course context played a key role in linking the quantitative (i.e., ESM survey) and qualitative (i.e., journal prompt data).

To perform this circle analysis, I relied on open-ended ESM survey responses, where students provided course-specific information. This suggests that including open-ended survey items in the ESM survey can be invaluable for contextualizing task-specific self-efficacy scores. Therefore, in the Fall 2021 data collection, we made a critical adjustment: instead of solely informing the writing of the journal prompts based on fluctuation in their self-efficacy scores, we incorporated their responses to the open-ended survey items to support writing the students' individualized daily journal prompts. This additional information was included with the aim of better understanding and interpreting the students' task-specific self-efficacy scores, thereby improving the process of writing the daily journal prompts.

4.7   Iteration Two: Fall 2021 Data Collection

In this section, I will discuss the logistics of the Fall 2021 data collection. The logistics of the Fall 2021 data collection closely mirrored those of the Fall 2020 data collection, with similar procedures, tools, and strategies used to ensure consistency in both iterations. Although the key differences between the two years have already been outlined, this section will provide a detailed look at how the data collection process was organized and executed for Fall 2021.

One key difference for the Fall 2021 iteration that has not yet been discussed was the research team. Although the faculty involved in the study remained the same as in Fall 2020, I replaced the two graduate students who were part of the research team in Fall 2020. This change in team composition allowed for a shift in the dynamics of the research process and may have influenced data collection and overall study execution.

Sample: The sample for Fall 2021 data collection was drawn from the same research-intensive bachelor's degree-granting institution as for Fall 2020 data collection. We solicited students who identified as STEM majors, including more than just physics majors, with priority given to this criterion. To identify these students, I first examined their major and then determined whether it placed them within the College of Natural Sciences. Once STEM majors were identified, students

were then considered based on their transfer status and/or first-generation college student status. This approach was motivated by my own experience as a first-generation college student. Although my initial focus was to explore self-efficacy among first-generation students, I ultimately did not engage with this criterion as extensively as originally planned. Within these criteria, I also aimed to include students with a diversity of STEM courses to capture a broader range of experiences within the field. Lastly, to further enhance the diversity of the sample, we also considered the organizations on campus to which the students belonged. In addition to these criteria, we extended participation offers to physics majors who had participated in the Fall 2020 study, to collect longitudinal data.

Duration of Data Collection: As mentioned previously, the Fall 2021 data lasted four weeks. The study was conducted consecutively for the first two weeks, followed by a week break, and then the last two weeks of data collection were completed consecutively. The decision to split the study into two parts with a break in between was intended to prevent participant fatigue and maintain data quality.

The two-week periods were strategically chosen to allow students time to settle into their classes, ensuring that they had some experience in their courses and were more situated in their academic routines. Additionally, the timing of data collection was designed to capture students' experiences on either side of midterms or around when they were taking their midterms, providing a clearer picture of how their self-efficacy might fluctuate during a critical point in the semester.

Administration of ESM Surveys and Individualized Daily Journal Prompts: The administration of the ESM surveys and individualized daily journal prompts in Fall 2021 closely mirrored that of the Fall 2020 data collection; however, there were two main differences between the two iterations. One of those differences was the switch from the LifeData application, used for delivering the ESM surveys in Fall 2020, to the Expiwell application in Fall 2021. The second difference between the data collection for Fall 2021 and Fall 2020 was in the adjustment of the time windows to respond to the ESM surveys. In the first two weeks of the Fall 2021 data collection, we reduced the response time from 45 minutes to 30 minutes. This change was made to increase the likelihood that the researchers could access the student's last response to the ESM survey on that day, allowing us to

probe this response if we wanted to with the individualized daily journal prompts. Otherwise, we would need to examine their last ESM survey response from the day on the next day if we weren't able to examine this before needing to send out their individualized daily journal prompts. However, students were provided with the option to report any issues with the study through an additional space in their individual OneNote notebooks. Some students noted that they were missing ESM surveys or struggled to complete them because they needed more time to respond. In response to this feedback, we extended the response time to 45 minutes for the last two weeks of data collection. We decided that it was more important to meet the needs of the students than to be able to examine the fourth survey response on each day. The students appeared to appreciate the additional 15 minutes. Table 4.2 shows the time windows for the Fall 2021 data collection.

Week Number	Daily Time Window	Time
Week 1 and Week 2	Daily Time Window 1	9:00 AM - 10:45 AM
	Daily Time Window 2	11:45 AM - 12:30 PM
	Daily Time Window 3	12:30 PM - 2:45 PM
	Daily Time Window 4	2:45 PM - 4:30 PM
Week 3 and Week 4	Daily Time Window 1	9:00 AM - 11:15 AM
	Daily Time Window 2	11:15 AM - 1:30 PM
	Daily Time Window 3	1:30 PM - 3:45 PM
	Daily Time Window 4	3:45 PM - 6:00 PM

Table 4.2 ESM Survey Distribution for Fall 2021: The daily time windows were distributed differently for Week 1 and Week 2 compared to Week 3 and Week 4. Each student was randomly prompted to complete the survey four times throughout a single day.

4.8 Setting Up the Next Three Chapters

4.8.1 Exploring the Impact of Open-Ended Survey Responses on Interpreting Task-Specific Self-Efficacy and Journal Prompt Development

Throughout my reflection of the preliminary data analysis, I have established the importance of incorporating open-ended survey items to better interpret task-specific self-efficacy scores during the data collection process, which in turn informs the creation of daily journal prompts. However, this raises further considerations, as the inclusion of qualitative data may shift the type of information we use to guide the journal prompt development.

In Fall 2020, our approach involved analyzing plots of raw task-specific self-efficacy scores.

When we observed fluctuations in students' scores, these fluctuations guided the content of the journal prompts. Yet, as indicated by the preliminary analysis, the open-ended survey responses - providing insights into students' tasks and courses - could alter the types of observations drawn from the ESM survey data. Consequently, this shift may change the nature of daily journal prompts.

This sets the stage for Chapter 5, where we will explore how integrating task- and course-specific open-ended ESM survey items influence our interpretation of task-specific self-efficacy scores and how this, in turn, affects the nature of daily journal prompts.

4.8.2 Investigating the Integration of Quantitative and Qualitative Data: Designing Daily Journal Prompts to Capture the Theoretical Foundations of Self-Efficacy

Although this analysis serves as a proof-of-concept for integrating quantitative and qualitative data, there is potential to improve the process of refining how we capture rich qualitative data. For example, more targeted prompts could encourage deeper reflections on specific moments captured in the ESM survey, offering valuable context about why or how certain experiences interact with self-efficacy. Additionally, analyzing multiple data points across different contexts (e.g., various courses or tasks) to be able to explore how self-efficacy develops over time.

Although Jane's responses confirm that Notification 8 represents a high moment of self-efficacy, a more detailed description of the experience would help researchers and educators understand the factors and context supporting her self-efficacy at that moment. More broadly, eliciting rich responses from multiple students about their particular experiences would provide valuable insight into how specific moments or events influence self-efficacy. This can be facilitated through carefully designed daily journal prompts. By examining the prompts written across Fall 2020 and Fall 2021 iterations of the mixed methods approach, we can ensure that the intention behind our prompts capture the theoretical foundations of self-efficacy (i.e., domains and sources) in relation to a specific moment.

In Chapter 6, I will report on the process of designing a codebook to systematically capture the ways in which our daily journal prompts are in line with the theoretical foundations of self-efficacy. Then, we will present an analysis of the codes applied to the two iterations of this mixed methods

69

approach (i.e., Fall 2020 and Fall 2021). This analysis serves two purposes: (1) to examine how the integration of open-ended survey responses impacted the design of the daily journal prompts in exploring the theoretical foundations of self-efficacy and (2) to provide deeper insight into the ways self-efficacy can be understood.

We will also examine how the integration of open-ended survey responses impacted the daily journal prompts written across the two iterations of this mixed methods approach (i.e., Fall 2020 and Fall 2021), providing deeper insight into the ways self-efficacy can be understood and measured.

4.8.3 Establishing the Need for Deeper Integration: The Role of Evidence in Linking Quantitative and Qualitative Data for Valid Self-Efficacy Claims

Although Jane's response to the daily journal prompt could be interpreted as confirming that, at Notification 8, she reported high task-specific self-efficacy, there is ambiguity regarding whether the moment Jane reflects on in the daily journal prompt's response is the same one she reported on the ESM survey. For example, Jane writes, "Yesterday in [physics course] I felt confident in my knowledge and experiment...", but it is unclear if she was referring to any other moments associated with her physics course throughout the day, which complicates the connection between the journal response and the ESM survey data.

This raises an important issue: the ability to link the ESM survey response with the daily journal prompts depends on how clearly the task and context are defined. Although both data sources are related to the same course, they do not necessarily reflect the same task or moment in time. To establish a valid connection between ESM data and the journal prompt and response, the task described in both would need to be aligned to be able to explore task-specific self-efficacy. However, a question arises: can the task be the same across different courses, or must it refer to the same task within the same course for the linkage to be meaningful?

Thus, the specificity of the connection between the ESM survey data and the journal response depends on the consistency with which the task and course information is shared between both data sources. To strengthen the credibility of linking these two data sources in terms of task-specific self-efficacy, it would be essential to identify or clarify the specific tasks involved in both the ESM

70

survey and the prompt and response to the journal. Having this task-level information would help ensure that the journal response corresponds to the same moment as the ESM survey response, ultimately allowing for a more robust claim that the two data sources are indeed linked.

The link between the ESM survey and the daily journal prompts is critical to interpreting the data in a meaningful way. If the goal is to use both data sources to provide a complementary or holistic view of a student's self-efficacy, it is essential that they focus on the same domains of self-efficacy, or at the very least, the specific domain of self-efficacy around which the claim is being built. Without this alignment, the data sources may not fully or accurately reflect the same aspects of the student's self-efficacy, which undermines their ability to offer a cohesive, integrated understanding of the student's self-efficacy. Therefore, ensuring that the task or domain in both the ESM survey and the journal prompt and response are the same is a necessary step to strengthen the validity or credibility of any conclusions drawn from combining these sources.

This idea of task alignment and data integration is what Chapter 7 addresses by exploring how different types of evidence, both quantitative and qualitative, can be integrated to provide a more accurate and holistic representation or understanding of a student's self-efficacy. Chapter 7 will provide a deeper investigation into the process of linking these data sources and offer strategies to improve the validity of the claims about self-efficacy that arise from these integrated data.

CHAPTER 5

EXPLORING THE IMPACT OF OPEN-ENDED SURVEY RESPONSES ON INTERPRETING TASK-SPECIFIC SELF-EFFICACY AND JOURNAL PROMPT DEVELOPMENT

5.1 Introduction

Research on self-efficacy (SE), or the belief in one's ability to successfully complete a task [18], has been shown to predict students' achievement in science courses [88], their persistence in science majors [89], and plays a role in students' science career choices [90]. Quantitatively, students' science SE has traditionally been measured through pre- and post-surveys to explore shifts in students' SE from the beginning to the end of the semester [41, 50, 51]. However, this approach limits our ability to investigate individual students' SE (REF) while understanding the impacts that particular events have on a student's SE [19] – these measures involve a large time gap between a student experiences an event versus when they report on their SE in relation to that event [19]. To better understand the relationship between particular events and an individual's SE, researchers have used qualitative methods such as interviews [19, 21]. While interviews invite participants to share the experiences they perceive as most salient to them from their past experiences, they are traditionally conducted in such a way that a time gap still exists between when students experience events and when they report their SE in relation to those events. With traditional quantitative and qualitative methods, participants draw on their memories of events a considerable time after these events have occurred, which can result in missing critical aspects of how these events are important in understanding one's SE.

To overcome the limitations of these traditional methods, researchers have utilized a method called the Experience Sampling Method (ESM) [80]. This is a technique for investigating the lived experiences of individuals in the moment. Researchers have used this method as a quantitative approach to investigate the gender gap in SE between women and men in different courses [26]. They found that women experienced lower SE within interactive-engagement physics courses than men, but this difference was not present when analyzing the comparison of their SE across other courses. Although this study opened the door to a new quantitative approach to studying SE, the

72

study was limited in that the work was unable to account for how specific features of instruction influenced students' SE. Thus, their work showed that ESM could capture students' fluctuations in their SE closer to when events occur, but an additional data source would be needed to be able to investigate how or why particular events influenced students' SE [26].

In this work, we will present a novel mixed methods research design that pairs the ESM with individualized daily journal prompts to mitigate some of the limitations of traditional research methods used to study SE. The combination of these two methods allows for real-time tracking of students' SE fluctuations while providing the opportunity for personalized reflections on their experiences in relation to those SE fluctuations. Throughout this paper, we will demonstrate how our data collection process changed over two iterations and how this change impacted our ability to investigate students' SE in relation to their daily experiences. We will explore this through a critical aspect of mixed methods research - the integration between the quantitative and qualitative data sources within our mixed methods design [13, 71]. Specifically, we will introduce two instances of integration that occurred during the data collection period:

1. The integration of the open- and close-ended items from the ESM survey. This builds a coherent thread by linking the students' numerical responses with their written responses.

2. The integration of the identified thread with the individualized daily journal prompts. Once key threads are identified from the ESM survey data, they inform the development of individualized daily journal prompts, which create a more contextually relevant and personalized reflection process, allowing students to reflect on moments from their day and the particular experiences that are influencing their SE.

This is not to say that all points of integration in our mixed methods design are discussed within this paper, but rather we are focusing on these two particular points, as they are critical in eliciting student reflections about moments from their day. Thus, this paper will explore how adding the first point of integration to the second iteration impacted our ability to investigate students' SE

73

in relation to their daily experiences. We will explore the points of integration by answering the following research questions:

1. How does the integration of the open- and closed-ended ESM survey data support the data team in identifying key threads in students' SE fluctuations?

2. How does the integration of open- and closed-ended ESM survey data influence the construction of the individualized daily journal prompts?

3. What are the advantages and limitations of our mixed methods research design?

We will begin by reviewing the background literature on SE and motivate the need for a mixed methods research design that intentionally explores the theoretical foundations of SE (e.g., domains and sources). In Section 5.3, we will present our overall research design, the contexts in which it was applied, and outline the differences between the two iterations of our data collection. We then will present our methods of analysis followed by the results of a nested and embedded case study. Finally, we will synthesize the advantages and limitations of this novel mixed methods research design while calling for more mixed methodological approaches when studying a complex construct like one's SE.

5.2 Background Literature & Purpose

In the following section, we summarize the theoretical foundations of SE and explore how researchers have employed both quantitative and qualitative methods to study students' SE within STEM education, particularly in physics. Given the complexity of the construct, SE, a mixed methods approach is necessary to capture both measurable trends and in-depth experiences, particularly when investigating how specific events influence students' SE. However, integrating quantitative and qualitative methods introduces its own intricacies; specifically, determining how to effectively combine these approaches in practice is a challenge that requires careful consideration to address the inherent limitations of each. This section will discuss the theoretical foundations of SE, the various methods used to study it, and the role that mixed methods can play in deeply understanding such an important motivational construct.

5.2.1 Theoretical Foundations of SE

SE refers to an individual's belief in their ability to successfully perform a task [18] and has been identified as a predictor of academic motivation, goal setting, performance, and persistence [91]. Specifically within the Science, Technology, Engineering and Mathematics (STEM) fields, SE has been shown to predict achievement [55, 88, 92], persistence in science majors [75, 89, 93, 94], and science career choices [90, 95, 96]. As such, students' beliefs in their academic abilities play a critical role in how they navigate academic spaces such as classrooms, tutoring centers, and research labs.

Examining how SE evolves as students interact with the academic systems can guide the design of environments that foster positive SE formation and contribute to students' overall success. However, in doing so, it is important to ground this examination within the theoretical foundations of SE as they provide insight into how students' beliefs develop across different contexts (e.g. domains) and from different experiences (e.g. sources).

5.2.1.1 Domains of SE

Specifically, Bandura emphasized that SE must be tailored to a specific domain that has varying levels of tasks within those domains [18]. For example, a person may have high SE in performing a particular task in one context (e.g., doing physics homework) but low SE in performing the same task in another context (e.g., doing chemistry homework). This emphasizes the need to clearly define the activity area of interest, analyze its different aspects, identify the capabilities required, and consider the range of situations where these capabilities might be applicable [59].

As such, researchers have operationalized SE in a variety of ways. Typically, the generalization of SE depends on the similarity of the tasks within a specific context, also recognized as task-specific SE [18, 25, 47, 97]. For example, discipline-based SE refers to an individual's confidence in their ability to perform tasks within a specific discipline or field of study [27]. As a result, many studies within physics education research have explored students' physics SE, one's confidence in their ability to perform tasks related to physics [35, 51] as well as comparing students' belief in their ability to perform similar tasks across the STEM disciplines [25].

Other researchers have distinguished SE according to the level of specificity of the task at hand. In addition to task-specific SE, this characterization includes course-level SE - an individual's belief in their ability to earn a passing grade in a specific course [19, 49], degree-level SE - an individual's belief in their ability to complete the requirements for a planned degree program [52], and career-level SE - an individual's belief in their ability to complete the educational requirements and job duties for the planned career [49, 98]. The differences in operationalization highlight the need for decisions regarding the scope and focus of the task while taking into account the domain-specific nature of SE. This has resulted in many questions about how to study SE domains [18, 27].

5.2.1.2 Sources of SE

In addition to examining the contexts in which SE evolves, it is also essential to consider how students' interpretations of the experiences they encounter within academia affect their SE. Bandura posited four general kinds of experience, or sources, that can impact one's SE [18].

The first source and generally the most studied are mastery experiences. These kinds of experience are those where students interpret their SE drawing on past success or failures on the same or similar kind of tasks [18, 19]. Researchers have claimed that this is the most influential source for interpreting one's SE [18, 20].

The second source Bandura posited is vicarious learning experiences [18]. These experiences occur when students are observing others perform a task, in which they use this information to evaluate their own likelihood of success at the same or similar tasks [12, 18, 19]. This may result in experiences that threaten or support students' SE depending on how they evaluate their performance relative to others.

Furthermore, social persuasion experiences are verbal and nonverbal messages from others about one's ability to perform a task [18, 19, 99]. Often, researchers explore this source by focusing on the encouragement or discouragement a person receives while performing a task [58]; however, since people are not always aware of the messages they encounter, studying implicit messages poses a challenge [100]. Consequently, many questions remain about how people interpret these social persuasion experiences in relation to their SE.

76

Lastly, physiological state experiences occur when individuals use their emotional state to evaluate their confidence in their ability to perform a task [18, 19]. It has been mentioned in other ways such as "emotional arousal", "arousing experiences" or "affective states" [101] and it accounts for how anxiety, stress, or other feelings can influence one's evaluation of their ability to perform a task. Researchers have had difficulty in disentangle the physiological state on its own from the other three sources [18, 19]; in turn, it has historically been challenging to study.

Overall, understanding these four sources is critical, as they shape how students interpret and internalize their experiences, which ultimately influences the development of their SE in different contexts. In the following section, we will summarize the methodological approaches that have been used to investigate students' SE and how the limitations of the various approaches interact with the theoretical foundation of SE.

5.2.2 Methodological Approaches to Examining SE

In general, SE has been studied using a variety of methodological approaches. Quantitative approaches, such as Likert-scale survey items, provide a value for a student's SE in relation to a task within a specific domain; however, these measurements are limited in investigating how specific moments influenced a student's SE [19]. In contrast, qualitative methods, such as interviews and thought-listing, offer rich, contextual insights into the factors shaping students' SE [99]; however, it is difficult to interpret the size of the impact using these approaches. In the following, we will discuss in detail how students' SE has been investigated while outlining the inherent limitations for each of these methodological approaches.

5.2.2.1 Quantitative Approaches

Self-efficacy is often measured quantitatively by using Likert scale surveys [20] which capture a student's SE at a single point in time. One-time measurements are typically intended to be used for comparing students' SE across different demographic groups [29, 33, 58, 102, 103] or to predict an outcome of success [18, 28]. For example, researchers have used multiple regression analysis to examine how factors such as gender, degree, and prior coursework predicted students' SE at the start of a semester [50]. Multiple-time measurements, such as a pre- and post-test design, are

leveraged to explore changes over a specific period of time, such as a semester or course sequence. Within the same study, Ainscough *et al.* measured SE using pre- and post-tests to identify shifts over the academic semester [50].

However, these approaches have limitations. One-time measurements provide a snapshot of a student's SE at a particular moment but fail to capture the evolving nature of SE, potentially overlooking prior experiences that influence it. Similarly, multiple-time measurements, while valuable for detecting changes in students' SE, may miss important events or experiences that specifically affected their SE. For example, a survey response might show a positive shift in SE, but this shift could be due to different underlying changes in individual students' perceptions of the task at hand [104]. In addition, these methods also tend to aggregate data at the group level, which generally limits the understanding of individual students' unique experiences and how they shape their SE.

In addition to the limitations of the type of measurement, the nuances of the theoretical foundations of SE present challenges to the design of a quantitative survey tool. With a lens toward the identified importance of the domain aspect of one's SE, the choices that are made when designing an SE survey have implications toward the operationalization of the domain and the tasks within the domain. For example, a general physics SE survey such as the Physics SE Survey (REF) can provide a broad overview of students' confidence in doing the physics that is taught in the classroom, but it may limit the kind of tasks physicists engage with more broadly. Recognizing that different content areas of physics may lead to different tasks, other researchers have more recently created targeted instruments to assess students' SE in understanding content areas such as stars and quantum mechanics. However, these tools focus on capturing the students' confidence in their conceptual understanding of these topic areas, which may not account for how other types of academic tasks may impact students' SE [37, 105]. Although there are many domain-specific SE surveys used by educational researchers [22, 23, 37, 47, 49, 105], balancing the number of items on a given survey while being transparent about the specificity of the domain can bring limitations to the survey design process.

Historically, there has been much less of a focus on SE surveys accounting for the identified importance of the sources of SE [20]. However, a well-known and widely used SE instrument that was developed with a sources of SE lens is the Sources of SE in Science Courses - Physics (SOSESC-P) [34]. This instrument assesses students' SE in physics using the four main sources of SE that were outlined in Section 5.2.1.2.

A limitation of this particular SE survey is its design around the source of social persuasion. The survey examines influences from instructors and peers but does not account for influences from other people in students' lives that may have an impact. Additionally, individuals may weigh the impact of the social persuasion experience differently depending on who is doing the persuading. For example, the instructor may have more weight on the student's SE than their peers because of their position within the classroom, or peers may have more weight as students may be more likely to have more interactions with their peers. This nuance of the social persuasion source of SE extends to the other sources as well. Mastery experiences, vicarious learning, and physiological state are also measured without considering the depth and context of students' interactions and experiences. As a result, while the SOSESC-P provides a useful tool for measuring the sources of SE, the survey may miss important contextual details as well as additional factors that students are using to interpret and evaluate their SE.

In general, surveys as a traditional quantitative approach to measure SE may focus on specific aspects of ones' confidence that researchers are interested in, but they often fail to capture the multitude of ways in which individuals interpret and evaluate their beliefs and experiences [20]. Thus, this critique points to the need for a more holistic approach that can account for various "personal, social, situational and temporal" factors influencing students' SE beliefs [20].

5.2.2.2 Qualitative Approaches

While quantitative measures can provide valuable insights into an individual's shift in SE, they have inherent limitations that can restrict a full understanding of the nuanced dynamics involved. As as result, a variety of qualitative approaches have been leveraged to gain deeper insights into students' experiences as it relates to their SE (REF). Although experiences can create opportunities

79

for developing SE, this does not mean they do (Lent et al., 1996; Sawtelle et al., 2012). In order for experiences to impact SE, individuals decide which information to focus on, evaluate its importance, and then combine it when judging their SE (Lent et al., 1996) – observations that only qualitative approaches can begin to capture.

To explore how individuals interpret and evaluate their SE, there are a diverse range of qualitative methods that have been utilized such as interviews (Citations), thought-listing (Lent et al., 1996; CITATION), open-ended items on surveys (Lin et al., 2017), weekly journals/prompts (CITATIONS), and video recordings (CITATIONS). Interviews, the most common qualitative approach to exploring one's SE, have been used to explore personal narratives and reflections, allowing researchers to capture the many nuances of the theoretical foundations of SE (Stoeckel et al., 2021; Zeldin and Pajares, 2000). For example, one interview study explored the personal narratives of 15 women who have chosen and continued their careers within STEM to understand how their SE beliefs influenced this decision. This approach allowed for an in-depth exploration of individual experiences, and found that social persuasion and vicarious learning experiences were crucial for women's SE beliefs.

Similarly, weekly journal prompts have also been employed to guide individuals in reflecting on experiences that support or threaten their SE (Fong and Krause, 2014; Stoeckel et al., 2021). One advantage of weekly prompts compared to interviews is the reduced time gap between the experience and reflection, allowing for more immediate insights. For instance, Fong and Krause (2014) utilized weekly journal entries in an educational psychology course, prompting students with questions such as, "What are the factors that make you feel confident?". This approach aimed to uncover broader sources of SE. In contrast, Stoeckel et al. (2021) focused on experiences related to students' abilities to complete weekly quizzes, providing prompts at the end of each quiz to explore what bolstered their confidence. As such, weekly prompts serve as a valuable platform for individuals to articulate their thoughts and experiences, encouraging reflection closer to when events occur.

A strength of many of these qualitative approaches to exploring SE is that they allow individuals

the flexibility to share their experiences in detail and how they are using information to assess their SE. By focusing on individual stories, researchers can identify how various experiences contribute to students' SE beliefs. This flexibility allows researchers to explore the ways in which social, cultural, and emotional factors may impact SE. In general, qualitative methodologies provide valuable insight into the intricacies of SE by moving from a simple numeric measurement to understanding the lived experiences of people more generally.

However, just like quantitative approaches, all methods come with limitations. While interviews and weekly journal prompts may offer rich data that reveal the nuances of one's SE, they often capture only snapshots that heavily rely on participants' memory recall. To mitigate the reliance on this recall, Sawtelle and colleagues were interested in capturing real-time influences on students' SE through the use of video recordings to examine opportunities that serve as influences on students' SE as they were solving physics problems. While video analysis may provide immediate insights into human interaction, they still rely on post hoc interviews for validation of a researchers interpretation of the video data.

In general, many qualitative approaches that have been used to investigate students' SE have this limitation of memory recall; how much time there is between a specific event and when a participant has the opportunity to reflect or validate the details of that moment. Moreover, understanding the impact of an event or how one weighs their various experiences that influence their SE is a challenge for many qualitative approaches (REF). Interviews, weekly journal prompts and even video recordings aren't able to account for a participant's perspective in real-time which leads to the need for more robust methodologies when exploring constructs such as SE. In the next section, we will present an argument for using a mixed methods approach when studying the development of students' SE.

5.2.3 The Value of Mixed Methods Research

The intentional combination of quantitative and qualitative methods is known as a mixed methods research design. However, beyond this general statement, researchers interpret and apply it in diverse ways. Creswell notes that definitions of mixed methods should encompass a variety of

different aspects such as the methods, the design, the purpose, and the philosophical orientation, but highlights that there is there is a lack of consensus between experts [64]. And although these aspects are critical, researchers often oversimplify qualitative and quantitative research by associating qualitative research with words and quantitative research with numbers [61]. Much of the literature identifies that the combination of these broad definitions with little consensus overlooks the nuance in how data is conceptualized within a design, which is essential for understanding the nature of inquiry of the research [61].

Tashakkori and Teddlie (2014) define mixed methods as research that collects and analyzes both qualitative and quantitative data sources, integrates the findings, and draws inferences from both sources [60]. Others emphasize core characteristics of mixed methods research such as the collection and analysis of both types of data based on the research questions, the integration of the data and the results, the organization of procedures and research designs, and the application of theory and paradigms to frame the study [64]. A common feature of both definitions above [60, 64] is the concept of integration. Different scholars conceptualize integration in a variety of ways and in turn, these differences in the conceptualization of integration can deeply influence the design and outcomes of a research study. Below, we will highlight the need for a clearer understanding of what integration entails in practice and our effort in contributing to the conversation as it relates to how we've designed our mixed methods approach for exploring the development of students SE in real-time.

5.2.3.1 Integration in Mixed Methods Research

Researchers employ a mixed methods approach because it adds value by combining quantitative and qualitative methodologies, aiming to address and overcome the limitations inherent in each [71]. This added value is achieved through integration, which allows for a more comprehensive understanding of the phenomenon of interest [2, 64]. Achieving meaningful integration at different stages of the study, as well as across different elements of the research design, helps capitalize on the full value of mixing methodologies to answer novel research questions [2, 106].

As the research about mixed methods continues to evolve, the integration of quantitative and

qualitative components remains a central focus. Researchers have proposed that greater levels of integration can improve the quality and rigor of mixed methods studies (Yin, 2006; Bazeley and Kempj, 2012; Teddlie and Tashakkorri, 2009), prompting efforts to develop frameworks and strategies that facilitate effective integration [62, 63, 69, 70, 107]. Part of this emphasis arises from the challenges associated with achieving meaningful integration in practice, with many researchers acknowledging that integration serves as both a process (the ability to combine quantitative and qualitative strands) and as a product (demonstrating that quantitative and qualitative strands are effectively integrated) [66]. A notable example is Vicki Plano Clark's identification of four integration strategies: (1) formulating integrative research questions, (2) aligning quantitative and qualitative data sources, (3) identifying points of integration, and (4) developing joint displays and mixed interpretations [70]. Our goal for this work is to focus on one of these strategies, specifically identifying points of integration, and how adding an additional point can influence the outcomes of our study.

5.2.3.2 Identifying Points of Integration

Identifying points of integration refers to determining specific points or moments in a study when the quantitative and qualitative strands intersect, interact, or influence each other [70, 108]. This strategy involves not only recognizing where the types of data meet but also considering how these points can be leveraged to generate more comprehensive insights. These integration points may occur at various stages of the research process, such as during the data collection, analysis or interpretation, and are crucial for ensuring that both quantitative and qualitative strands inform and enrich one another.

For example, Fetters and Molina-Azorin (2017) describe the data collection dimension of integration as collecting both qualitative and quantitative data with intentional alignment to the overall mixed methods research procedures [63]. This means that the process of data collection is not just about gathering different types of data in isolation; rather, the collection of each dataset is purposefully designed to complement and interact with the other. This intentional approach ensures that both strands of data are integrated in a way that supports the research questions, methods, and

theoretical framework of the study, contributing to a holistic understanding of the phenomenon of interest. For example, in our context, to capture a more nuanced picture of SE, it is important to complement the quantitative and qualitative data that explore the underlying perceptions, beliefs, and contextual factors influencing students' SE. Integrating these strands can offer a more holistic view of how and why SE develops and changes over time.

5.2.4 Our Contribution to the Integration Conversation

In this paper, we aim to demonstrate the impact of adding a new point of integration to our mixed methods research design intended to study the development of students' SE. Our goal is to be more transparent about what, where, and how integration occurs within the data collection phase of our design. We will use the analytic technique known as "following the thread" which generally begins with an initial analysis of data to identify key themes and analytic questions for further exploration [69, 72, 73]. However, we are extending this analysis technique beyond its traditional conceptualization by applying it during the data collection phase of our study. In general, we will discuss how having an additional point of integration can support the identification of key threads which in turn, influences how to gain meaningful insights into the development of students' SE. By introducing our novel mixed methods design, our study contributes to the ongoing conversation about enhancing the integration process and provides a practical example of how theoretical strategies can be implemented in a complex research design.

5.3 Research Design & Context

To be able to explore students' SE closer to when events occur, we developed an explanatory sequential mixed methods design that utilized surveys via the ESM followed by an individualized daily journal prompt. In this section, we will provide a general overview of the research design, a detailed description of the specific methods of collection for each data source (e.g. quantitative and qualitative), and the context in which we explored how adding a point of integration during data collection supported the writing of the individualized individualized daily journal prompts.
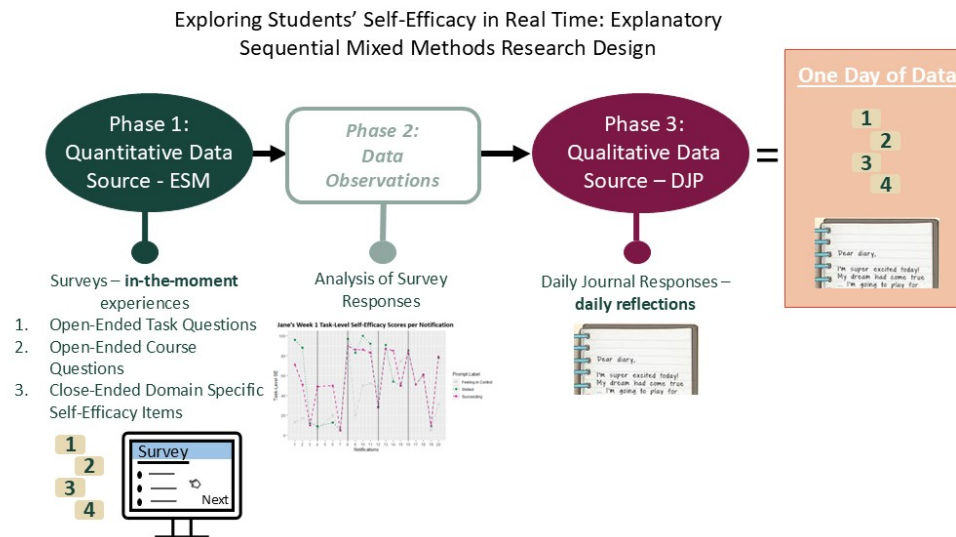
Figure 5.1 The explanatory sequential mixed methods design for one day of data collection. This design lead with a quantitative data collection and analysis that informs a qualitative data collection.

5.3.1 Overview of the Research Design

Figure 5.1 shows a general snapshot of one day of data collection. Our design is led by the quantitative data source – ESM surveys intended to probe in-the-moment experiences (Phase 1). In our design, we randomly notified participants four times throughout the day via a Smartphone application called ExpiWell (REF). For each notification they were asked to respond to short, open- and close-ended survey items regarding their SE.

Once the ESM survey responses were collected, we observed each student's ESM survey responses by plotting their task-specific SE scores (Phase 2) and looking for noticeable increases or decreases in their scores over the course of the day. These fluctuations allowed us to identify moments where the student experienced a change in their SE related to specific moments in their day. The goal of this is to observe patterns that can identify key trends in their SE that guide the creation of an individualized daily journal prompt. The implementation of Phase 2 changed over two iterations of data collection and will be explained in further detail in Section 5.3.2.

Following the completion of the four in-the-moment ESM surveys and the researcher's observations of the survey responses, we crafted individualized daily journal prompts for each participant (Phase 3). The daily journal prompts were written to probe into moments identified from the ESM surveys and to solicit information about students' SE in more detail. This qualitative data source was intended to encourage the participants to reflect on their experiences from the day in relation to their SE.

5.3.1.1 Phase 1: Quantitative Data Source – ESM

Generally, the ESM is a method to examine both the context and content of an individual's lived experienced by asking the participants to provide self-reports regarding their experiences at different moments during their daily lives [80]. In applying the ESM, one can capture a participant's episodic representation of the context and factors influencing them rather than relying on their memories after a longer time gap. As described above, we used open- and close-ended survey items as the means to collect the quantitative data in our mixed methods research design.

The open-ended survey items were used to collect information about the students' external context during a particular moment. For example, we asked them *"What is the main thing you are doing?"*, *"Where are you doing this activity?"*, and if this activity was for a course *"What is the name of the course?"*. The closed-ended survey items were used to collect information about the internal content of their experience and were adopted items from other research studies that probe into domain-specific SE [25, 49]. The task-specific SE items were shown to be valid and reliable measures of task-specific SE within a high school context [25] while the course-level and career-level SE items were shown to be valid and reliable within a physics context [49]. The full ESM Survey is presented in Appendix A; however, for the purposes of this work, we will only be using the open-ended survey items and the close-ended task-specific SE items, shown in Table 5.1.

5.3.1.2 Phase 3: Qualitative Data Source – Individualized Daily Journal Prompts

Individualized daily journal prompts are a method used to capture students' daily experiences while closing the gap between when specific events took place and when they were discussing their SE in relation to those events [109, 110]. In pairing the daily journal prompts with the ESM

86

surveys, our aim was to gain insights into why students' SE is fluctuating across the day. There may not always have been moments from the ESM survey to investigate, but in these cases, we chose to provide students with generalized daily journal prompts to investigate their SE.

As a result daily journal prompts covered a wide range of ways in which we elicited more details about the students' daily experiences. For example, we wrote prompts to investigate further information about a particular moment, to inquire about their SE in relation to a specific domain (i.e., task, course, degree, career) and to explore explanations for an increase or decrease in their SE scores that we observed from the ESM survey. In general, to write individualized daily journal prompts, we leaned on the complex theoretical foundation of SE such that our prompts probed various domains and the four sources of SE [18]. For example, if we observed a shift in the ESM data where the student was working with others on physics homework, then we would write prompts to investigate how these interactions, either through a vicarious learning or social persuasion experience, influenced their belief in their ability to do their physics homework.

5.3.2 Phase 2: An Added Point of Integration

We employed our mixed methods research design across two iterations, Fall 2020 and Fall 2021. Across these two iterations, we modified Phase 2 in our overall design, specifically how the observations from the ESM survey responses were used to write the individualized daily journal prompt. As shown in Figure 5.2, we solely used the closed-ended task-specific SE scores for

Survey Block	ESM Questions
Open-Ended Task	(Q1) What is the main thing you are doing right now? (Q2) Where are you doing this activity?
Closed-Ended Task-Specific Self-Efficacy	(Q3) How skilled are you in the activity? (Q4) Do you feel in control of the situation? (Q5) Are you succeeding at what you are doing?
Open-Ended Course	(Q6) Is the activity you are doing for a course you are currently enrolled in? (Yes, No) (Q7) What is the name of the course? Please write the course label and number (e.g. PHY123) in answering this question.

Table 5.1 Task-Specific ESM Survey Questions: Q1 and Q2 are the open-ended task items, Q3, Q4, and Q5 are the closed-ended task-specific SE items measured on a scale from 0 - 100, and Q6 and Q7 are the open-ended course items.
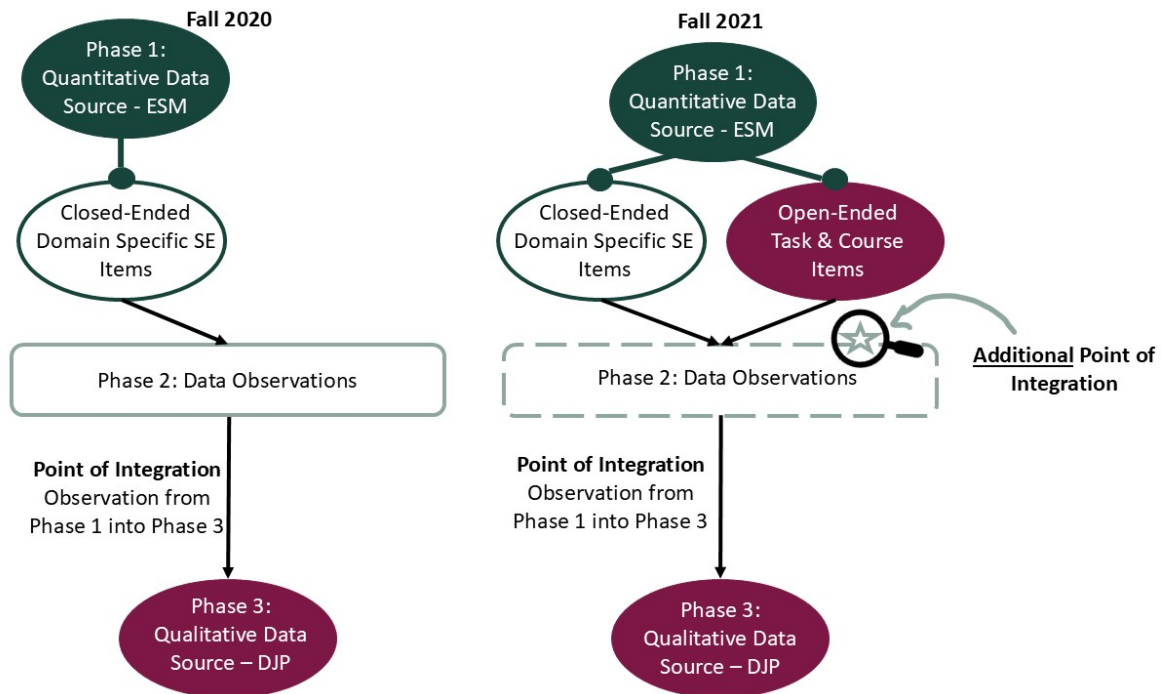
Figure 5.2 Mixed Methods Data Collection Procedure Across Fall 2020 and Fall 2021. The black arrows indicate the flow of procedure (e.g., moving from Phase 1 to Phase 3). In both iterations, Phase 1 of data collection involves gathering quantitative data through the ESM surveys. The next phase, "Data Observations," refers to the process by which researchers draw observations from the ESM data to inform the writing of the daily journal prompts. The Fall 2020 iteration involved only drawing observations from the ESM survey items, highlighted in the picture through showing what data was collected and used in Phase 2. In contrast, in Fall 2021, researchers used both the closed-ended and open-ended ESM survey item responses together to draw observations from the ESM survey data, as highlighted in the diagram through showing both types of data collected and used in Phase 2. The inclusion of open-ended items in Fall 2021, and the merging of these with the closed-ended SE items, adds an additional point of integration, enriching the process of writing the daily journal prompts by combining qualitative responses with the quantitative task-specific SE data. The diagram does highlight a separate point of integration that is occurring overall which is using the ESM data observations to inform the writing of the daily journal prompts, shown through "Point of Integration Observation from Phase 1 into Phase 3".

the observations that were made to inform the writing of the daily journal prompt in Fall 2020. During this iteration, we assumed that these scores directly reflected students' confidence in the specific task they are performing at the exact moment they complete the survey but only wrote individualized daily journal prompts around shifts that we observed in the plotting of the task-specific SE responses. There was no specificity of the context in these individualized daily journal prompts which was a limitation of this iteration of data collection.

Overall however, the intent of using the ESM is to investigate the context of students' daily lived experiences and we recognized that we were not leveraging the open-ended task items to ensure the connection between the specific task reported and the task-specific SE responses. In Fall 2021, the open-ended task survey responses were an added piece of the observation that was used to inform the writing of the individualized daily journal prompt. For example, if a student reported meeting with an academic advisor (a moment that we hypothesize would influence their SE), we could explore this interaction through a contextualized daily journal prompt.

Through a nested and embedded case study design, we examine how this additional point of integration influenced the evolution of our overall mixed methods approach—specifically, how it supported the research team in identifying key analytic threads in students' SE fluctuations and informed the construction of individualized daily journal prompts (daily journal prompts) designed to elicit meaningful reflections on students' experiences. To do this, we compare how the ESM data functioned across two iterations of data collection—Fall 2020 and Fall 2021—in order to understand how the methodological evolution shaped the types of insights generated. To situate this comparative analysis, we first provide an overview of each data collection period and outline the structure of our research design in both iterations.

5.3.2.1 Fall 2020 Data Collection

In Fall 2020, data collection involved administering ESM surveys multiple times per day alongside daily journal reflections. Students completed the ESM survey four times a day randomly within four time windows between 9 AM and 6 PM. The ESM surveys measured domain-specific SE (i.e., task-specific, course-level, and career-level), capturing students' perceptions of their

confidence in their abilities in the moment.

After 6 PM, individualized daily journal prompts were delivered to the students. The journal prompts in this iteration were designed primarily based on students' task-specific SE scores, if there were any observable fluctuations in the students' task-specific SE scores. This allowed us to explore the fluctuations in their task-specific SE within their daily experiences. The prompts aimed to elicit reflections that expanded on how or why students rated their task-specific SE in this way.

5.3.2.2 Fall 2021 Data Collection

The Fall 2021 data collection maintained the structure of ESM surveys and daily journal prompts but introduced a key methodological refinement: the use of students' own language from their open-ended survey responses to examine the ESM task-specific SE scores and was used to shape the journal prompts.

This added point of integration allowed for greater individualization of the journal prompts. In Fall 2021, prompts were not only tailored based on SE scores but also crafted using students' own wording from their open-ended survey responses, making them more personally relevant. This approach created a direct link between the survey and journal data. By incorporating students' language, we bridged the gap between their real-time survey responses and their journal reflections. For example, if a student reported working on a physics assignment in a survey, we would craft a prompt specifically about their experience with that assignment, prompting them to reflect on their SE related to that task. This connection encouraged students to elaborate on specific moments they reported in their ESM surveys and explore how those moments influenced their confidence in a particular context. This methodological evolution reflects our ongoing effort to strengthen the alignment between ESM data and journal reflections, ultimately improving the mixed methods approach to exploring SE in real time.

In Fall 2021, given our positionality as physicists, we recruited from the upper division physics courses as we were interested in physics transfer experiences and to allow students who participated in the Fall 2020 iteration to participate in the Fall 2021 iteration. In addition, to diversify the overall experiences we can investigate with this mixed methods approach, we expanded our recruitment

90

of participants to include students enrolled in other STEM courses. We received a list of STEM courses within the institution that had high numbers of post-transfer students participating in them; thus, identified the five courses with the largest post-transfer student attendance - introductory statistics , introductory calculus , introductory chemistry , first organic chemistry within the organic chemistry sequence, and introductory biology. From this list, we recruited from the introductory chemistry course anticipating an overlap with the students in the biology and calculus courses; the organic chemistry course and statistics course may have reduced the kinds of majors that could participate in the study so we did not focus our efforts there.

From these courses, we asked faculty to email, on our behalf, the research opportunity to their students. Interested students were invited to complete a recruitment survey, sent out as part of this email. The recruitment survey included items that asked for their transfer information (e.g. how many credits they had transferred, where they had transferred from), first-generation status (i.e., the education levels of their parents and if they identified as a first-generation), major/minors, class standing, and the STEM courses they were taking. Demographics (e.g. gender, race, disability and others) were not collected during recruitment stage of the study.

Approximately 255 students completed the recruitment survey by the deadline. We invited 30 students to participate, prioritizing transfer and/or first-generation college students who were STEM majors, starting with physics students, supporting the design with investigating degree-level SE through including students from different majors. The first 20 were selected based on their minors and courses to ensure diverse STEM experiences. For example, students minoring in languages (e.g., Spanish, Chinese) or in STEM fields like data science or computer science were favored for diversity and breadth of coursework. We also considered students involved in programs supporting transfer, first-generation, or marginalized communities, as these could influence their STEM SE experiences. For the final 10, we prioritized program involvement, transfer/first-generation status, and STEM major/minors. Programs were considered to enhance sample diversity, influence students' SE, and avoid selecting students with similar demographics. Of the 30 students, 24 transferred credits from another institution, including community colleges,

two-year, or four-year colleges. The remaining 6 did not transfer any credits. Sixteen students identified as first-generation college students, though some considered themselves first-generation if their parents had some college experience but didn't graduate. Two students were uncertain, and 12 did not identify as first-generation.

The 30 students were first asked to completed the pre-sign-up survey, which explained the study, obtained consent, and provided instructions for signing up for the Expiwell smartphone application. After creating Expiwell accounts, students completed the initial sign-up survey, which included degree and career-level SE data. Of the 30 invited students, 18 completed the initial sign-up survey and continued into the primary data collection. Of these, 15 students transferred credits and 11 identified as first-generation. Table 5.2 shows the majors and minors of the 18 students who participated in the study. Among the 18 students, one student was a double major, and 11 were pursuing minors, with some minoring in two disciplines. Three students had non-STEM minors, including two in language fields.

Name of Majors and Minors	Number of Majors	Number of Minors
Physics	6	0
Astrophysics	2	0
Chemistry	0	1
Mathematics	0	2
Biology	5	0
Data Science	1	1
Computational Modeling	0	7
Computer Science	0	1
Engineering	2	0
Neuroscience	2	0
Non-STEM Majors	0	3

Table 5.2 Shows the majors and minors of the 18 students who participated within the Fall 2021 study. For Biology, the institution offered distinct majors for different sub-disciplines of biology, and we chose to represent all those distinct majors within "Biology" (i.e., Biology, Biochemistry, and Molecular Biology, etc.) below. Similarly, the institution offered distinct majors for different sub-disciplines of Engineering (i.e., electrical, civil, computer, etc.), and we chose to represent all those distinct majors within "Engineering" below.

As an incentive, students received $50 per week of data collection. In Weeks 1 and 2, students

earned the full incentive by completing 16 out of 20 Expiwell notifications (i.e., ESM surveys) within 30 minutes of receiving the prompt. Some students communicated with the research team via an additional OneNote page, given to them to report issues with the Expiwell app given this was the first time we used this particular application or ask questions, that 30 minutes to respond to the survey was difficult. To address these concerns, the incentive requirement was adjusted in Weeks 3 and 4: students needed to complete 15 out of 20 notifications and had 45 minutes to respond. Notifications were sent from 9:00 AM to 6:00 PM. Two students continued to experience issues but were eligible for incentives by reporting these problems. Students with limited participation (one to two days) received a partial $20 incentive for the week.

5.4 Our Case Study Approach

In this study, we employ a case study approach as the analytic method to explore how the integration of different types of data sources shape the development of individualized daily journal prompts. This section starts by justifying the use of a case study by drawing on key principles from the case study literature, explaining how the case study approach is appropriately suited to our research questions. The next section then outlines the overall structure and rationale for our specific case study approach (nested and embedded case study). The embedded units allow us to compare how threads identified from the ESM data in Fall 2020 differ from those in Fall 2021. The nested structure allows us to track how the changes in integration in the data collection across the two iterations impacted what we observed from the data and how these were used to inform the writing of the individualized daily journal prompts.

5.4.1 Justifying a Case Study Approach as an Analytic Method

We chose to use a case study approach as it offers a strong fit for our focus on examining a bounded phenomenon—the evolution of the data collection process—across two iterations (Fall 2020 and Fall 2021) of employing a mixed methods approach. This approach supports our interest in tracing how integration practices shaped the construction of our qualitative data collection tool. This decision is supported by Yin (REF) and Stake (REF), who both emphasize the value of the case study approach when the goal is to understand complex processes situated within real-world

contexts.

Stake (REF) recognized the complexity involved in selecting a case study approach by offering three broad categorizations: intrinsic, instrumental, and collective case studies. He emphasized that many case studies do not fit neatly into a single category and proposed these types not as rigid classifications, but as a way to help researchers consider the purpose behind choosing a case study. This framing highlights that selecting a case study approach is as much about clarifying the researcher's intent as it is about the design itself. Together, our rationale and Stake's typology position the case study as both a flexible and purposeful approach to studying dynamic, real-world phenomena.

Similarly, Yin (REF) argued that researchers should explicitly justify their use of a case study approach as part of designing a strong and credible study. For Yin, articulating the rationale for selecting a single or multiple case design is essential to the overall coherence and rigor of the research. His work reinforces the idea that case study approaches are particularly valuable when research questions are descriptive or explanatory in nature, or when the goal is to 'illuminate a particular situation'.

Stake (1978) emphasized that case studies are particularly valuable when the goal is to come to know the particular in a way that allows for recognition across other contexts. Rather than producing generalizable findings in a traditional sense, case studies generate naturalistic generalizations—insights that readers can relate to their own contexts through personal resonance and contextual similarities. In this way, case studies invite readers to make meaning from the situated knowledge they present.

In our work, the individualized daily journal prompts were written based on the kinds of observations that could be drawn from the ESM survey responses. These observations were constrained by what was visible in the data at each point in time. However, by examining what kinds of information were used to craft these prompts—and how that information varied between iterations—we surface the decisions and interpretations embedded in the data collection process itself. This is particularly important in studying students' social-emotional experiences, as it allows

Phase 2: Data Observations

* Provides the task being done during the moment in relation to a specific course

~ Examines one's task-specific self-efficacy towards a task situated within a course

Open-Ended Course Question Response

Situates the reported moment within a course

Open-Ended Task Question Response

Provides the task being done during the moment

#Contextualizes the survey responses with course information

@Contextualizes the survey responses with task information

Close-Ended Task-Specific Self-Efficacy Scores

Examine for increases or decreases in task-specific self-efficacy scores

Exploring the Additional Point of Integration - merging the qualitative and quantitative ESM Data

Figure 5.3 Visualizing Integration Across ESM Data Sources in Fall 2021. This Venn diagram illustrates the additional point of integration introduced in Fall 2021 by establishing intentional communication between three Experience Sampling Method (ESM) data sources: (1) open-ended task survey responses, (2) closed-ended course survey responses, and (3) closed-ended task-specific self-efficacy (SE) scores. Each data source is represented as a circle, and overlaps between them indicate instances where these data strands were brought into communication during data collection. Stars within the overlaps represent specific integration points, with each symbol (e.g., #, @,) denoting a distinct configuration of data sources in dialogue. The green star, with squiggly in the middle, at the center marks the fullest integration, where all three sources communicate—allowing us to explore how task-specific SE is situated within a particular course context and further elaborated through student reflections. This integration point was not present in Fall 2020, where only the closed-ended task-specific SE scores were considered in isolation. The figure thus highlights the methodological shift made in Fall 2021 to embed integration as both a design and analytic principle.

us to interrogate how methodological shifts influence what gets seen, captured, and ultimately interpreted.

ADD TEXT ABOUT FIG 5.3 HERE!

The comparative nature of our study, tracing these shifts between Fall 2020 and Fall 2021, offers insight into how evolving practices of integration contribute to the ways SE is studied in relation to students' everyday lives.

5.4.2 Case Selection and Boundaries

The case selection for this study was informed by the work of Creswell (REF) and Merriam and Tisdell (REF), who each describe principles for defining and selecting cases in qualitative research. Creswell emphasizes the importance of establishing a clear rationale for case selection, while Merriam and Tisdell highlight the necessity of bounding a case—by time, context, or activity—for it to be analytically meaningful. These guidelines were critical in shaping the structure and focus of our study.

In alignment with our research questions, the case for this study centers on the evolving data collection process. Specifically, we examine how integration was introduced and operationalized during two sequential iterations of data collection: Fall 2020 and Fall 2021.

- In the Fall 2020 iteration, only closed-ended ESM survey items were used to inform the development of individualized daily journal prompts.

- In the Fall 2021 iteration, both open- and closed-ended qualitative ESM survey items were utilized to guide the creation of these prompts.

These two iterations are nested within the overarching case which enables a temporal comparison of how methodological decisions about integration unfolded across time, under consistent research goals and within the same broader project.

While the nested design allows us to examine how the overall process shifted between iterations, this study also employs an embedded case study approach to analyze more granular units within each iteration. Specifically, within each iteration, we focus on the ESM data and the individualized daily journal prompts as the embedded units of analysis. These embedded units allow us to investigate how different sources and configurations of ESM data (also known as "threads") directly influence the content and construction of the qualitative prompts.

By analyzing how observational and interpretive threads were identified from the ESM data and, in turn, how these threads informed the design of the individualized daily journal prompts, we are able to assess the methodological impact of data integration practices. This is critical

because changes in how ESM data is analyzed during data collection have downstream effects: they influence the observations identified, which in turn shape the journal prompts. Since these prompts generate the qualitative data central to our mixed methods approach, any changes in their formation directly impact the types of claims we can make—especially those related to students' experiences and the interpretive validity of the study.

5.4.3 Embedded Units of Analysis

The study is bounded by the selection of the embedded units—namely, the ESM data and the individualized daily journal prompts. The selection process involved two interrelated stages: (1) selecting the individualized daily journal prompts that were informed by the ESM data, and (2) identifying the corresponding ESM data. I began with the selection of the daily journal prompts, both because of the involved nature of determining which prompts to highlight from the Fall 2021 iteration, and because I was methodologically interested in how the prompts were crafted—particularly as this process reflects the integration of open- and closed-ended ESM data, which is central to this study.

While the generation of ESM data is relatively transparent, a core contribution of this mixed methods approach lies in how we intentionally designed individualized daily journal prompts by drawing from these data. In Fall 2020, prompts were constructed using the closed-ended ESM items alone, whereas in Fall 2021, we introduced an additional point of integration by linking students' closed-ended task-specific SE scores with their open-ended survey responses, shown in Figure 5.3. This integrated interpretation then informed the writing of the individualized prompts. As such, the prompt selection process reflects both an analytical and design-based integration of methods. The following sections detail the selection process for each stage.

5.4.3.1 The Direct Connection to the ESM Data Source

I began by categorizing the daily journal prompts that were informed by the ESM data. My participation in the Fall 2021 data collection process provided me with an in-depth understanding of how the prompts were generated and therefore the familiarity with the specific language and phrasing used when ESM data was used to inform a prompt. This language often included phrases

that explicitly linked the survey data to the individualized daily journal prompts. Below are examples from the Fall 2020 and Fall 2021 data collection, respectively.

Fall 2020

Today in the survey data you had a few "very confident" responses and then a few "not confident" responses - can you tell us a little bit about what was different from those early responses to the later ones? - Fall 2020 Daily Journal Prompt

The prompt above includes language that clearly indicates it was shaped by ESM data. For example, it begins with the phrase "Today in the survey data...," which explicitly signals the prompt's basis in the participant's daily survey responses. In addition to this general reference to the survey data, it also incorporates specific content from the participants' ESM responses. For example, it notes fluctuations between "very confident" and "not confident" responses, suggesting variability in SE scores throughout the day.

Fall 2021

On Friday, we noticed that you spent most of the day doing [computational modeling course label] homework. We noticed an upward trend in how skilled and successful you felt when you were working on that homework.

- *Can you take us back to Friday and discuss what was happening while you were working on your [computational modeling course label] homework?*

- *How confident were you in doing this [computational modeling course label] homework?*

- *How did that homework impact your confidence in your ability to succeed in your [computational modeling course label] course?*

- Fall 2021 Daily Journal Prompt

This Fall 2021 daily journal prompt provides clear evidence that it was informed by participants' ESM data. First, the opening sentence—"On Friday, we noticed that you spent most of the day doing [computational modeling course label] homework"—indicates that the prompt was generated based on data about the participant's activity patterns across ESM check-ins that day. This suggests that survey responses likely included repeated mentions of working on a particular course's homework, which were then used to construct the prompt.

Second, the statement "We noticed an upward trend in how skilled and successful you felt when you were working on that homework" draws directly on ESM data, reflecting the task-specific SE items collected throughout the day. The reference to an "upward trend" implies that the prompt was not only using a single data point, but rather synthesizing across multiple ESM responses to detect a pattern.

Finally, the follow-up questions are clearly shaped by that data-driven observation. They invite the participant to elaborate on their experiences during the specific activity that was flagged in the ESM data (doing homework), and ask them to reflect on their confidence both in that moment and in the broader context of the course. Together, these elements show how the prompt was tightly linked to ESM data, both in terms of timing ("On Friday"), content (working on a specific course's homework), and the domain-specific measures on the ESM survey (a perceived increase in skill and success over time).

5.4.3.2 Selecting the Embedded Units

Fall 2020: Once the individualized daily journal prompts were categorized, the next step was to select which ones to use for this case study. After reviewing the Fall 2020 data, it was determined that only three of the daily journal prompts were informed by the ESM survey data. These three prompts stood out due to their direct references to participants' survey responses. These prompts were selected for inclusion in the case study because they offered the clearest evidence of data-informed journaling and provided insight into how real-time ESM data was integrated into daily journal prompts. Because there were only three ESM-informed prompts, all were included in the embedded case study. This approach allowed for a more comprehensive exploration of whether

these prompts were examining different threads within the Fall 2020 data. By including all three, the case study could more fully examine how participants' SE experiences were explored, given that only the closed-ended ESM responses was used to inform the writing of the individualized daily journal prompts. This ensured a focused analysis of how the data collection process itself shaped the prompts and the participants' reflections.

> Example 1: *Today in the survey data you had a few "very confident" responses and then a few "not confident" responses - can you tell us a little bit about what was different from those early responses to the later ones? - Fall 2020 Daily Journal Prompt*

> Example 2: *Today in the survey data it appears that you had a rough patch in your day where you felt less in control and skilled than at other moments - can you recall that moment today and tell us about what was happening? - Fall 2020 Daily Journal Prompt*

> Example 3: *By your last notification today in the app it was clear that something had changed in your day. Can you tell us anything about what affected that change? - Fall 2020 Daily Journal Prompt*

Fall 2021: While the Fall 2020 prompts were limited to those directly informed by fluctuations in task-specific SE scores, the Fall 2021 dataset offered a wider range of prompts due to the expanded use of the open-ended ESM data. As such, when selecting Fall 2021 prompts for comparison, our objective was to capture this range of threads that were distinct from one another but still rooted in the integration of the open- and closed-ended ESM survey data. By selecting prompts that highlight a distinct thread from the ESM survey data to be followed up on in the daily prompt, we can better understand how the changes in the data observation process, which drew on different sources of ESM data during the data collections of Fall 2020 and Fall 2021, impacted the nature and focus of the journal prompts.

The first selected the individualized daily journal prompt is provided below. This individualized daily journal prompt was based on the integration of the open- and closed-ended ESM survey data

(Green Star in Figure 5.3). This prompt not only addressed fluctuations in SE but also probed into the student's experience, making it a more representative example of how prompts can capture in-the-moment changes in SE. We call this thread – Singular Task Across a Singular Day.

> **Daily Journal Prompt Embedded Unit #1:** *On Friday, we noticed that you spent most of the day doing [computational modeling course label] homework. We noticed an upward trend in how skilled and successful you felt when you were working on that homework.*
>
> - *Can you take us back to Friday and discuss what was happening while you were working on your [computational modeling course label] homework?*
>
> - *How confident were you in doing this [computational modeling course label] homework?*
>
> - *How did that homework impact your confidence in your ability to succeed in your [computational modeling course label] course?*

The second individualized daily journal prompt is shown below. This prompt was selected because it invites the student to reflect on their SE across multiple academic tasks experienced throughout the day, all within the same course. Unlike prompts that explicitly ask students to compare different tasks within the same course, this prompt provides a more open frame — giving students the agency to identify distinctions between tasks and to articulate how those distinctions shaped their sense of capability. In this way, the prompt allows students to surface and make meaning of intra-course variation in SE on their own terms.

> **Daily Journal Prompt Embedded Unit #2:** *Yesterday it stood out to us that your confidence in your abilities in [honors undergraduate research seminar course label] was notably higher. Could you tell us what was happening yesterday? What were you doing? Who were you working with? What impacted your sense of confidence?]*

Lastly, the final thread from Fall 2021 involved fluctuations in students' SE across multiple courses (e.g. Fluctuations Across Multiple Courses). For this thread, the task-specific SE scores fluctuated not within a single course or activity, but across different courses or tasks (e.g., in class for one course, looking over notes for another course). This is where the task open-ended ESM survey items course open-ended ESM survey items came into play together, allowing us to probe not only the task-specific SE but also the broader context of the students' experience across different courses.

While the task open-ended ESM item was not necessarily focused on a specific activity repeated over time (as in some of the other cases), the fluctuation in SE across multiple courses still reflected a moment-by-moment shift in the student's SE. In one particular unit, we chose to develop an individualized daily journal prompt that considered how these multiple task SE fluctuations across two distinct courses might influence a student's overall academic experience. This represents an important nuanced exploration of SE, using the integrated data sources to understand how different tasks across various courses shape a student's SE. This individualized daily journal prompt was the only prompt with this thread, making it a natural choice for demonstrating how the mixed methods approach was applied in practice. The prompt is shared below.

Daily Journal Prompt Embedded Unit #3: *Over the last 3 days, your responses in the app indicated you were working on both [Cell and Molecular Biology Course Label] and [General Chemistry Course Label]. However, you indicated you were less skilled and felt less successful on [Cell and Molecular Biology Course Label] than [General Chemistry Course Label.] Can you tell us a little bit more about why you feel differently in these two classes?*

Thus, for the Fall 2021 iteration, we identified the embedded units based on whether the task and course open-ended ESM survey items were integrated with the closed-ended ESM survey items; specifically, the task-specific SE items. Once we established this level of integration occurred, we then turned to selecting individualized daily journal prompts based on the thematic threads

identified from these integrated data sources. This process for selecting the embedded units from the Fall 2021 iteration allowed us to bring in the following threads: Singular Task Across a Single Day, Multiple Tasks within the Same Course, Fluctuations Across Multiple Courses. By ensuring that the prompts addressed distinct thematic threads, we have set up a nested and embedded case study to explore the threads identified from the ESM data, and in turn, examine how these threads were used to shape the nature of the individualized daily journal prompts.

5.4.4   Comparing Embedded Units

In the above sections, I have identified the individualized daily journal prompts from both iterations that I selected for the embedded part of my case study approach. In this section, I will describe how these prompts were paired in the two iterations to establish the embedded case analysis. The prompts selected for comparison were not simply matched based on similar surface topics (e.g., confidence or task engagement), but rather on their shared methodological function. Specifically, I prioritized prompts that (1) were grounded in similar ESM observations, (2) invited opportunity to discuss those observations, and (3) provided opportunities to reflect about how the integration of different data sources informed the threads that were built.

For example, I intentionally paired one Fall 2020 prompt with one of the Fall 2021 prompts as demonstrated below.

Fall 2020 Individualized Daily Journal Prompt: *Today in the survey data it appears that you had a rough patch in your day where you felt less in control and skilled than at other moments - can you recall that moment today and tell us about what was happening?*

Fall 2021 Individualized Daily Journal Prompt: *Yesterday it stood out to us that your confidence in your abilities in [honors undergraduate research seminar course label] was notably higher. Could you tell us what was happening yesterday? What were you doing? Who were you working with? What impacted your sense of confidence?*

These two prompts were paired because both explored a specific moment in the student's day that reflected a meaningful shift in task-specific SE—either an increase or a decrease. Both prompts invited the participant to narrate what was happening in that moment, providing context for how they interpreted the change. The Fall 2021 prompt, however, incorporated more scaffolding by ending with a specific question—"What impacted your sense of confidence?"—demonstrating the increased intentionality in prompt construction based on integrated ESM data. In this way, the pairing reflects how the same type of observational insight could lead to more refined and targeted follow-up as the methodology evolved.

By structuring the comparisons in this way, the analysis makes visible how the threading process became more intentional and how that intentionality shaped the individualized nature of the prompts—allowing for a more nuanced exploration of students' SE experiences. Table 5.3 shows the embedded units that will be discussed in the following results section.

Identifier	Fall 2020 Daily Journal Prompt	Fall 2021 Daily Journal Prompt

Embedded Unit 1	Today in the survey data you had a few "very confident" responses and then a few "not confident" responses - can you tell us a little bit about what was different from those early responses to the later ones? –Jane	On Friday, we noticed that you spent most of the day doing [computational modeling course label] homework. We noticed an upward trend in how skilled and successful you felt when you were working on that homework. Can you take us back to Friday and discuss what was happening while you were working on your [computational modeling course label] homework? How confident were you in doing this [computational modeling course label] homework? How did that homework impact your confidence in your ability to succeed in your [computational modeling course label] course? –Robert

Embedded Unit 2	Today in the survey data it appears that you had a rough patch in your day where you felt less in control and skilled than at other moments - can you recall that moment today and tell us about what was happening? –James	Yesterday it stood out to us that your confidence in your abilities in [honors undergraduate research seminar course label] was notably higher. Could you tell us what was happening yesterday? What were you doing? Who were you working with? What impacted your sense of confidence? –Ted Brown
Embedded Unit 3	By your last notification today in the app it was clear that something had changed in your day. Can you tell us anything about what affected that change? –Lenny	Over the last 3 days, your responses in the app indicated you were working on both [Cell and Molecular Biology Course Label] and [General Chemistry Course Label]. However, you indicated you were less skilled and felt less successful on [Cell and Molecular Biology Course Label] than [General Chemistry Course Label.] Can you tell us a little bit more about why you feel differently in these two classes? –Rose

Table 5.3 This table presents the embedded units selected for analysis, each consisting of one individualized daily journal prompt. For example, Fall 2020 Embedded Unit 1 includes the prompt: "Today in the survey data you had a few..." This comparison pairs Embedded Unit 1 from Fall 2020 with Embedded Unit 1 from Fall 2021.

5.4.4.1 Relating the Framing of Methods and Results: Embedded Units and Student-Centered Narratives

While the methodological comparisons in this study are organized around embedded units to examine the evolution of the data collection process, the results themselves remain grounded in the lived experiences of the student participants. Each embedded unit reflects not just a methodological shift, but a moment in a student's journey. As such, in the results section, I refer to students by pseudonym where appropriate to preserve the narrative continuity of the data and to foreground the student voices at the heart of this study.

This narrative decision is also embedded in the logic of the study's design. The ESM data and individualized daily journal prompts are inherently tied to each student, with their responses forming the basis for identifying threads and guiding data integration. Because both forms of data originate from the same participant—who, in effect, serves as the point of integration between the ESM surveys and the individualized daily journal prompts—and because the data are intentionally linked in both collection and analysis, it is both methodologically sound and ethically appropriate to maintain the student connection in how the data is presented. Thus, the analysis itself is intentionally narrated with respect to the students whose data shaped those prompts. This approach provides both methodological clarity and an ethically grounded representation of the voice of the participants.

5.5 Threads of Inquiry Influence the Writing of the Daily Journal Prompts

5.5.1 Examining the ESM Survey Data: Following Threads

In this section, we describe the process of identifying key threads from ESM data, which can be used to inform the writing of individualized daily journal prompts. We will explore how the types of threads we can follow differ when analyzing only the closed-ended ESM data versus when both open- and closed-ended ESM data are available. Specifically, we will compare three sets of embedded units: (1) Embedded Unit 1 from Fall 2020 (Jane's ESM Data) compared to Embedded Unit 1 from Fall 2021 (Robert's ESM Data), (2) Embedded Unit 2 from Fall 2020 (James's ESM Data) Compared to Embedded Unit 2 from Fall 2021 (Ted Brown's ESM Data), and (3) Embedded Unit 3 from Fall 2020 (Lenny's ESM Data) compared to Embedded Unit 3 from Fall 2021 (Rose's

ESM Data). This comparison will highlight how the range of threads available for investigation expands when we incorporate both data types from the ESM survey. From this analysis, I will show the following:

1. When we examine the closed-ended ESM data, we are limited in the kind of threads we can follow from the ESM survey data. We can only explore changes in the task-specific SE scores (e.g., fill in with example here).

2. When we examine the open- and closed-ended ESM data, we can explore a broader range of threads to be able to inform the writing of the daily journal prompts. Thus, the integration of the open- and closed-ended ESM data sources enables a deeper and more nuanced exploration of the students' SE and the contexts influencing it.

In turn, we can use this information to craft more targeted and relevant daily journal prompts to guide students in reflecting on the factors influencing their SE.

5.5.1.1 Embedded Unit 1 Comparison – The "Singular Task Across a Single Day" Thread

Jane from Fall 2020: On Day 1 of Week 2 data collection, Jane completed the survey four times; her task-specific SE responses are plotted in Figure 5.4. In examining Jane's Task-Specific SE Scores from Week 2, the research team observed her data shift from being in the 80s (more confident in Notifications 1 and 2) to the 20s (less confident in Notifications 3 and 4). Thus, the overall thread the team followed up on using a daily journal prompt was this decrease in Jane's task-specific SE scores across two moments throughout one day of data collection.

Robert from Fall 2021: In Fall 2021, we observed there an increase in Robert's task-specific SE across Day 5 of Week 3 (shown in Figure 5.5). To investigate this variation in Robert's SE in Fall 2021, we also examined his open-ended survey response to pinpoint if we wanted to investigate a specific moment he reported or if we wanted to investigate the moments as a collective.

Table 5.4 shows Robert's responses to the open-ended question, "What is the main thing you are doing?". In examining Robert's responses to these questions, we observed he was working on homework across the day. As such, using the responses to this open-ended survey question
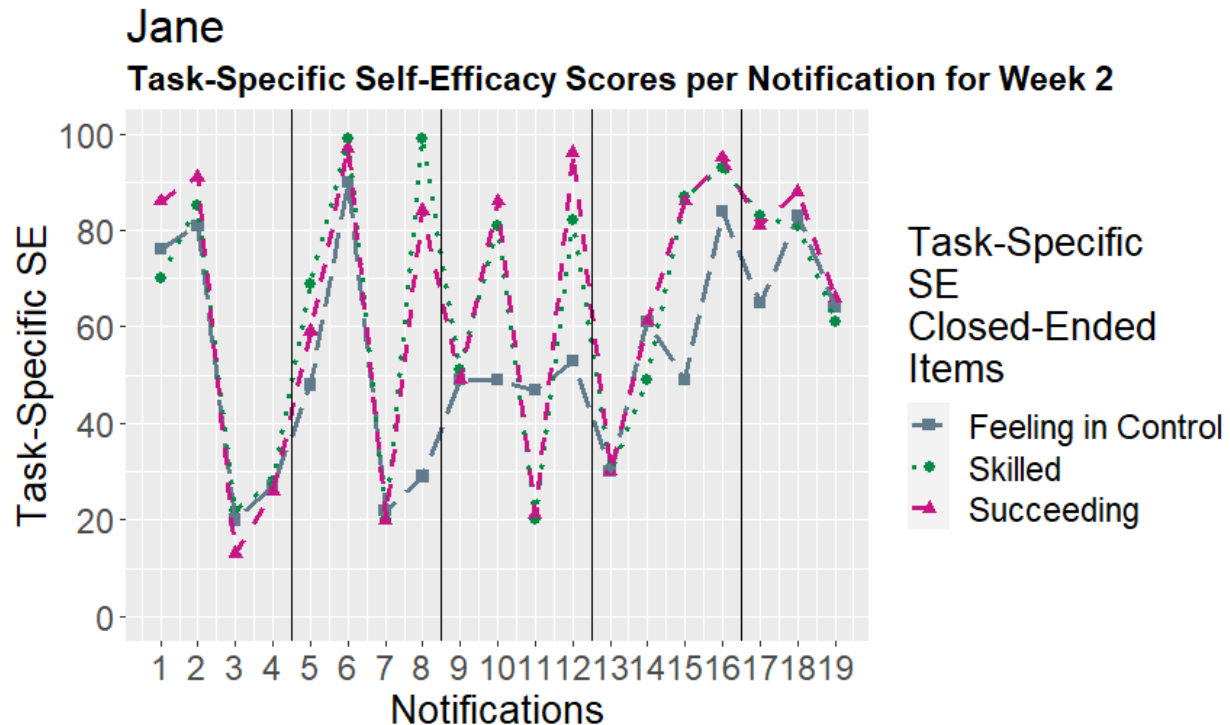
Figure 5.4 Jane's Fall 2020 Week 2 Task-specific SE

supported us in being able to contextualize his fluctuation in his task-specific SE scores with a task (i.e., homework).

While this provided some context to the fluctuation in his task-specific SE scores across the day, there were many questions raised. Was Robert working on homework within a particular course or homework across multiple courses? This sort of information changes the ways we interpret Robert's SE. For example, if Robert was working on one homework from a course, then a variation in his task-specific SE scores would tell us that something happened across the day for Robert's task-specific SE to increase or decrease towards completing that homework for that course. If Robert was working on multiple homeworks within a course, then this may tell us that Robert may be more or less confident in his homework depending on the content. We could imagine this may be the case if one homework is going over content he has learned before and thus is calling on mastery experiences to become more confident in one homework versus the other.

To account for this, we turned to use the course open-ended question - "What is the name of the course?". Robert's responses to this question are shown in Table 5.4. In this case, Robert's
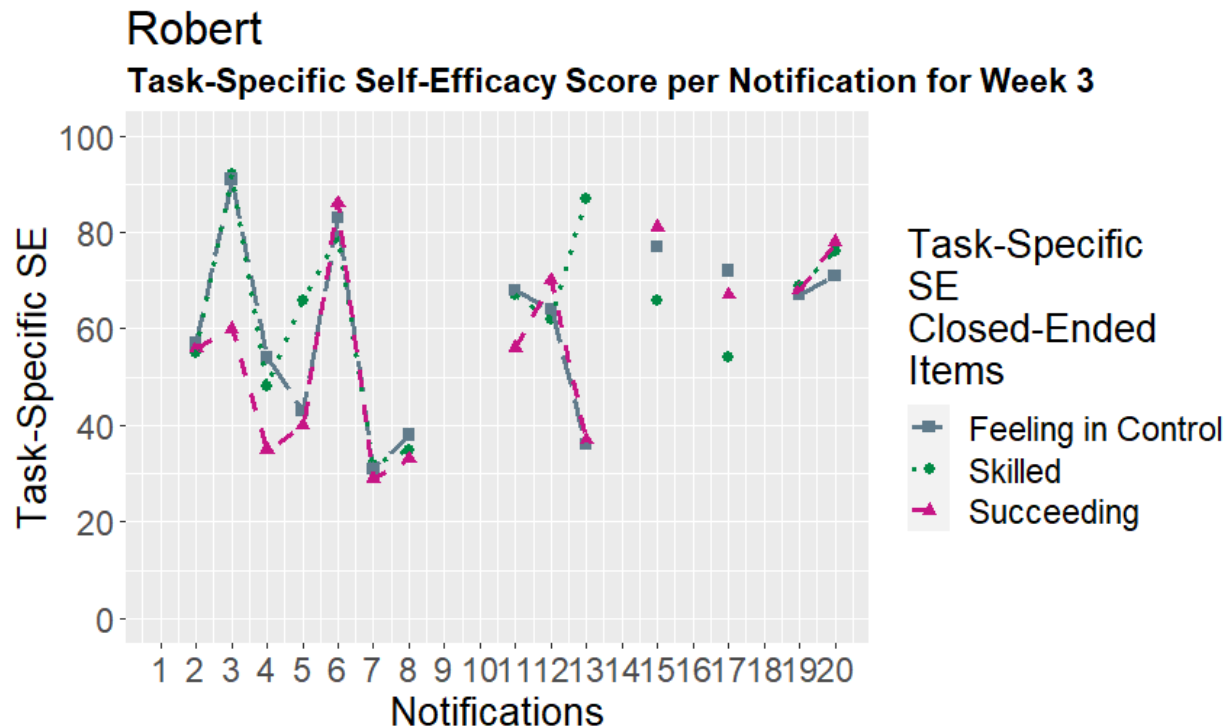
Figure 5.5 Robert's Fall 2021 Week 3 Task-specific SE

homework was associated with his computational modeling course across all three notifications on Day 5. The course open-ended question proved useful to associate the notification or moment he was reporting on with a specific course. As a result of using the course open-ended question in addition to the task open-ended question, we knew Robert was working on homework for his computational modeling course across the day. Next, we could pair these pieces of information with the inference we made from examining Robert's task-specific SE plot.

Meta-inferences: Robert's SE appears to be increasing as he works on his computational modeling

Notification	Reported Task	Reported Course
17	working on a [computational modeling course label] homework assignment	[computational modeling course label]
19	Homework	[computational modeling course label]
20	Homework	[computational modeling course label]

Table 5.4 Robert's reported tasks and courses associated with Notifications 17, 19, and 20. Notification 18 is not shown here as he did not complete a survey.

homework across the day. The meta-inference ties together multiple pieces of data – his task-specific SE scores were increasing across the day, he reported working on homework (task), and he reported this activity was for his computational modeling course (course) – to build a thread to follow up within his daily journal prompt. The inference we drew from Jane's data was, "A decrease in task-specific SE from the first to the third and fourth notification," only discusses her closed-ended ESM survey data.

The meta-inference from Jane's data does not hold any information in relation to the moment(s), limiting the kind of information we can follow up on with the daily journal prompts. Robert's case, however, illustrates the added depth and context that qualitative data brings to the analysis. Thus, when only examining the closed-ended ESM data, we can only examine fluctuations in one's task-specific SE scores, but when we integrate the open- and closed-ended ESM survey data, we can contextualize the variation in one's task-specific SE scores in relation to a specific moment or a series of moments from the students' day.

5.5.1.2 Embedded Unit 2 Comparison – The "Multiple Tasks within the Same Course" Thread

James from Fall 2020: On Day 1 of the Week 2 data collection, James completed the survey four times; his task-specific SE responses are plotted in Figure 5.6. In examining James's Task-Specific SE Scores from Week 2, the research team observed his data decrease across the day from Notification 1 to Notification 3. Thus, the overall thread the team followed up on using a daily journal prompt was this decrease in James's task-specific SE scores across two moments throughout one day of data collection.

Ted Brown from Fall 2021: In Fall 2021, there were many threads from analyzing Ted Brown's ESM survey data as shown in Figure 5.7; however, we focused on the fluctuation in task-specific SE scores from Notification 3 to Notification 4. To contextualize this fluctuation, we turned to examining Ted Brown's qualitative ESM survey data.

As shown in Table 5.5, Ted Brown reported working on research projects during each notification. To explore whether these projects were the same or related, we analyzed his course open-ended

James

Task-Specific Self-Efficacy Scores per Notification for Week 2
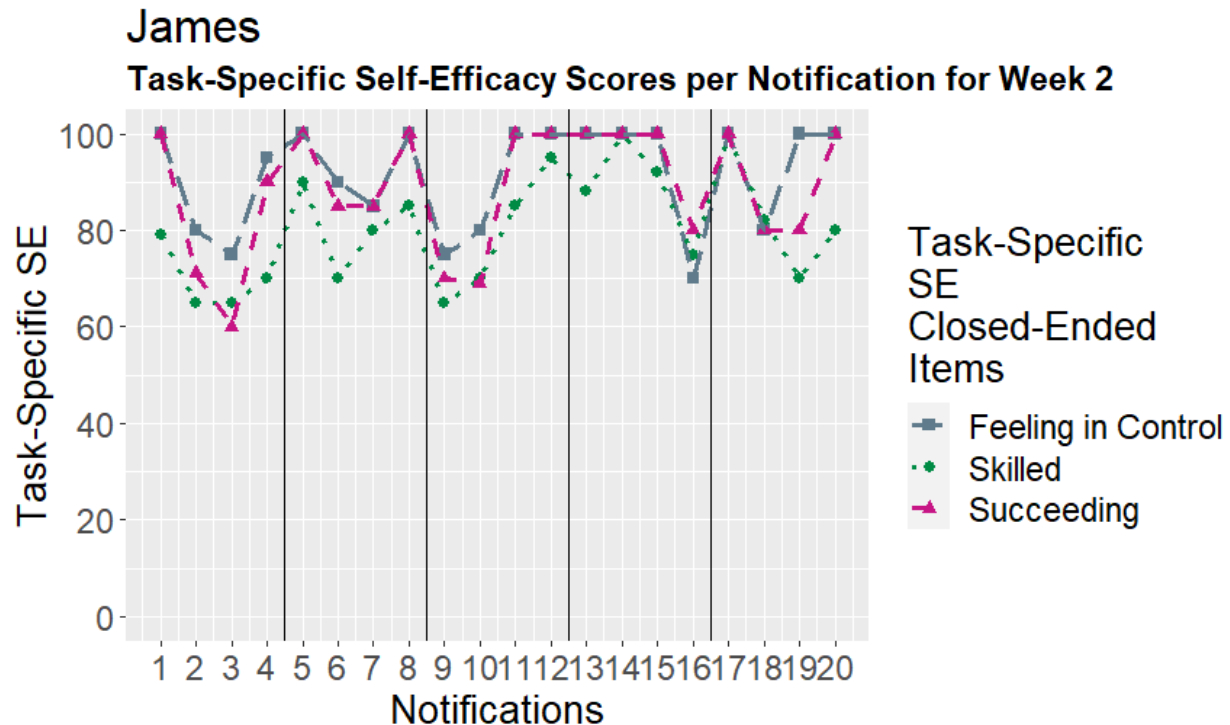
Figure 5.6 James Fall 2020 Week 2 Task-specific SE

survey responses. Notification 3 and 4 on Day 1 were associated with his [research course]. Thus, using the open-ended survey responses allowed us to contextualize Ted Brown's task-specific SE scores with information about the task (research projects) and the course (research course).

Meta-inferences: Ted Brown's SE appears to be increasing as he is working on research project(s) for his [research course]. The meta-inference ties together multiple pieces of data – his task-specific SE scores were increasing across the two notifications, he reported activities related to his research project (task), and he reported these activities were for his [research course] – to build a thread to follow up within his daily journal prompt. The inference we drew from James's data, "A decrease in his task-specific SE scores from Notification 1 to Notification 3," only discusses his closed-ended ESM survey data.

As such, the meta-inference from James's case is less informative for crafting a journal prompt. James's case, examined using only close-ended data, demonstrates the constraints of focusing solely on fluctuations in SE scores without contextualizing these scores, which may drive those changes. In contrast, Ted Brown's data, when examined holistically through the lens of both open- and
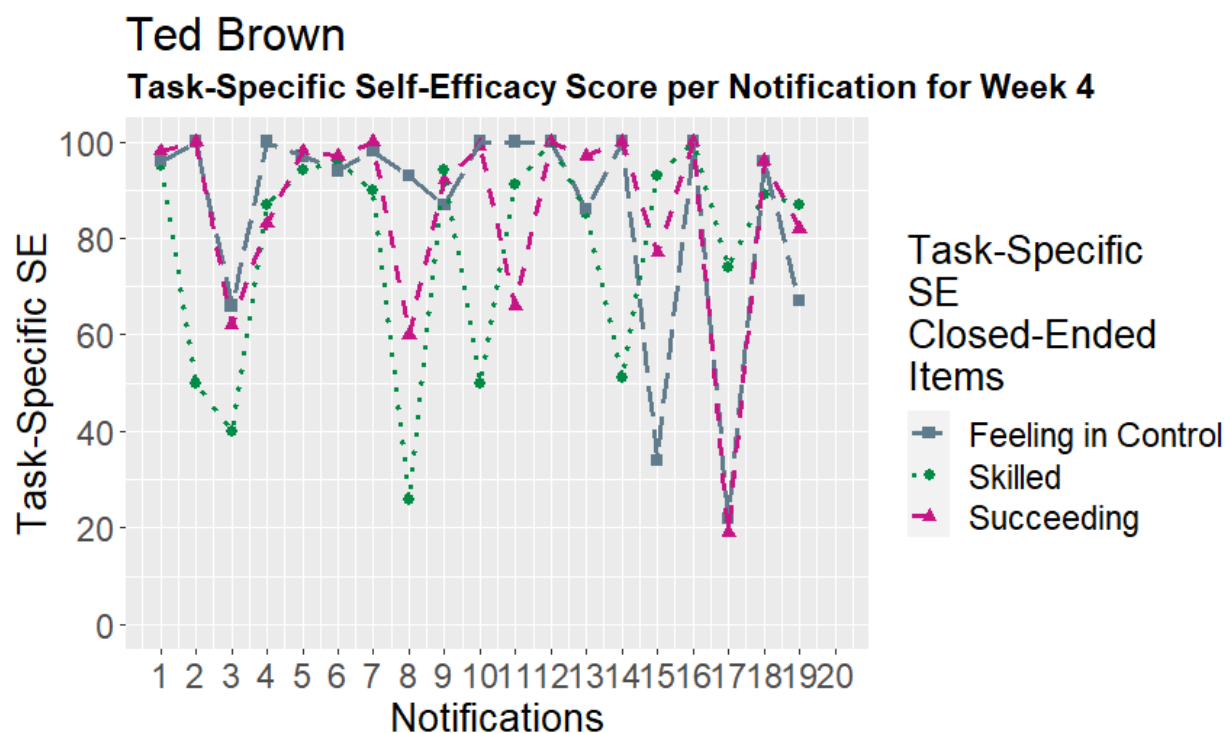
Figure 5.7 Ted Brown Fall 2021 Week 4 Task-specific SE

close-ended ESM data, presented multiple possible threads. This case underscores the value of understanding the multifaceted nature of SE – how fluctuations in SE might differ depending on the activity (e.g., research project, quiz, different parts of a research project) the student engages in and/or the course the activity is associated with. Thus, Ted Brown's case illustrates the added depth and context that qualitative data brings to the analysis of one's SE.

Notification	Reported Task	Reported Course
3	working on my research project	[research course label]
4	putting my findings for this week for my research project into a presentation	[research course label]

Table 5.5 Ted Brown's reported tasks and courses associated with Notification 3 and 4.

5.5.1.3   Embedded Unit 3 Comparison – The "Fluctuations Across Multiple Courses" Thread

Lenny from Fall 2020: On Day 4 of Week 1, Lenny completed the survey multiple times throughout the data collection; his task-specific SE responses are plotted in Figure 5.8. From examining Lenny's Task-Specific SE Scores from Week 1, we observed a dramatic increase between Day 3

(Notifications 9 thru 12) and Day 4 (Notifications 13 thru 16). Thus, the overall thread the team followed up on using a daily journal prompt was this increase in Lenny's task-specific SE scores across two days of data collection.
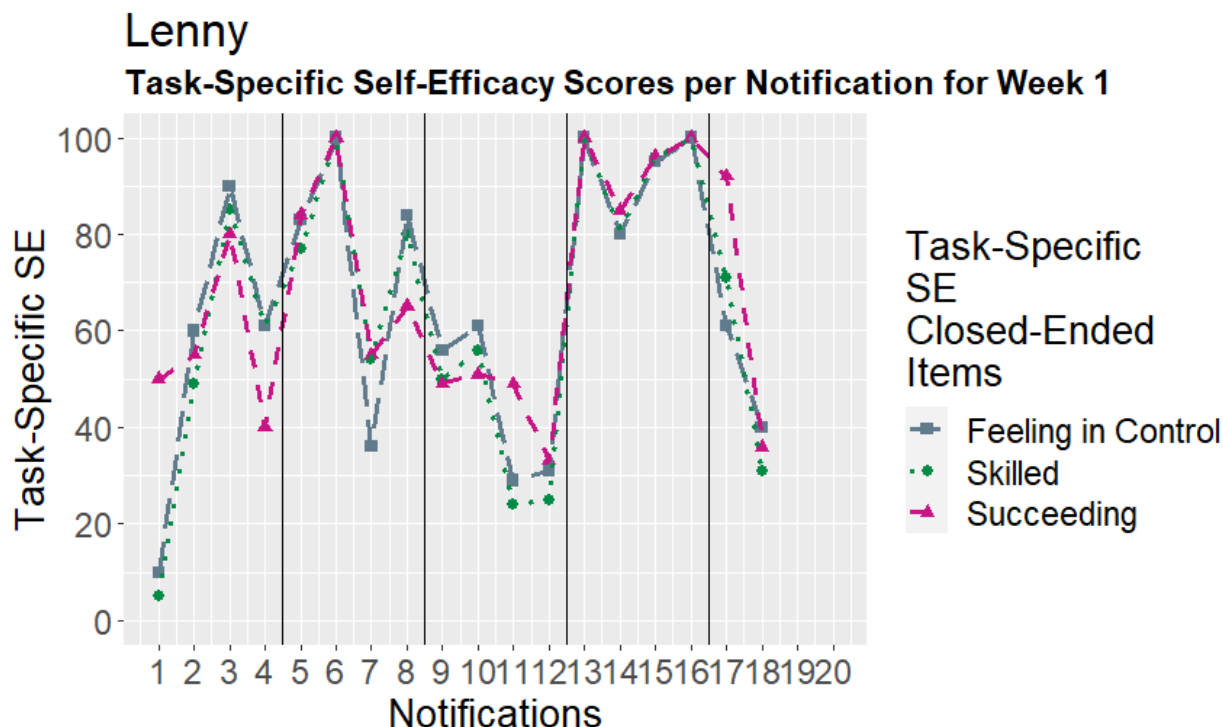
Figure 5.8 Lenny Fall 2020 Week 1 Task-specific SE

Rose from Fall 2021: When examining Rose's Task-Specific SE responses across Days 1, 2, and 3, we observe some fluctuations (shown in Figure 5.9. Specifically, Notification 6 on Day 2 shows a potential decrease in her task-specific SE responses; quite different from the rest of her responses thru Day 3 of data collection (Notifications 1 thru 12).

To further elaborate on this variation, we examined her open-ended survey responses to pinpoint if we wanted to investigate a specific moment or a series of moments she had reported, or if we wanted to investigate the moments as a collective. Table **??** shows Rose's responses to the open-ended question "What is the main thing you are doing?" (some notifications were not included if they were not related to academics). This revealed that she was working on task across various courses: biology, chemistry, and math. Notably, at Notification 6 on Day 2, when the research team observed a potential decrease in her SE, Rose reported working on a task for her biology course.

114
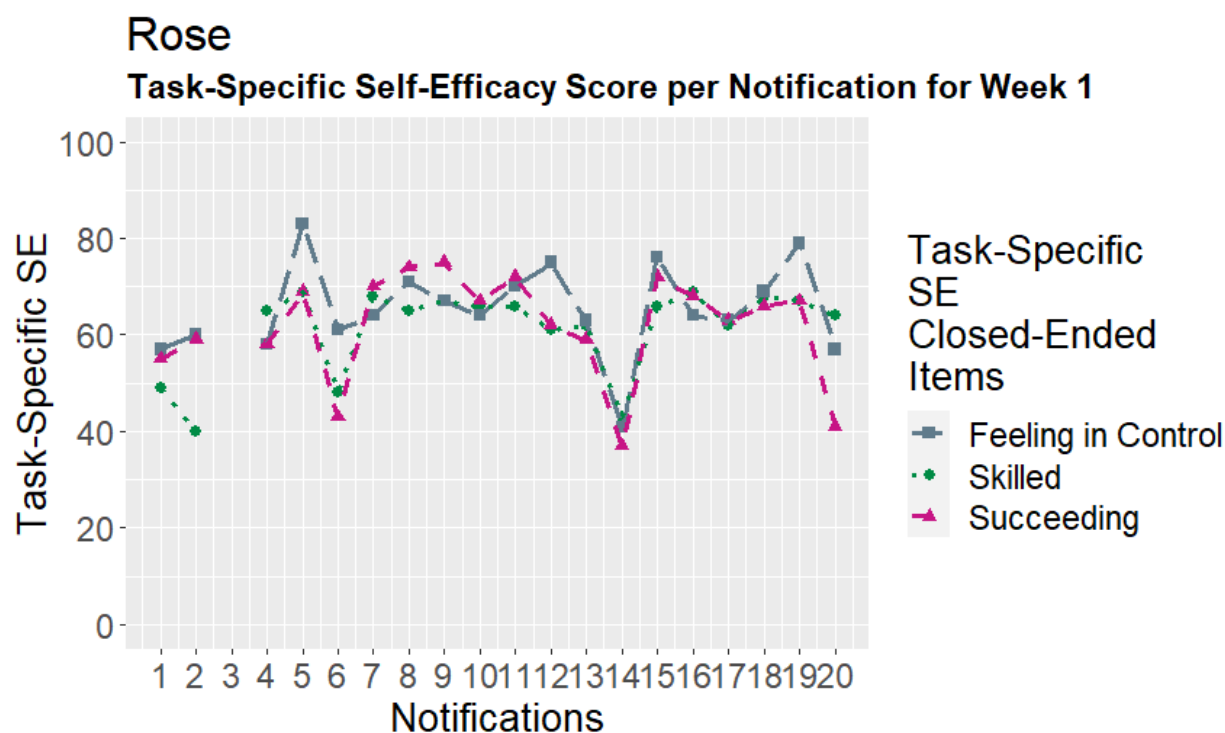
Figure 5.9 Rose Fall 2021 Week 1 Task-specific SE

This contextual information, provided from the task open-ended survey responses, was crucial for interpreting the variations in her task-specific SE scores as this tells us it may not be one particular task but rather we may be observing changes in her task-specific SE scores based on course (i.e., Rose scores tasks in one course differently than tasks in another course). Without this context, the team only know that her task-specific SE fluctuated but would not have information in relation to those task-specific SE fluctuations.

Notification	Reported Task	Reported Course
1	I've mainly been doing school work or working	[biology course label]
6	In class for [biology course label]	[biology course label]
7	Looking over notes for chem recitation	[chemistry course label]
8	Walking to my chem recitation	[chemistry course label]
10	Studying for my math quiz Friday	[math course label]

Table 5.6 Rose's reported tasks and courses associated with each notification for her Week 1, Day 1 through Day 3 of data collection. Only notifications related to academics are shown.

For example, at Notification 1 on Day 1, Rose simply reported, "I've mainly been doing school

115

work or working," which provided little insight into that specific task, demonstrating a potential limitation with the task open-ended survey item. However, when the team cross-referenced this with the course-related open-ended survey responses, it became clear that she was working on tasks for her biology course. Thus, to fill in some context, we still need to consider the course open-ended survey item. To fill in some context, we still need to consider the course open-ended survey item. As such to further contextualize the moments reported on these surveys, the team then examined the survey items asking whether this was course related, and if so, to ask them to report which course the activity was related. Table 5.6 shows Rose's response to this question. The research team observed that for Notification 1 and 6, which may be potential decreases in Rose's task-specific SE, she was working on tasks associated with her biology course. Then, for two of the notifications, Rose was working on tasks associated with her chemistry course; for one of the notifications, Rose was working on a task associated with her math course. Overall. Rose reported across multiple surveys that she was working on her biology and chemistry course across the last three days. Thus, the course question helps us contextualize her responses further by situating their scores and task information from the task open-ended survey item response within a particular course. Next, we can pair these pieces of information with the inference we made from examining Rose's task-specific SE plot.

Meta-inferences: Rose was working on a variety of tasks associated with her biology course and her chemistry course, and that her task-specific SE scores, regardless of the activity, were reportedly lower for tasks associated with her biology course then her chemistry course. The meta-inference ties together multiple pieces of data - her task-specific SE scores, her reported tasks, and her reported courses - to build a thread to follow up within her daily journal prompt. The inference we drew from Lenny's data was "A increase in his task-specific SE scores across two days of data collection" only discusses his closed-ended ESM survey data.

As such, through the lens of Rose's data, when the team considered the open-ended ESM survey data in addition to the close-ended ESM survey data, we were able to identify threads in which the task-specific SE scores were contextualized. In Rose's case, the team was able to examine

for potential shifts in her task-specific SE scores across different courses. Thus, by integrating the open-ended survey data (qualitative responses) and the closed-ended survey data (quantitative responses), this allowed for the thread to move from observing fluctuations in scores (for example, Lenny's case) to describing the fluctuations in scores across different courses (for example, Rose's case).

5.5.1.4 Summarizing Threads of Inquiry for Daily Journal Prompts

Through the cross-case analysis presented above, we compared three sets of embedded units. Across the Embedded Units 1-3 from the Fall 2020 data collection, the kind of threads we could follow up on from the ESM survey data was limited because of only using the closed-ended ESM data. Specifically, we could only explore fluctuations in the task-specific SE scores (e.g., A decrease in Jane's task-specific SE from the first to the third and fourth notification). However, the Fall 2021 cases demonstrate how integrating both open- and closed-ended ESM survey data allows us to contextualize fluctuations in SE by linking them to specific tasks, courses, and/or situations. This integration illustrates the added depth and context that qualitative ESM data brings to the analysis, emphasizing the value of using a mixed methods approach.

Additionally, each Fall 2021 embedded unit highlighted different kinds of threads that could be identified through this integration, a result of adding the point of integration (i.e., integrating the open- and closed-ended ESM survey data). The varied nature of these threads is summarized below:

1. Embedded Unit 1 from Fall 2021 (Singular Task Across a Single Day): Robert had a singular task (homework assignment), and his fluctuation reflected how his SE fluctuated as he worked through is homework across multiple moments in a singular day. This fluctuation provided insight into how his SE was influenced by the task at different times, which we could explore further with an individualized daily journal prompt.

2. Embedded Unit 2 from Fall 2021 (Multiple Tasks Within the Same Course): Ted Brown had a project, and his fluctuation in SE reflected how his confidence changed across the

activities related to his project within the same course. We can't assume that "working on his research project" and "preparing a presentation" are the same tasks. This case highlights the importance of understanding the tasks within the same course may have different impacts on SE, and we cannot assume that all project-related tasks are equivalent.

3. Embedded Unit 3 from Fall 2021 (Fluctuations Across Multiple Courses Over Time): In Rose's case, the fluctuation in her task-specific SE scores across multiple days were related to her task-specific SE scores fluctuating across different courses, in which we would want to examine these contexts further. Her data illustrated how SE can fluctuate depending on the type of course, and how examining these fluctuations over time across different contexts would require further exploration.

These different threads demonstrate that when we integrate the open- and closed-ended ESM data, we are able to explore a broader range of threads that inform the writing of the daily journal prompts. This broadening happens because integrating the data allows us to link fluctuations in task-specific SE scores (closed-ended ESM data) to particular tasks, courses, or situations (qualitative ESM data). As a result, the integration of these data sources provides a more nuanced understanding of students' SE and the contexts that influence it, enabling the creation of more tailored and relevant daily journal prompts to guide students in reflecting on the factors affecting their SE.

5.5.2 Impact of Data Integration on the Nature of the Daily Journal Prompts

Having identified the threads, built from integrating the open- and closed-ended ESM survey data, we now turn to how these threads inform the creation of individualized daily journal prompts. In the following section, we will demonstrate how the nature of the daily journal prompt changes depending on whether the data used to construct it is purely quantitative (i.e., based on fluctuations in task-specific SE scores) or integrates qualitative insights from students' open-ended ESM survey responses. By examining specific examples of individualized daily journal prompts (Embedded Units 1-3 from Fall 2020 and Fall 2021) created for each student, we will show that the type of thread found from data analysis directly informs the formulation of individualized daily journal

prompts, with more detailed and context-specific prompts developed when both qualitative and quantitative ESM data are integrated.

To illustrate this, we will present both the individualized daily journal prompts and the corresponding student responses for all six of the selected embedded units. Including both the prompts and the responses is important to demonstrating how integration of data types shapes not only the construction of the prompts but also how students engage with them. By analyzing both the prompts and the responses, we can show the impact of more personalized, context-sensitive prompts on student reflection, highlighting how students connect fluctuations in their SE to the specific tasks, events, and experiences they encountered.

We begin by showing the individualized daily journal prompts that were written based on the different threads built. Through these examples, we will demonstrate that when only quantitative ESM data is considered, the focus is on fluctuations in task-specific SE scores. Although these fluctuations provide important information, they are de-contextualized from specific moments and the tasks they refer to. However, when threads based on the integrated quantitative and qualitative ESM survey data are used, the individualized daily journal prompts become more personalized and context sensitive. These prompts explore the specific tasks, courses, and/or events tied to the fluctuations in SE allowing for deeper student reflection.

Thus, when comparing the individualized daily journal prompts across the three sets of embedded units (i.e., 1. Embedded Unit 1 from Fall 2020 (Jane's data) and Embedded Unit 1 from Fall 2021 (Robert's data), 2. Embedded Unit 2 from Fall 2020 (James's data) and Embedded Unit 2 from Fall 2021 (Ted Brown's data), and 3. Embedded Unit 3 from Fall 2020 (Lenny's data) and Embedded Unit 3 from (Rose's data)), we will show that the quantitative-only prompt is more general, asking about changes in SE scores without context. In contrast, the integrated prompt is richer and more specific, asking students to connect their SE scores to the actual experiences that influenced them. This occurs because the qualitative data provided the context that allowed the individualized daily journal prompts to become a tool for guided self-reflection. Rather than simply asking about shifts in scores, the integrated individualized daily journal prompts encourages students to reflect

on the "why" or "how" behind these fluctuations, leveraging the value of an explanatory sequential mixed methods approach.

By highlighting these differences, we will show that the data observation process which was used to build threads, and then these threads influenced the written prompts. Specifically, the integration of qualitative data leads to more personalized and context-rich prompts that guide students in reflecting on their individual experiences. This approach fosters more meaningful daily journal prompts, in turn, identifying specific factors that influence students' SE.

The key takeaway here is that integrating quantitative and qualitative ESM data not only broadens the range of threads that can be found in the ESM data observations but also shapes the nature of individualized daily journal prompts. Using both quantitative and qualitative ESM data sources, we move beyond simple numerical changes in SE to explore how these changes are linked to specific tasks, events, and/or contexts. This shift results in more personalized journal prompts that encourage students to reflect on the experiences that influence their SE, rather than just the fluctuations themselves.

5.5.3   Comparing Embedded Unit 1 from Fall 2020 (Jane's Data) to Fall 2021 (Robert's Data)

Embedded Unit 1 from Fall 2020 (JANE's data): Jane's individualized daily journal prompt was informed by the thread, drawn from her quantitative ESM survey, in which Jane's task-specific SE scores from Notification 1 and Notification 2 to Notification 3 and Notification 4 decreased. As a result, her daily journal prompt, shown below, was written to elicit information about the differences from the earlier to the later responses in relation to her SE.

> *Today in the survey data you had a few "very confident" responses and then a few "not confident" responses - can you tell us a little bit about what was different from those early responses to the later ones?* - **Jane's Fall 2020, Week 2, Day 1 Daily Journal Prompt**

We can see here in the daily journal prompt that the writer (i.e., third author) used language within the daily journal prompt to link the daily journal prompt to the ESM surveys. First, they

started by using "Today in the survey data", which would tie this daily journal prompt to Jane's ESM survey responses from that day. Then to probe into the decrease within Jane's task-specific SE scores from Notification 1 and Notification 2 to Notification 3 and Notification 4, they described her scores saying "you had a few "very confident" responses and then a few "not confident" responses - can you tell us a little bit about what was different from those early responses to the later ones?" This language focuses on the fluctuations within Jane's scores throughout the day.

The advantages of the language used here are that we have clear language that ties the daily journal prompt to the ESM data and we have language that indicates that we noticed a difference in her scores. However, the limitation here is that language solely focuses on Jane's scores, which is a reflection that the thread being followed also is solely based on an observation drawn from Jane's task-specific SE scores. As a result, this prompt does not require Jane to discuss how particular moments in her day are influencing her SE, which was one of the main aims in designing this mixed-method approach. This becomes more evident when considering Jane's response to this daily journal prompt, provided below.

> *My physics 184 has a great help room and instructors and I really feel like I am learning and mastering the class. My physics lab 1919 is really frustrating and there feels like there is no where to get help or find out information or expectations it is just hard to know what to learn or what to study.* - **Jane's Fall 2020, Week 2, Day 1 Daily Journal Prompt Response**

Within Jane's daily journal prompt response, she discussed the differences in her feelings about these two courses and her overall experiences within the courses she was taking at the time. Jane did not reflect about a moment or event from her day (that is, she did not reflect about the moments she reported at any of the Notifications), and there was no language from Jane linking together her daily journal prompt response and her numeric survey responses. The only language that linked the numeric survey responses and the daily journal prompt was the way the team had formulated her daily journal prompt. Thus, while the methodology was able to pick up a snapshot of her feelings

and overall experiences within the courses she was taking that semester, providing SE information about her course level SE across these two courses, SE information about how the specific evidence or event within one of these courses from her day was missing. Although the ESM quantitative survey responses capture the task-specific nature of Jane's SE, the daily journal prompts were not able to further explore those moments reported in the surveys, potentially as a result of not being written with context that could have been provided through the task and course open-ended survey responses.

Embedded Unit 1 from Fall 2021 (ROBERT's Data): Through Robert's daily journal prompt and response, we will demonstrate how using the thread, capitalized on integrating the task-specific SE scores and the task and course open-ended survey response, resulted in writing tailored, context-specific journal prompts and a deeper exploration of his SE. From Section A. Examining the ESM Survey Data: Following Threads, we reintroduce the thread that was built from Robert's integrated quantitative and qualitative ESM data, shown below.

Robert's task-specific SE appears to be increasing as he works on his computational modeling homework across the day. - **Thread Identified By Researchers from Robert's Fall 2021, Week X, Day Y ESM Data**

Using this thread, the research team would have written the following daily journal prompt to further investigate this observation.

On Friday, we noticed that you spent most of the day doing [computational modeling course label] homework. We noticed an upward trend in how skilled and successful you felt when you were working on that homework.

1. *Can you take us back to Friday and discuss what was happening while you were working on your [computational modeling course label] homework?*

2. *How confident were you in doing this [computational modeling course label] homework?*

122

3. *How did that homework impact your confident in your ability to succeed in your [computational modeling course label] course?*

- Robert's Fall 2021, Week X, Day Y Daily Journal Prompt

In examining this daily journal prompt, the researchers drew on the ESM data in several ways as a result of using a thread that combined the quantitative and qualitative ESM survey data. By using the language, "we noticed an upward trend in how skilled and successful you felt when you were working that homework", which addressed the thread, in which we recognized that he reported increases in his task-specific SE scores (i.e., skill and success being two of the items representing task-specific SE) as he was working on his homework. This prompt could be written about the computational modeling homework across the day because we utilized the task and course open-ended ESM survey items to contextualize his task-specific SE scores.

Then, because we are using the task and course open-ended items which support us as researchers in understanding what he is doing in the moment when he completed each of the surveys on Friday, we had the opportunity to write different kinds of questions investigating the role of that homework in impacting Robert's SE. The first question is prompting Robert to share what was happening in the moments he was working on his homework. Then we wrote two different questions prompting Robert to discuss his SE towards the homework, and then to discuss how this particular task impacted his course-level SE for that class. Through questions like these, we can probe into the difference between his SE towards the task versus the course, and how they relate to one another potentially. Thus, by integrating this ESM data, researchers can (1) craft personalized journal prompts that encourage students to reflect on the specific moments and events that shaped their SE (i.e., attending office hours helped him recover his confidence in doing this homework) and (2) journal prompts can be tailored to investigate specific aspects of SE (task-specific, course-level, or the relationship between them) providing more granular insights, which can be seen in his response to the daily journal prompt below.

There was a problem on the homework I was having a hard time finding a solution

*to, I swore I was doing everything right, but no matter what I couldn't get the result I
expected. Turns out I wasn't doing everything right and all I needed was a 2 minute
explanation form the help room to get me back on track and I believe after the help I
was feeling more successful because before that I was unable to move on until I found
the solution to that problem since I hate leaving stuff blank/incomplete.*

*Once I found the solution to the problem I was stuck on I was feeling a lot more confident
in my abilities to do the homework since all the problems were easy in comparison to
the one I had just spent many hours dwelling on.*

*I feel the assignment has made me a little less confident in my ability to succeed as
most of the content of the homework seemed new to me (since learning the content
in the class has been difficulty, I think [computational modeling course label] is best
taught in an online format, but I also understand that for other students the in person
has been easier) and that has made me a little bit afraid of how I'll do on the final. -*
Robert's Fall 2021, Week X, Day Y Daily Journal Prompt Response

Robert was able to share about his experience in working on this homework across the day as
seen in the first part of his response, where talks about having a hard time finding a solution to a
homework problem. As such, Robert discusses the use of officer hours, illustrating how seeking
help can support his SE towards doing the homework, specifically, when a problem is threatening
his SE towards this homework. Without integrating the task and course open-ended items from
the ESM survey, researchers might miss identifying the thread that led to writing this daily journal
prompt. The thread enables the fluctuations captured in the quantitative ESM data to be situated
within the context of specific tasks, courses, or events from the students' days; in turn, leads to
individualized daily journal prompts that prompt students to reflect on specific events that contribute
to these shifts in SE. In this case, Robert's daily journal prompt and response demonstrates how
with the additional point of integration (i.e., using the qualitative ESM survey responses) supports
investigating the detailed sequence of events that shaped his task-specific SE scores across the day –

how his SE towards his homework increased from one moment to the other moment after attending officer hours.

In comparing Jane's daily journal prompt and response and Robert's daily journal prompt and response, this comparison highlights the strengths and limitations of solely using the quantitative ESM survey data versus integrating the quantitative and qualitative ESM survey data to explore SE. In Jane's case, the journal prompt, as a result of the thread, focused solely on the fluctuations in her task-specific SE scores but was not able to bring in language to prompt her to reflect about the specific experiences that might have caused those shifts. As a result, in her response, while we learn valuable insights about her different experiences within her two physics courses and potential sources influencing her SE, this general response does not necessarily support understanding the fluctuations in her task-specific SE that we were interested in understanding.

However, in Robert's case, the journal prompt, as a result of a thread developed from integrating both task-specific SE scores and task and course open-ended survey responses, was able to investigate his task-specific SE scores in relation to the specific moments that he reported on across his day in his surveys. Additionally, as a result of the thread we followed from his ESM survey data, we could write daily journal prompts further diving into his experience on working on his computational modeling course homework. These questions potentially did different work, so his first question was asking for him to share about his experience of working on the homework broadly, his second question was asking for him share his level of task-specific SE towards doing this homework, and his third question was written to understand how a homework impacts his course-level SE. Overall, Robert was encouraged to reflect on the specific moments that contributed to his SE; specifically, exploring his task-specific SE, his course-level SE, and how they relate to one another, providing richer insights into Robert's overall SE and how his SE changes. Then, in Robert's response, he is able to share how a problem was threatening his SE and that office hours supported in recovering from this threat to his SE towards the homework. Robert also shared though that this homework did threaten his SE towards his final exam. Through the prompt and his response, this highlights the complexity of one activity that may impact Robert's SE; more

125

importantly, demonstrates that to investigate the complexities of SE, we need daily journal prompts that are written to investigate the role of activities and events in shaping students' SE. Thus, by comparing these two cases, we see that an integrated analytical approach of linking his quantitative and qualitative ESM data to identify a thread to follow up on within the daily journal prompts can lead to a deeper understanding of SE – beyond simply tracking changes in scores to investigating the context and experiences that shape those changes.

5.5.4 Comparing Embedded Unit 2 from Fall 2020 (James's Data) to Fall 2021 (Ted Brown's Data)

Embedded Unit 2 from Fall 2020 (JAMES's Data): In Sub-Section 5.5.1, we discussed how we analyzed James's task-specific SE scores for fluctuations; hence, identified a thread from this data to inform the writing of the daily journal prompts. Below is the thread drawn from James's ESM quantitative data that we aimed to investigate further through the journal prompt.

James's task-specific SE scores begin decreasing across the day from Notification 1 to Notification 3. Notification 3 may be a potential dip. - **Thread Identified By Researchers from James's Fall 2020, Week 2, Day 1 ESM Data**

As a result of this thread, the research team would have written the following daily journal prompt. The aim of this prompt would have been to elicit a description from James about his experiences that would lead to a decrease in his task-specific SE scores.

Today in the survey data it appears that you had a rough patch in your day where you felt less in control and skilled than at other moments - can you recall that moment today and tell us about what was happening? - **James's Fall 2020, Week 2, Day 1 Daily Journal Prompt**

In the daily journal prompt, the writer used language to tie the daily journal prompt to the ESM survey responses. First, they started by using "Today in the survey data" indicating to James that we were interested in further understanding his survey response from the day. Then, to investigate the decrease in James's task-specific SE scores, language was used to draw attention to the decrease

126

- "had a rough patch in your day where you felt less in control and skilled than at other moments." We asked him to share his experience in that moment by saying "can you recall that moment today and tell us about what was happening?"

The advantage of the language used in this prompt is that it clearly connects the daily journal prompt to the ESM survey data, highlighting the fluctuation in James's SE scores. The prompt indicated the change in his task-specific SE scores, which helps situate the investigation in a specific context. However, the limitation here is that while the prompt draws on the overall ESM data (e.g., "Today in the survey data") and directly references the dip in James's task-specific SE (e.g., "felt less in control and skilled"), it does not specifically address the moment reported in Notification 3. As a result, there is a risk that James may not reflect on the exact moment he reported in the survey at Notification 3. This is because the prompt does not explicitly reference that particular instance, which is important because of the goal of this mixed-methods approach is to investigate how specific moments influences SE. Without this contextualization, the prompt risks being too general, making it more difficult to tie the journal response back to the precise moment in the ESM data we are interested in investigating. Bringing in James's prompt response without other data provides evidence that tying the ESM survey data with the daily journal prompt and response can be challenging without the support of the open-ended survey responses.

> *I was trying to learn how to do the problems in differential equations section 6.1. At the time I kept getting the wrong answer on the first problems of the homework set and I was getting frustrated.* - **James's Fall 2020, Week 2, Day 1 Daily Journal Prompt Response**

In reading James's daily journal prompt response, he reflected on a specific moment from his day: "I was trying to learn how to do the problems in differential equations section 6.1. At the time, I kept getting the wrong answer...". He also uses language - "I kept getting the wrong answer...and I was getting frustrated" - supporting the idea of a low moment in his task-specific SE. However, it is difficult to directly connect this response to his low task-specific SE scores reported at Notification

127

Notification Number	Reported Task
1	making food
2	I'm in [math class] lecture taking notes
3	solving differential equations with a 2x2 matrices
4	solving differential equations with diagonalizable matrices, using real and unreal eigenvalues

Table 5.7 James's reported task associated with each notification for his Week 2, Day 1 data collection.

3 without more context from James. For instance, we cannot be certain that the moment James describes in his journal response is the same moment he reported on at Notification 3. Without this confirmation, we cannot definitely link the response to the specific fluctuation in his SE scores. To establish a stronger connection, we would need to cross-check his journal response and ESM survey response with the open-ended ESM survey responses he provided at Notification 3. This would help confirm the relationship between the two and ensure the trustworthiness of the linkage between the journal prompt and the ESM survey data, but this is being done in a post hoc analysis situation rather than during the data collection to inform the writing of the daily journal prompt itself. To show this, Table 5.7 shows his responses to the open-ended ESM task survey responses on Week 2, Day 1.

In Notification 3, James reported working on "solving differential equations with a 2x2 matrices," which aligns with the moment described in his journal response. This supports the trustworthiness of the connection between the ESM survey data and the daily journal prompt and response. However, had we integrated both the quantitative and qualitative ESM survey responses earlier in our analysis, rather than conducting post hoc analysis, we could have more effectively explored the complexities of James's SE in relation to his differential equations homework. For example, we could have examined what factors were influencing his confidence in solving homework problems and how his SE in homework affected his overall confidence in the course. By introducing the specific activity James reported in Notification 3 into the journal prompt, we could have more directly linked the daily journal prompt to the precise moment in these ESM survey data, allowing a deeper exploration of his experience. This approach could also have opened up additional questions, such

as those related to the differences between various types of math courses or his previous experience with math content.

James's case highlights that while using the ESM quantitative survey data to write the individualized prompts may support investigating a moment more holistically, this does not guarantee that we can integrate the ESM quantitative survey data with the individualized daily journal prompts to holistically explore how moments are shaping students' SE. Further, without providing more details around the moment in which we aim to explore, we can only ask James to provide us with more information about what was happening in the moment, limiting our ability to investigate the complexities of one's SE. For example, this could have resulted in James only focusing on recalling the moment and what was happening, but may not have resulted in James discussing his SE in relation to this moment. Overall, while the quantitative survey data from the ESM can be used to identify possible high or low moments of SE, we are limited in investigating the moment without the open-ended ESM survey items, which support the exploration of the complexities of SE in relation to that moment.

Embedded Unit 2 from Fall 2021 (TED BROWN's Data): In Sub-Section 5.5.1: Following Threads, we discussed how we integrated Ted Brown's task-specific SE scores with the task and course open-ended ESM survey responses to build a thread to inform the writing of his daily journal prompt. In the following, we restate that thread.

> *Ted Brown's SE appears to be increasing as he is working on research project(s) for his [honors undergraduate research seminar course label] course based on the two different snapshots of his task-specific SE related to doing activities for this course.*
> - **Thread Identified By Researchers from Ted Brown's Fall 2021, Week 4, Day 1 ESM Data**

Drawing on this thread, which integrated both quantitative task-specific SE scores and qualitative open-ended ESM survey responses, the research team crafted the following daily journal prompt for Ted Brown. The aim was to confirm the observed increase in his SE and to gain deeper insight into

129

the factors that contribute to this increase, providing a richer understanding of his SE dynamics.

> *Yesterday it stood out to us that your confidence in your abilities in [honors under-graduate research seminar course label] was notably higher. Could you tell us what was happening yesterday? What were you doing? Who were you working with? What impacted your sense of confidence?* - **Ted Brown's Fall 2021, Week 4, Day 2 Daily Journal Prompt**

In examining the individualized daily journal prompt written for Ted Brown, the researchers drew on the ESM data in several ways as a result of using a thread that combined the quantitative and qualitative ESM survey data. Here, the researchers contextualized the increase in scores by only including the information from the course question - "...your confidence in your abilities in [honors undergraduate research seminar course label] was notably higher". This addressed the thread in which we recognized that his task-specific SE scores increased from one moment to the next moment in the day, and while both moments were about activities in the same course, Ted Brown did not provide enough information for us to know if the tasks themselves were the same, related, or completely different. This led us to not contextualize the prompt with task information. However, a limitation of this approach is the lack of task-specific detail - while we wrote a prompt that tied the increase in his SE to the course level, we did not write a prompt that further explored his task responses to the ESM survey data at those two points in time. This ambiguity reduces the grainularity of the prompt and potentially limits our ability to fully understand how specific tasks influenced his SE.

We did not discuss the research project specifically within this daily journal prompt. In part, this could have been due to one of his task open-ended responses only telling us that he was working on his research project for his [honors undergraduate research seminar course]. Even though Ted's open-ended response at Notification 3 tells us he was working on the research for his honors course, this amount of detail does not support understanding the specific tasks or context surrounding the work at this moment. This left room for ambiguity regarding the nature of the work on the project

he was doing (e.g., analyzing data versus building the presentation) or anyone he was working with, which are all factors that may influence Ted's SE. However, this ambiguity also highlights an important aspect of our methodology: it underscores the need for careful prompt design. By integrating both quantitative and qualitative survey data, we can tailor daily journal prompts to specifically probe these uncertainties, asking for more detail about the task or course context to better understand the factors influencing SE. This flexibility in our approach allows us to bridge these gaps and delve deeper into the complexities of SE. In this case, using the thread, we wrote a prompt for Ted that gave him the freedom to discuss the tasks in relation to this course and his increase in SE.

At the same time, there is an advantage in not directly drawing on language of the tasks Ted reported in Notification 3 and 4. While he may not provide additional information about those moments, the daily journal prompt provided enough contextual detail to prompt Ted to discuss his SE situated within the [honors undergraduate research seminar course] experience on Day 1. This also allowed him the agency to describe what specifically contributed to the increase in his SE. By writing the daily journal prompt in this way, we avoided leading Ted to attribute his increased SE to the research project itself or those particular tasks from Notification 3 and 4, thus reducing the potential for researcher bias in framing the prompt (e.g., assuming that working on the project directly impacted his SE.

Moreover, because Ted was working on the project that day, we anticipated that he would naturally share how the project influenced his SE, even without a direct prompt to do so. This approach assumes that Ted's response would naturally draw upon the relevant aspects of his experience, which might not be as explicitly captured if we had directly asked about the project. Additionally, given that data collection occurs across multiple days, we can always follow up with Ted on subsequent days to probe further into both his ESM data and his previous daily journal prompt responses, allowing us gather richer context over time or even write another daily journal prompt to further probe into his experience around these tasks and this course again.

What made my confidence get much higher in that course was a combination of two

131

things. One of them was that for an assignment that I was given for the research project

I went above and beyond of what was asked of me and did more than just look at some

of the data that I was given but I made a full PowerPoint presentation about my analysis

and conclusion off of this. The second point that gained me confidence from this was

that my research mentor was very impressed and made a point to compliment many

specific details about both me even putting together a presentation and also how I set

it up overall. This all made me gain a lot of confidence in myself and in my work. -

Ted Brown's Fall 2021, Week 4, Day 2 Daily Journal Prompt Response

Ted Brown confirmed that his SE in the course had increased by stating "What made my confidence get much higher in that course was a combination of two things." One of these experiences involved going "above and beyond" by building a full PowerPoint presentation to present his analysis and conclusions, rather than simply reviewing data. Ted reported on building this presentation one of the survey responses; specifically, shared this information in the task open-ended ESM survey question. His daily journal prompt response supports why building the presentation may be a moment of high task-specific SE in his day or at the very least important to his SE. The second experience that contributed to Ted's increased confidence was receiving positive feedback from his research mentor about the presentation and its setup; unfortunately, this experience was not captured by the ESM data collection. Most likely because he was not given the survey at the time he was meeting with his advisor or directly after that meeting.

This example demonstrates that while having contextualized threads from integrating the quantitative and qualitative ESM survey data supports writing contextualized daily journal prompts, prompts that only include course information and score information result in different advantages and limitations. An advantage of this prompt is that Ted Brown shared multiple experiences, one mastery experience and one social persuasion experience, that influenced his SE within the course still. Another advantage of this kind of thread is that Ted was able to share an experience with a research mentor, even though this was not captured in his ESM survey data, demonstrating that an overall advantage of the daily journal prompts is being able to learn about other experiences

that may influence students' SE. There is value in having Ted share this experience, regardless of whether it is captured in the ESM data, because this experience provides us with information about how Ted's SE is shaped as well as allowing Ted to report on experiences closer to when those experiences occurred still. The limitation of capturing this experience like this in the daily journal prompt but not in the ESM survey data is that we can't use this to understand Ted's scores.

In the next couple of paragraphs, I will explore the limitation associated with not being able to interpret Ted's scores through the use of the daily journal prompt and response, and discuss how the integration of quantitative and qualitative ESM data demonstrates that 1) qualitative data alone is insufficient, 2) tailored, context-specific journal prompts are essential, 3) integration provides richer insights, and 4) SE is complex, requiring integrated approaches to capture its full complexity. Additionally, I will show that threads can also be identified from the daily journal prompt response, and then be used to write new daily journal prompts.

Based on Ted's daily journal prompt response, we identified two key experiences - one mastery experience (i.e., going above and beyond by building the presentation) and one social persuasion experience (i.e., feedback from the research mentor) - that influenced his SE. These experiences served as the foundation for two key threads that we sought to explore further. First, we were interested in understanding how Ted weighted these experiences in shaping his SE, and what informed his evaluation of their relative importance. This was also driven in part by knowing that men tend to place more emphasis on mastery experiences then social persuasion experiences (CITE). Second, we aimed to follow the thread of his research mentor's feedback to gain deeper insight into how specific comments or social persuasion influenced Ted's SE. Following up on these threads, identified from the previous daily journal prompt, resulted in this individualized daily journal prompt for Ted Brown:

> *Yesterday, you told us about your two experiences in your [honors undergraduate research seminar course label] (going above and beyond what was asked of you and your interaction with your research mentor) that increased your confidence in your [honors undergraduate research seminar course label] course.*

When thinking about these two experiences, how would you rank which one had more of an impact on your confidence in your [honors undergradaute research seminar course label] course?

Yesterday, you also told us that your research mentor complimented many specific details of the presentation and your overall setup. Was there a specific compliment or comment from your research mentor that impacted your confidence? - **Ted Brown's Fall 2021, Week 4, Day 3 Daily Journal Prompt**

This follow-up daily journal prompt is an important tool for investigating Ted's SE because it builds directly on the experiences identified in the initial ESM data and daily journal prompt response. Ted's integrated ESM data, which showed an increase in his task-specific SE scores across two distinct activities related to his honors undergraduate research seminar course, provides the foundational thread that led to the formulation of the follow-up prompt. The Fall 2021 integration of quantitative ESM scores with qualitative ESM open-ended survey responses allows for a more nuanced exploration of the why and how behind Ted's SE fluctuations. While the daily journal prompts alone can provide deep insights into the experiences that influence SE (that is, in Fall 2020 we still could have written follow-up daily journal prompts from previous daily journal prompts and response without the ESM data), the integration of the quantitative ESM scores with the qualitative ESM open-ended survey responses allows for a more nuanced exploration of how and why SE changes occur. In particular, the follow-up daily journal prompt enables us to probe deeper in the relative impact of mastery experiences versus social persuasion on Ted's SE, something that is enhanced by the added temporal context and trend information provided by the quantitative data. This becomes more evident when considering Ted's responses to this daily journal prompt:

I would say the positive interaction with my mentor had a more significant impact on my confidence, but both of these things complimented each other well. Where because

I felt proud of my work, I was very happy that someone else agree[d], and because [some]else complimented my work it made me more proud of it.

One specific comment that she made was that putting all my results into a clean PowerPoint was the absolute correct to do for any project like this and that she was especially impressed that I had put all my conclusions clearly as well. These compliments greatly reinforced my confidence in that my judgment and work in the course is good. - **Ted Brown's Fall 2021 Week 4, Day 3 Daily Journal Prompt Response**

From Ted's response shown above, we gain a more nuanced understanding of how he values the two experiences that contributed to his increased SE within the course. Ted reports that the positive interaction with his mentor had a more significant impact on his confidence, although he acknowledges that both the mastery experience and social persuasion experience worked in tandem to bolster his SE towards the course. His reflection highlights that the mentor's feedback not only reinforced the pride he felt in his work but also amplified that feeling by validating his judgment. This reinforces the idea that SE can be shaped by both mastery and social persuasion experiences, and his response underscores the complementary role of mastery experiences and social persuasion in building his SE, suggesting that an interaction of these factors, rather than a single experience, plays a critical role in shaping his confidence.

However, this leads to some important questions about how we interpret task-specific SE scores within the ESM surveys used in this study. Ted told us that both experiences contributed positively to his SE, but how does this translate into his ESM scores? For instance, since the two experiences happened in close succession, should we expect to see a consistent increase in his SE across his data? Or, because the mentor's feedback was perceived as more meaningful to him, would his SE score be higher after this interaction, reflecting the greater weight he places on social persuasion in shaping his confidence? Another consideration is whether Ted treats these experiences as independent, rating them separately based on their individual impacts, and thus potentially giving a higher score to the experience that had the most significant influence on his SE. This highlights that while the

daily journal prompts and responses provide useful insight into Ted's SE, the qualitative data alone are not enough to fully capture the complexities of his SE.

This question points to a limitation in the design of this study: the ESM data and the journal prompts do not always capture the same experiences. In this case, the interaction with Ted's mentor, which played a key role in his SE, was not directly recorded through the ESM task-specific SE scores. Instead, the data only reflects his work on building his presentation. This discrepancy highlights both an advantage and limitation in how different data sources (ESM and daily journal prompts) may capture separate or the same moments of experience, and how these moments may interact to shape students' scores. Without capturing the mentor meeting directly in the ESM data, we are left with an incomplete picture of how these experiences contribute to Ted's SE.

Some of this disconnect may stem from the phrasing of the journal prompt itself, which was designed to probe into Ted's experiences within the course. Specifically, the research team structured the daily journal prompt to encourage Ted to reflect on events or experiences from the previous day, without prompting him to discuss specific tasks or activities he had already reported through the ESM survey data. This highlights that a thread that is more focused on course without the task information may result in a thread that investigates particular moments in relation to the course over the course of the day that the ESM data may not have captured. Thus, an advantage of this kind of thread is that the daily journal prompt could expand upon the ESM data through capturing a moment that is influential for the student's SE that the ESM data did not capture. Another kind of advantage of a thread is that the daily journal prompt could expand upon the ESM data through capturing a moment that is influential for the student's SE that the ESM data did capture and could expand upon this moment through explaining how and why this moment influenced the student's SE.

Ted's case highlights that there may be significant value in linking the ESM quantitative data with the qualitative data from the open-ended ESM survey responses to build a thread. By integrating these different data sources, researchers can gain a more comprehensive understanding of the student's SE and identify key experiences that may influence their SE. Once this thread is

established, it becomes possible to craft a daily journal prompt that can explore the complexities of SE. In this case, we decided to investigate his experiences within the course across the day, without necessarily needing to include the task-specific details. Even when the thread contains task-specific and course-related information as a result of this integration, researchers have the flexibility to choose which aspects of the thread to include in the daily journal prompt, leading to various types of daily journal prompts.

For example, with Ted, we chose to write a daily journal prompt that explored the increase in his scores in relation to his experiences within the course throughout the day. However, we could have alternatively crafted a prompt that encouraged Ted to discuss, in greater detail, the specific work he was doing for his research project from the survey response to the latter ESM survey response, and specifically how that influenced his SE. This illustrates a richer thread - one that includes both task and course-related open-ended responses from the ESM survey - opening up more options for how researchers explore a student's SE with the daily journal prompts.

In contrast, James's case highlighted a limitation of relying solely on ESM quantitative data. Without the integration of the open-ended ESM survey data, the daily journal prompt could only focus on the changes in his task-specific SE scores. While this gives us some insights into James's SE, there is no way to contextualize this information within the task or even the broader course experience in the daily journal prompt. We essentially got lucky that James did this work with his daily journal prompt response, in that his response did provide more SE information about a moment he reported on the ESM survey. While his response did the work we wanted in terms of expanding up on a moment he reported on in the survey, we cannot definitively say that the daily journal prompt response captured the complexity of the moment as related to James's SE scores. Specifically, how do we know that the daily journal prompt response accurately reflects the underlying factors influencing James's SE scores without still relying on the qualitative information provided in the ESM survey? Thus, while using a richer thread with both task and course-related open-ended responses offers more flexibility and depth in writing daily journal prompts, it also emphasizes the importance of carefully integrating both types of data to accurately capture and

reflect the student's lived experience of SE.

As shown with Ted's case, this is not to say that we necessarily have to use the task and course open-ended ESM survey responses. Writing a more general prompt around Ted's SE scores in relation to the course, without incorporating the task-specific details, still offered flexibility for Ted to discuss various aspects of his experience. Notably, it allowed him to reflect on his interaction with his mentor, which was not captured by the ESM data but was critical for his SE. However, when a thread includes both task and course qualitative information from the ESM survey responses, it provides researchers with additional options to explore and investigate a wider range of influences from the student's day that may be influencing their SE. This shows that daily journal prompts can expand upon the ESM data in two ways: (1) by expanding on the SE information related to moments already reported in the survey, as seen in James's case, and (2) by capturing additional moments from the student's day, as was partially the case for Ted.

5.5.5 Comparing Embedded Unit 3 from Fall 2020 (Lenny's Data) to Embedded Unit 3 from Fall 2021 (Rose's Data)

Embedded Unit 3 from Fall 2020 (LENNY's Data): Two threads were identified in Lenny's ESM data, as discussed in Section A. After examining his quantitative ESM data, the two threads identified in his task-specific SE scores were: (1) a dramatic increase from the 20s-50s on Day 3 to the 80s-100s on Day 4, and (2) a possible decrease at Notification 14 on Day 4.

As the first author analyzed Lenny's data post-collection, the daily journal prompt could have addressed either of these threads or been phrased more broadly, allowing Lenny to explain what was influencing his task-specific SE without explicitly focusing on either thread. The latter approach would let Lenny describe the changes in his SE without us biasing him to discuss one thread or the other, assuming that the two threads were influenced by different factors. Appearing to use this approach, the research team developed the following daily journal prompt to further explore Lenny's task-specific SE fluctuations.

By your last notification today in the app it was clear that something had changed in

your day. Can you tell us anything about what affected that change? - **Lenny's Fall**

2020 Week 1, Day 4 Daily Journal Prompt

In evaluating the daily journal prompt, the writer refers to the ESM surveys using the phrase - "By your last notification today in the app." This language links the daily journal prompt to the ESM survey data and signals to Lenny that we are interested in understanding his responses. The writer then follows up with "it was clear that something had changed in your day. Can you tell us anything about what affected that change?" This statement acknowledges the change in Lenny's task-specific SE scores, without explicitly identifying the two threads. Instead, it conveys to Lenny that we have observed changes in his ESM responses and are asking him to reflect on what influenced those changes, without presupposing the cause.

Although the intention of the daily journal prompt was to have Lenny share what influenced changes in his task-specific SE scores, the prompt does not directly focus on these variations. Instead, it asks Lenny to describe what changed during his day. The daily journal prompt mentions the ESM survey data with the phrase, "By your last notification today in the app," but then shifts focus to broader changes in Lenny's day with "it was clear something had changed in your day." This shift prompts Lenny to reflect on changes that occurred specifically on that day, rather than on Day 3 and Day 4. This limitation may limit the ability to fully capture the key moments or emotional responses from Day 3 that may have influenced Lenny's SE.

Another limitation of this approach is that, while one of the identified threads focuses on changes from Day 3 to Day 4, the prompt does not explicitly make it clear whether Lenny can reflect on changes across these two days. The prompt doesn't necessarily prevent him from doing so, but the open-ended phrase encourages reflection on the current day, which may unintentionally exclude important information from previous days - particularly regarding shifts in SE scores between Day 3 and Day 4. As a result, Lenny might not discuss his SE directly or may omit key moments he reported on the ESM surveys, especially those from Day 3. Even if Lenny shares moments, they may not align with the survey responses. This means that integrating the ESM data with the daily journal prompt responses to build a coherent narrative about what influenced Lenny's SE could be more challenging. Since the daily journal prompt responses do not explicitly reference or

discuss the ESM data, it becomes difficult to directly link the two. While the daily journal prompt responses may suggest potential factors influencing Lenny's SE, without explicit connections to the ESM data, we cannot fully integrate these insights to capture a complete understanding of the factors driving his SE across different days.

Relying exclusively on quantitative ESM data to build the thread that informed the writing of this daily journal prompt limited the ability to contextualize and identify patterns or changes in task-specific SE scores, which then limits us in being able to contextualize and follow up on those patterns with daily journal prompts. Thus, the daily journal prompts and its responses are limited in how they may help us understand students' SE. In the absence of this integration, the research risks producing a partial or skewed understanding of what drives changes in task-specific SE, potentially overlooking important contextual, emotional, or situational influences that quantitative data cannot capture. These limitations become evident when turning to Lenny's response to his daily journal prompt.

> *"I started today with my Vietnamese class and it was a review class where we got to speak to native Vietnamese speakers. It was an overwhelming time for me since I've only had 4 weeks of Vietnamese and am at about a young toddler level of understanding. I couldn't understand any of the fluent speech and it felt like a train wreck trying to communicate with them. After the class I got to work on my senior thesis which is heavily based in programming and data science. Areas that I am and feel competent in. What I was doing drastically changed throughout the day. I went from doing something I have very little experience in to something I have years of experience in."* - **Lenny's Fall 2020, Week 1, Day 4 Daily Journal Prompt Response**

In his daily journal prompt response, Lenny explained that the change in his day was due to his different feelings toward his Vietnamese class compared to his senior thesis work. His response focuses on describing the change in his day rather than necessarily reflecting on how it related to the change in his ESM scores. For example, he described his experience in the Vietnamese class,

where he had the opportunity to speak with a native Vietnamese speaker, as "overwhelming." He mentioned that he "couldn't understand any of the fluent speech," which reflects his perception of his limited ability in the language. He even referred to the experience as "a train wreck trying to communicate with them." This contrasts with his senior thesis work, which he feels more competent in, as it involves areas "that I am and feel competent in." This statement is a reflection of his SE - Lenny feels confident in his ability to handle his senior thesis work. He summarized the shift he experienced throughout the day by saying, "What I was doing drastically changed throughout the day. I went from doing something I have very little experience into something I have years of experience in." Lenny shared what we had asked for, which is what has changed across his day, and although he related this shift to his SE, we must consider whether this change is reflected in his task-specific SE scores.

To link Lenny's experiences with the ESM survey scores, he would have needed to report on these specific experiences within the ESM surveys that were completed on the same day. Although his SE may have changed throughout the day as he described, this change may not have been captured within the ESM scores themselves. This illustrates that relying solely on the ESM quantitative scores and the daily journal prompts is insufficient. To make that connection, we would need to know what Lenny was doing at the exact moment he completed the surveys. Reviewing his open-ended ESM survey responses would allow us to determine whether they align with his experiences in his Vietnamese class or shortly after his class, as well as whether we captured his work on his senior thesis project. Although the research team did not use the task and course open-ended ESM survey items during the data collection, they would later need to consider both the task and course open-ended ESM survey responses in a post-hoc analysis to confirm whether these experiences are reflected in Lenny's ESM scores.

Examining Lenny's response to the task open-ended ESM survey question in Table X, we see that his experience in the Vietnamese course earlier on Day 4 was not captured in the ESM data. However, we did capture him practicing or working on his Vietnamese on Day 3, as reflected in Notifications 11 and 12. This aligns with the lower task-specific SE scores for Day 3, as compared

Notification Number	Reported Task
13	Walking to the store
14	Physics senior thesis research
15	Senior thesis work
16	Getting KFC

Table 5.8 Lenny's reported task associated with each notification for his Week 1, Day 4 data collection.

to the scores on Day 4, when he was working on his senior thesis. Lenny's daily journal prompt suggests that these changes in his SE are consistent with his description of his experiences in the Vietnamese course versus his senior thesis work: "I went from doing something I have very little experience in to something I have years of experience in." He also directly contrasts his competence in programming and data science, which he applies to his senior thesis, with his limited competency in Vietnamese, noting that after "only 4 weeks of Vietnamese," he is at a "young toddler level of understanding."

Overall, there is a disconnect between Lenny's ESM data and his daily journal prompt response. On Day 3, he reported practicing for the ESM surveys, but in his daily journal prompt response, he discussed his experience in class, where he had to speak with a native Vietnamese speaker. While we can infer from his daily journal prompt response why his task-specific SE scores were low on Day 3, Lenny did not reflect on the specific experience mentioned in the ESM data. This disconnect highlights a limitation in how we've implemented the ESM along side the daily journal prompts. Using the ESM along side the daily journal prompts means we can capture a wide range of experiences. However, when these experiences aren't aligned or reflected consistently across both the ESM data and the daily journal prompt responses, we lose the ability to capture an in-depth understanding of the moment reported on the ESM. This doesn't necessarily reflect a flaw in the ESM itself but rather a challenge in how we integrated it with the journal prompts. While the advantage of this daily journal prompt and response is that it captures a holistic view of Lenny's SE changing across the day, the limitation arises when the data from the two sources doesn't align, making it difficult to link specific experiences with SE fluctuations. Thus, there are benefits and limitations to how we link and interpret a student's SE when the data isn't fully synchronized,

limiting the advantage of using a mixed methods approach.

We have discussed a bit about how Lenny's response to the daily journal prompt provides valuable insight into his SE across different tasks and courses. From his response, we learn that Lenny feels incapable of speaking Vietnamese when interacting with a native speaker. He described this experience as "overwhelming," which reflects a perceived threat to his SE in this area. In contrast, he feels confident in his ability to handle his senior thesis, which involves programming and data science — areas where he has more experience and competence. This suggests that Lenny's SE for his senior thesis is likely higher than his SE for his Vietnamese course. Therefore, we might expect to see lower task-specific SE scores for tasks related to his Vietnamese course compared to those related to his senior thesis. Additionally, his response indicates that Lenny draws on mastery experiences—his years of experience in data science and programming—as a key factor in assessing his SE. However, we did not explore other potential experiences, such as Lenny's broader engagement with the Vietnamese language, which may also shape his SE.

While the journal prompt provided useful insights into Lenny's SE, it's important to note that the question didn't explicitly ask him to reflect on it. Lenny's response about how his confidence changed across the day could be seen as a fortunate coincidence — it just so happened that he linked the change in his day to changes in his SE. This is not a flaw in Lenny's reflection, but it raises an important point: without specifically prompting students to reflect on their SE, it's possible that other students may not mention it at all. This challenges the consistency of the prompt's effectiveness in capturing SE across different individuals. Although Lenny's response adds validity in that he was not directly led to discuss his SE, it also suggests that the prompt may not always elicit this reflection from every student, making it harder to compare responses or generalize findings.

In sum, Lenny's data provides valuable insights into the advantages and limitations of using ESM in conjunction with daily journal prompts. While the ESM data captured fluctuations in his task-specific SE, the journal prompt was less precise in directly linking these changes to specific experiences. The disconnect between the ESM data and Lenny's self-reported experiences in the journal highlights the challenges of integrating multiple data sources, especially when they aren't

fully synchronized. While this mixed-methods approach offers a broader view of Lenny's day, it also restricts our ability to draw clear, consistent connections between particular moments and SE fluctuations. These findings emphasize the importance of refining the integration of ESM data and journal prompts, particularly by considering open-ended responses from ESM surveys in future analyses.

Embedded Unit 3 from Fall 2021 (ROSE's Data): As discussed in Section A. , we observed that Rose was working on a variety of tasks associated with her biology course and her chemistry course, and that her task-specific SE scores, regardless of the activity, were reportedly lower for tasks associated with her biology course then her chemistry course. To explore this thread, the research team wrote the following daily journal prompt for Rose.

> *Over the last 3 days, your responses in the app indicated you were working on both [Cell and Molecular Biology Course Label] and [General Chemistry Course Label]. However, you indicated you were less skilled and felt less successful on [Cell and Molecular Biology Course Label] than [General Chemistry Course Label]. Can you tell us a little bit more about why you feel differently in these two classes?* - **Rose's Fall 2021, Week 1, Day 3 Daily Journal Prompt**

Upon examining this daily journal prompt, we see that the research team made a concerted effort to link it to the ESM surveys in a broader sense, with the phrase "Over the last 3 days, your responses in the app. . . " This phrase connects the journal prompt to Rose's ESM data over the past few days, framing the conversation about her task-specific SE within the context of the data we had already observed. The prompt goes on to acknowledge that Rose had been working on tasks for both her biology and chemistry courses, and highlights that her SE scores for tasks associated with biology were lower than those for chemistry. By directly stating that Rose appeared to report that she felt "less skilled" and "less successful" in biology, the prompt was designed to invite Rose to reflect on the reasons behind these differences.

The phrasing of this prompt was intentional - it steered Rose toward discussing the potential

disparity between her perceived skill and success in the two courses, which aligns with the task-specific SE scores that we had seen in her ESM data. Additionally, the prompt implicitly sets up a comparison between the two courses and their impact on Rose's SE. This directionality in the prompt might limit the scope of her reflection to a narrow focus on just the two courses, without accounting for other potential factors influencing her SE. Her daily journal prompt response reflects these decisions.

> *I feel differently about BS161 and CEM141 because of my recent scores in both of the classes. In high school, my strongest subject was AP biology, and my weakest subject was AP chemistry. With that in mind I thought that's how my college education would be as well. However, my chem class is easier for me to understand, rather than my biology class. In chemistry the lectures are over zoom, but I have in-person recitations that really help me understand. Also, we get weekly homework assignments that help me understand the material. However for my biology class, it's completely online, and we don't have homework assignments. Learning depends on self studying, and the prequizes given twice a week. I've been struggling learning how to self teach myself through an online platform.* - Rose's Fall 2021, Week 1, Day 3 Daily Journal Prompt Response

Rose's response to the daily journal prompt provides a detailed, personal explanation for the differences in her SE between her biology and chemistry courses. Her prior experiences with AP Biology and AP Chemistry shaped her expectations for college courses. Despite struggling with chemistry in high school, Rose finds her chemistry class easier to understand than her biology class in college, which contrasts with her initial expectations. One key factor contributing to this shift is the learning environment. Rose mentions that her chemistry course is delivered through Zoom lectures, supplemented by in-person recitations, which she finds helpful for understanding the material. In contrast, her biology course is entirely online, with no in-person support. This self-directed nature of the biology course has made it more challenging for Rose to learn effectively,

145

especially since it lacks homework assignments and relies heavily on pre-quizzes and self-study. As a result, the opportunities available to support Rose in building her SE differ across the two courses. Although she was stronger in biology during high school, her lack of confidence in her ability to self-teach in an online format has lowered her SE in her college biology course. On the other hand, her positive experiences in chemistry, particularly with its structured learning support, have bolstered her SE. Weekly homework assignments and interactive recitations provide frequent feedback and opportunities to solidify her understanding, likely contributing to her higher SE in chemistry. In contrast, the online structure of the biology course, which lacks regular assignments and relies on self-teaching, has likely contributed to lower SE.

While Rose's journal response provides insights into why her task-specific SE scores were lower for biology and higher for chemistry, the daily journal prompt response did not explicitly capture any specific moment Rose reported on her ESM surveys in relation to these fluctuations. We can infer that the challenges Rose described in her biology course contributed to lower SE scores for the tasks she reported. The lack of support - specifically, the absence of regular homework assignments and in-person interaction - likely made the biology course tasks more difficult, which are most likely reflected in her lower task-specific SE scores for the biology-related tasks.

Conversely, her higher task-specific SE for chemistry tasks aligns with her description of the course's supportive learning structure. The availability of in-person recitations and the regular homework assignments likely helped boost her confidence in handling chemistry tasks, which is most likely reflected in her higher task-specific SE scores for chemistry tasks.

Even though Rose's daily journal prompt response provides useful insight into her SE, the disconnect between her ESM data and the daily journal prompt responses remains a limitation. The ESM surveys captured SE at specific points in time, but without direct references to those moments in Rose's daily journal prompt responses, we cannot conclusively link her experiences discussed in the daily journal prompt response with the task-specific SE fluctuations. As a result of the daily journal prompt response, we can only infer how her feelings towards biology and chemistry impacted her task-specific SE.

Some of this disconnect may stem from the phrasing of the journal prompt itself. The prompt guided Rose to explain her differing feelings towards these courses, which may not fully capture the nuanced fluctuations of SE. For example, Rose does not mention specific moments of success or struggle that could have influenced her SE at particular points during the week. This highlights the challenge of relying on the daily journal prompts to investigate dynamic changes in SE throughout the day or across multiple days.

Overall, the advantage of Rose's journal prompt is that it helps us move towards understanding her task-specific SE fluctuations by shedding light on how course structures may influence her SE. Specifically, the presence or absence of key supports such as weekly homework assignments or in-person recitations that may contribute to higher or lower task-specific SE scores. Her daily journal prompt response provides valuable insights into how these factors shape her course-level SE toward the biology and chemistry courses. However, a limitation of the prompt is that it does not guide Rose to reflect on specific moments within those classes that could influence her SE at particular points. More importantly, it doesn't link her responses to the particular moments of her ESM surveys, which may reveal how her SE fluctuated throughout the day or across multiple days. Instead, Rose's daily journal prompt response focuses on her general feelings towards the courses. While this still provides useful context, it does not allow for a precise connection to the moments and their associated task-specific SE as reported in the ESM data. This case also underscores that using daily journal prompts solely using course information may not prompt students to discuss specific moments that directly affect their SE. It suggests that including task and course-related information in prompts (as done in the open-ended ESM surveys and their use in building threads in Fall 2021) may be more likely to elicit responses that address specific moments in time, providing deeper insights into students' SE.

In contrast, Lenny's daily journal prompt was more broad as we did not include course or task information, but rather asked him what had changed across his day to understand what might influence his ESM responses to change. His daily journal prompt prompted him to recall specific events related to his Vietnamese course in relation to the skills needed for his senior thesis project.

Unlike Rose's more general reflection about two courses, Lenny's prompt successfully related his SE to a concrete moment in a classroom and to working on a project. Similarly to Rose, Lenny's daily journal prompt response and ESM data were not able to be directly linked either, but even so, his detailed exploration of specific events in the courses helped provide valuable insight into his SE, including his task-specific SE scores, and highlighted the importance of prompts in eliciting rich, moment-focused reflections. The comparison between Rose's and Lenny's cases suggests that while prompts can reveal insights into course-level SE, task-focused prompts that ask for reflections on specific moments within the course are more likely to yield responses that directly tie the task-specific SE fluctuations to key events reported on ESM surveys. This is critical for being able to understand how specific moments are impacting students' SE development.

5.5.5.1 Key Takeaway: Integrating Quantitative and Qualitative Data Leads to Richer, Context-Sensitive Daily Journal Prompts

The cross-case analysis of the Embedded Unit 1 from Fall 2020 (Jane's Data) and Fall 2021 (Robert's Data), Embedded Unit 2 from Fall 2020 (Jame's Data) and Fall 2021 (Ted Brown's Data), and Embedded Unit 3 from Fall 2020 (Lenny's Data) and Fall 2021 (Rose's Data) illustrates the limitations and advantages of relying on threads from solely using the quantitative ESM survey data versus integrating the quantitative and qualitative ESM survey data to built the thread that informs the creation of the individualized daily journal prompts.

From the Embedded Units 1-3 from the Fall 2020 data collection, the individualized daily journal prompts were constrained by the sole reliance on the quantitative ESM survey data, which provided valuables insights into the fluctuations in SE but lacked the contextual richness needed to fully understand these changes. The individualized daily journal prompts created in these cases linked to the broad trends in the ESM data, using language like "Today in the survey data you had a few very confident responses and then a few not confident responses", "Today in the survey data it appears that you had a rough patch in your day where you felt less in control and skilled", and "By your last notification today in the app it was clear something had changed in your day," but did not incorporate task-specific or course-level details that might have clarified the moments driving these

fluctuations. As a result, while the prompts helped to explore shifts in SE or, in Lenny's case, shifts in his day, they did not allow for a deeper investigation into the specific tasks, events or experiences influencing these changes.

In contrast, the integration of quantitative and qualitative ESM survey data to build threads in the Embedded Units 1-3 (Robert's, Ted Brown's, and Rose's Data) from Fall 2021 provided greater flexibility and depth in creating individualized daily journal prompts. By incorporating the open-ended ESM survey responses, researchers were able to identify more nuanced threads, linking task-specific and course-level experiences to fluctuations in SE. This allowed for the creation of prompts that encouraged students to reflect on specific moments within their coursework or particular tasks, leading to richer, more contextually informed responses.

This analysis demonstrates that while quantitative ESM data can provide valuable insights into changes in SE, it is the integration of quantitative and qualitative ESM data that truly enriches the process. By using both types of data, researchers can develop individualized daily journal prompts that not only track shifts in SE but also delve into the specific experiences, tasks, and contexts that shape these fluctuations. Ultimately, this integrated approach allows for a deeper and more complete understanding of students' SE, allowing for exploration of the specific moments when these shifts occur.

Furthermore, the Embedded Unit 3 from Fall 2020 and Fall 2021 illustrate that, while prompts can reveal insights into course-level SE, task-focused prompts – especially those that ask students to reflect on specific moments within a course – are more likely to yield responses that tie the task-specific SE fluctuations to key events reported in the ESM surveys. This highlights the importance of crafting daily journal prompts that not only explore broad trends but also probe the concrete task-related moments that may drive fluctuations in SE.

Additionally, the Embedded Unit 3 from Fall 2020 (Lenny's individualized daily journal prompt), and Embedded Unit 2 from Fall 2020 (James's individualized daily journal prompt) and Fall 2021 (Ted Brown's individualized daily journal prompt) demonstrate that individualized daily journal prompts can encourage students to reflect on specific moments without necessarily

including task-specific information. These reflections, however, may not always align with the ESM data if those moments were not captured in the surveys. Therefore, one key takeaway is that individualized daily journal prompts can expand on the ESM data in two ways: (1) by expanding on the SE information related to moments already captured by the surveys or (2) by capturing new moments that may influence SE but were not initially reflected in the ESM data.

The Embedded Unit 2 from Fall 2021 (Ted Brown's Data), in particular, underscores the value of integrating both types of data. By building a thread from both quantitative and qualitative ESM survey data, the individualized daily journal prompt allowed for a deeper exploration of SE fluctuations. This also highlighted that through well-crafted prompts, the data captured in the individualized daily journal prompts could uncover additional threads, providing a richer understanding of the student's experiences, to be followed up on in the individualized daily journal prompts.

Overall, this cross-case analysis emphasize the advantage of combining quantitative and qualitative ESM survey data: while the quantitative ESM data alone provides a valuable snapshot of SE fluctuations, the qualitative ESM data provides contextual information to situate the fluctuations within moments from the students' lives leading to threads that enhance the richness of the prompts. This integrated approach not only strengthens our understanding of SE but also allows researchers to investigate its complexities and the nuanced factors that shape it when these threads are used to inform the writing of the individualized daily journal prompts. However, as seen in the Embedded Units 1-3 from Fall 2021, decisions about how to write prompts must carefully consider the complexity of SE as a result of integrating both types of data supporting research in crafting more meaningful and specific individualized daily journal prompts. Prompts can be written to guide students to reflect not just on how their SE changed, but also on why those changes happened, tied to specific tasks or experiences they encountered. In essence, having both types of data available helps ensure that the prompts are designed to encourage more informative and meaningful reflections from students. These reflections provide a clearer picture of how their experiences and specific moments in their day influence their SE.

5.6 Discussion and Conclusion

The nested and embedded case presented in this chapter highlights the methodological significance and practical benefits of integrating the open-ended (qualitative) and closed-ended (quantitative) ESM data when designing individualized daily journal prompts. This integration serves not as a strategy for enhancing the threads observed from the ESM data and the contextual richness of prompts but also as a way to deepen our ability to explore SE in relation to real-time experiences.

Through comparisons between the Fall 2020 and Fall 2021 embedded units, we see a clear progression in the quality and specificity of the daily journal prompts as a result of the methodological shift. In the 2020 cases, where only closed-ended (quantitative) ESM data were available, researchers were limited to tracking general trends in task-specific SE. Prompts based solely on this data could point to when a shift occurred but lacked the contextual grounding needed to explore why the shift took place. As a result, student reflections were often broad and disconnected from the specific events or tasks that likely influenced their SE.

In contrast, the 2021 embedded units show how integrating open-ended (qualitative) ESM responses enriches the threads that inform daily journal prompts. By linking task-specific SE fluctuations to particular academic experiences—such as working on a research project, preparing for a presentation, or navigating multiple courses—researchers could construct prompts that were both more targeted and more meaningful. These prompts guided students to reflect not only on changes in their SE, but also on the sources and contexts driving those changes, aligning directly with the dissertation's aim of making visible the domain-specific and source-specific nature of SE.

Importantly, the findings also reveal that individualized daily journal prompts do more than follow threads—they can generate new ones. In several cases, the individualized daily journal prompts provided the opportunity to explore in greater depth the experiences reported in the ESM data, allowing for follow-up questions or prompts that could further illuminate the factors influencing changes in SE and how students interpret these factors. This suggests that the relationship between the ESM surveys and journal prompts is not strictly unidirectional. Rather, prompts based on integrated data can open new analytical pathways, expanding our understanding of students' SE

experiences beyond what was initially captured in survey responses. This is a critical insight, as it positions the prompts not only as tools of reflection but also as instruments for extending the analytic reach of the research.

The analysis demonstrates that the integration of qualitative and quantitative ESM data significantly enhances the individualized daily journal prompts, making them more context-sensitive and aligned with the moment-to-moment fluctuations in self-efficacy. By using both types of data, this approach offers a deeper, more nuanced exploration of SE, situated within the real-time experiences of students. Moreover, the study contributes to the broader mixed methods literature by illustrating how integration can be practiced during the data collection phase, providing greater clarity and transparency in how integration supports the interpretive aims of the research.

Another critical takeaway from this research is the iterative nature of designing a mixed methods approach, which is a complex and evolving process. Designing a study that effectively integrates both quantitative and qualitative strands requires ongoing, intentional exploration of how to best combine the two data types. Each iteration of this study revealed new insights into how integration can be refined to better reflect the complexities of SE experiences. For example, the integration of open- and closed-ended ESM survey data in the Fall 2021 cases allowed us to craft more nuanced, context-specific journal prompts that encouraged deeper student reflections. These reflections were rooted in the specific tasks, courses, and experiences that shaped their SE.

This iterative approach highlights that designing mixed methods educational research is not a one-time decision but rather a dynamic, evolving process. The integration of data types, in this case, shaped not only how we approached the study but also what could be made visible in understanding students' self-efficacy. The shifting nature of self-efficacy itself—the way it fluctuates based on specific academic tasks, courses, and personal experiences—required flexibility and adaptability in our design. As the study progressed, the ways we crafted the individualized daily journal prompts evolved in response to both the quantitative ESM data and the emerging qualitative insights, ensuring that we captured a more complete picture of SE.

Furthermore, this chapter demonstrates that integration is not just a theoretical goal, but a

practical necessity when researching complex psychological constructs like self-efficacy. The integration of quantitative and qualitative data allowed for the exploration of SE in ways that would have been difficult or even impossible with only one type of data. Through iterative refinement, we ensured that the individualized daily journal prompts were not only reflective of shifts in SE but also contextually grounded in students' lived academic experiences.

Overall, this study underscores the value of an iterative approach to mixed methods research design—one in which integration is not treated as a single methodological step, but as an ongoing, adaptive process that evolves alongside the study itself. By continuously refining how quantitative and qualitative data interact—particularly through the integration of open- and closed-ended ESM data—we were able to design more context-sensitive, reflective instruments that made visible the nuanced and dynamic nature of students' self-efficacy. This approach demonstrates how real-time integration can shape not only the research design, but also what becomes possible to observe, interpret, and understand about students' academic experiences. In doing so, this chapter advances our methodological approach to studying SE and sets the stage for deeper explorations of how integration practices can be used to illuminate complex psychological constructs. Future work should continue to explore and extend these strategies, investigating how iterative, mixed methods designs can refine the tools we use—and the insights we gain—in educational research.

CHAPTER 6

DESIGNING DAILY JOURNAL PROMPTS TO CAPTURE THE THEORETICAL FOUNDATIONS IN SELF-EFFICACY

This study is intended to be submitted for publication in *Physical Review Physics Education Research*, a journal published by the American Physical Society (APS) and is dedicated to publishing significant research related to the teaching and learning of physics (REF). As such, this chapter was written with this intention in mind and was only edited for formatting to align with the requirements of the dissertation.

6.1 Introduction

Although research on self-efficacy (SE), one's confidence in their ability to successfully perform a task, has been highlighted for its critical role in student achievement [88], persistence in science majors [89], and career choices [90], current qualitative methods for exploring SE often come with limitations. Traditional qualitative approaches, such as interviews, allow participants to reflect on key events influencing their SE. They are typically conducted long after the events occur, leading to a potential disconnect between the real-time experiences and the reflections shared. This disconnect is in part a result of participants needing to reconstruct their past experience and, in turn, resulting in missing nuances of how their SE fluctuates in real time. To address this limitation, we have been designing a mixed methods approach, in which the primary qualitative data source leverages individualized daily journal prompts as a tool to explore students' SE closer to their experiences. In this work, we focus on the design and implementation of individualized daily journal prompts to explore how we capture the experiences and context-dependent nature of SE.

Although few studies have used daily journal prompts to examine SE, they offer a promising approach by allowing students to document their reflections shortly after key experiences, minimizing the time delay inherent in traditional methods (REF). However, existing research has often used weekly journals, which may not capture the immediacy that daily prompts can provide. Thus, this work will examine how the individualized daily journal prompts can serve as a tool to gain deeper and more immediate insights into students' academic journeys through the lens of SE, offering a

richer understanding of the factors that influence their SE in real time.

To explore SE in real time, we employed an explanatory sequential mixed methods design, combining the Experience Sampling Method (ESM) with individualized daily journal prompts. This approach allowed us to track students' fluctuations in their domain-specific SE through ESM surveys and then, through individualized daily journal prompts students were allowed to reflect on their experiences in relation to their SE. The integration of these methods across two iterations of data collection (Fall 2020 and Fall 2021) allowed us to track how the addition of an additional point of integration provided deeper insights into how students' SE evolves in response to academic events. The addition of this contextual information is key to understanding the shifts in how we write and use these prompts to capture the nuances of SE (see Chapter 5).

To further explore how individualized daily journal prompts investigate SE, we developed a codebook to assess how their content and structure align with theoretical foundations, specifically the sources and domains of SE. This tool enables a systematic analysis of the prompts themselves, revealing patterns in how they were designed to elicit real-time reflections on SE. The codebook also facilitates a broader conversation about the advantages of the journal prompts, particularly their potential to complement other instruments, in our case the ESM surveys, in capturing both the domain-specific nature of SE and offering additional insight into its source-based dimensions. The guiding research questions for this study are:

1. How can we intentionally design prompts to explore the theoretical foundations that underpin SE (e.g., domains and sources)?

2. In what ways do the individualized daily journal prompts capture the theoretical foundations that underpin SE?

3. In what ways do the affordances of daily journal prompts enable deeper exploration of the sources and domain-specific nature of SE?

In the following section, we will review the relevant background literature on using journals as a tool for qualitative data collection. We then outline the data collected for this work as well

155

as the overall design and implementation of the individualized daily journal prompts. Specifically, we will highlight the reasoning for the development of a codebook, leading to the next section, which will summarize the development process. Here, we will discuss how the codebook was developed, justifications for the structure of the codebook, and the inter-rater (IRR) reliability process. Furthermore, we will present the codebook by showing its application to a few specific prompts, demonstrating how the codebook illuminates differences between the types of prompts designed across the two iterations of the data collection. Lastly, we will address the limitations of the current codebook, and suggest future directions for improving the design of the individualized daily journal prompts to support capturing the nuances of SE.

6.2 Background Literature & Purpose

6.2.1 Expanding Reflective Journaling: From Learning to Affective Experiences

Broadly, reflective journaling is a process in which individuals document their thoughts, feelings, and experiences, often with the intention of gaining a deeper understanding of the individuals and their emotional and behavioral responses (REF). The practice of reflective journaling has been recognized in various fields, from education to X (REF).

Although the core concept of reflective journaling remains the same - engaging with personal experiences through writing - different fields and individuals may approach the practice in distinct ways. Some may focus on affective or emotional exploration (REF), while others might emphasize cognitive reflection or creative reflection (REF). Within one study they used reflective journaling as a means of collecting qualitative data and conceptualized this as an activity "describing a recent experience and unpacking salient aspects (e.g., people, resources, activities) that affected learning, and doing so in an ongoing manner over time" (REF).

This definition of reflective journaling emphasizes learning, aligning with its established value as a pedagogical tool that nurtures student learning (REF). It facilitates self-awareness and personal growth by encouraging people to document and reflect on their experiences over time (REF). This practice not only helps to understand one's learning journey but also contributes to the development of critical thinking and reflective skills (REF or own argument).

However, we know that for students to succeed in higher education, and more specifically in the sciences, their SE plays a critical role [51, 88–90]. Thus, I have expanded the definition of reflective journaling to include not only learning but also motivational constructs, such as SE. The aim of this broader definition is to allow for a most holistic reflection that goes beyond learning processes and incorporates the emotional and psychological elements that shape student experiences.

Even with this expanded definition, reflective journaling ultimately focuses on the process of describing and identifying important elements of experiences. When using journal prompts as a qualitative research method, the goal is often to encourage participants to do more than simply recall their experiences; rather, it invites them to reflect on and analyze their interpretations or communicate their perceptions of those experiences (REF). Journals also allow for the exploration of experiences at a different scope, as they reduce the time between when participants experience an event and they document it (REF). This suggests that not all experiences need to be documented, but rather key moments that carry greater significance in the reflection process. It is these moments and their salient aspects that we care most about in the research process.

Thus, I have further modified the definition of reflective journaling for our work to be the following:

"Reflective journaling is the practice of systematically describing and analyzing personal experiences, with particular attention to key moments, while unpacking salient aspects (e.g., people, resources, activities) that influence learning and other motivational constructs factors in an ongoing manner over time. "

Building on this understanding of reflective journaling, we now turn to its specific role as a research method, highlighting its distinctions, strengths, and applications in qualitative inquiry.

6.2.2 Journaling in Research: Distinctions, Strengths, and Applications

Journaling offers distinct advantages over other qualitative methods such as interviews and observational methods, making it a valuable qualitative tool for capturing participants' experiences. Unlike interviews, which rely on retrospective recall, journaling allows participants to document

their thoughts and experiences either in the moment or closer to the moment, reducing memory bias and preserving details that might otherwise be lost (Bolger et al., 2003; Monk et al., 2015; McCombie et al., 2024; Bartlett and Milligan, 2015; Mackrill, 2009). While journals excel as capturing real-time insights, they can also serve as a reflective tool for participants to explore past experiences, making them versatile for both immediate and retrospective data collection (Mackrill, 2007). Furthermore, one of the key benefits to journaling is its flexibility, allowing participants to document their experiences, thoughts, and feelings at their own pace and in their own words, in contrast to structured data collection methods such as interviews, focus groups, or questionnaires (Bijoux and Myers, 2006). Compared to observations, journaling grants access to participants' internal thoughts and emotions, perspectives that may not be externally observable. Observations also require the researcher's presence, whereas journals do not, making them a practical and unobtrusive method for gathering rich, contextualized data over time [111].

Journaling is also recognized for its ability to capture rich, real-time data about participants' experiences and phenomena, directly from the participants' perspectives (Lutz and Paretti, 2019; McCombie et al., 2024; Rudrum et al., 2022). One key advantage of this method is its ability to document experiences as they occur, offering valuable insights into day-to-day life. Researchers have highlighted that journaling is particularly useful for capturing micro-level experiences - those subtle, everyday events that might otherwise be overlooked (Kaun, 2010; Lutz and Paretti, 2019).

Journals enable researchers to collect rich, contextual data on participants' daily lives while centering their voices and perspectives. Journal prompts can be adapted and refined over time, allowing them to evolve in response to the emerging themes and events in the data (Rudrum et al., 2022; Bolger et al., 2003). This flexibility adds to the depth and relevance of the data collected, ensuring that it remains closely aligned with the participants' lived experiences.

In research, the ESM is commonly used as a method for collecting data closer to when participants' experiences occur. Some scholars have specifically linked journals to ESM (REF). While journaling can be integrated into an ESM study, we intentionally use the term "journals" to clarify that, in our context, journaling serves as a complementary data source rather than a component of

the ESM itself. Although both approaches involve repeated data collection, the focus of ESM is typically quantitative statistical analysis, while journaling focuses on qualitative insights (McCombie et al., 2024).

Journaling's combination of flexibility, immediate reflections, and its longitudinal capacity makes it a powerful complement to other qualitative methods in research. Integrating journals with other data collection methods is common practice (McCombie et al., 2024). For example, journals have been used to triangulate findings when paired with interview data, adding credibility to the results (Clayton et al., 2018; Kragh et al., 2017; Voriadaki et al., 2015). Others have used journals to generate topics that inform interviews (Denno et al., 2021; Deslandes et al., 2015; Long et al., 2016). When combined with other methods, journals take advantage of their unique strengths. To illustrate, journals can be used to elicit rich and immediate reflections, while interviews can be used to further investigate and generate discussion (Craig et al., 2017). As such, journals can expand on and help validate findings from other research methods.

Within physics education research, there appears to be limited use of journal reflections to explore affective variables such as SE. One study discussed using prompts on the students' weekly quizzes to probe their perceptions of their abilities (Stoeckel and Roehrig, 2021). The specific prompt was: "If you believe you are at mastery, what strategies or actions helped? If not, what will you do to improve?" This prompt explores the perception of students about their abilities and asks them to weigh their confidence in their abilities by describing their mastery or how to gain mastery. This particular prompt focuses primarily on mastery experiences, as their prompt encourages students to reflect on experiences that made them feel more or less confident in their mastery of the content. This prompt, however, may or may not capture other kinds of experience that students link to the mastery of the content. For example, how their experiences impacted their SE towards taking the quiz. Here, we extend on their work to explore how journal prompts can be expanded and refined beyond the source of mastery to capture a broader range of experiences.

159

6.2.3 Using Journal Reflections to Explore Self-Efficacy

Many have recognized that qualitative and quantitative research have different strengths and weaknesses (Merriam and Tisdell, 2016; Maxwell, 2013); as such, commonly using one method over another can limit our understanding of SE (Usher and Parajes, 2009; Webb-Williams, 2017). Qualitative research strengths lie in its process orientation towards the world, in which it is used to explore the people, situations, events, and process connecting these phenomena (Maxwell, 2013). This includes exploring how people interpret their experiences and how and what they attribute meaning to (Merriam and Tisdell, 2016).

SE has traditionally been studied using quantitative methods (Klassen and Usher, 2010; Tschannen-Moran and Woolfolk Hoy, 2001; Webb-Williams, 2017; MORE REF). However, many have discussed the benefits of leveraging qualitative and mixed methods for SE research (Zeldin and Pajares, 2000; Usher, 2009; Klassen and Usher, 2010; Chen and Usher, 2013, add PER REF). For example, in a study leveraging interviews, Usher (2009) claimed that "an interview approach offers a phenomenological lens through which the development of efficacy beliefs can be viewed and enables researchers to examine the different conditions under which students process and appraise their experiences as particular junctures in their schooling." However, Webb-Williams (2017) countered that this approach offers a smaller sample size and lower generalizability of the findings.

While many studies have leveraged interviews as the means to collect qualitative data about SE, few studies have used journal reflections to explore students' SE (Huang et al., 2014; Fong and Krause, 2014). Within the context of physics education research, it is especially true that journals are rarely used to explore students' SE (REF). As such, we are interested in leveraging journals as a viable method for capturing students' SE.

6.2.4 Summary of the Purpose of the Study

This work is critical in exploring how to design individualized daily journal prompts to elicit responses from students regarding their SE in relation to their day-to-day lives. While the ESM survey provides structured, momentary snapshots of SE, the individualized daily journal prompts offer an opportunity to explore students' experiences in greater depth. To systematically examine

the prompts, we developed a codebook to analyze the structure and content of the prompts in both iterations of data collection. The following sections summarize the development process; specifically, outline the data collection across the two iterations, the design of the individualized daily journal prompts, and then describes how the codebook was developed and used to demonstrate the role of the individualized daily journal prompts in exploring the theoretical foundations that underpin SE.

6.3   Methods

6.3.1   Context of the Larger Study

This study is situated within the broader investigation of SE in academic contexts, with a particular focus on how students' SE fluctuates in real time in response to their daily experiences. The larger study employs a mixed methods approach to explore SE dynamics using both the ESM (via survey data) and daily journal reflections. The primary objective was to understand how students' beliefs in their abilities evolve over time and in response to specific academic tasks, environments, and interactions. Within this larger study, individualized daily journal prompts serve as the key qualitative component, providing deeper insights into students' real-time experiences beyond what is captured in the ESM survey responses.

6.3.2   Mixed Methods Approach to Explore Self-Efficacy in Real-Time

To investigate the dynamic nature of SE, we employed an explanatory sequential mixed methods design, where quantitative ESM survey data informs the qualitative daily journal reflections. This approach allows us to identify trends in students' SE through the survey responses and then use journal prompts to further explore or explain these patterns.

The use of our mixed methods is particularly well suited for capturing SE in real-time because it enables the study of momentary fluctuations (via ESM surveys) and contextualized reflections (via daily journals). The ESM surveys provide real-time snapshots of SE tied to specific academic tasks, while the journal prompts allow students to elaborate on their experiences, challenges, and interpretations of those moments.

Across both the Fall 2020 and Fall 2021 iterations, the integration of these data sources was

critical to examine how students interpret and articulate their SE experiences in real time. However, the methodology evolved in Fall 2021, incorporating students' own language from their open-ended survey responses to further tailor the journal prompts and expand on the fluctuates in their task-specific SE scores we were observing. This refinement represents an additional point of integration and was intended to enhance the ability to capture the nuances of SE (see Chapter 5).

6.3.3 Data Collection

In Fall 2020, data collection involved administering ESM surveys multiple times per day alongside daily journal reflections. The students completed the ESM survey four times a day randomly within four time windows between 9 AM and 6 PM. The ESM surveys measured domain-specific SE (i.e., task-specific, course-level, and career-level), capturing students' perceptions of their confidence in their abilities in the moment.

After 6 PM, individualized daily journal prompts were sent to students. The journal prompts in this iteration were designed primarily based on students' task-specific SE scores, if there were any observable fluctuations in their scores. This allowed us to explore the fluctuations in their task-specific SE within their daily experiences. The prompts aimed to elicit reflections that expanded on how or why students rated their task-specific SE in this way.

A key feature of the Fall 2020 design was that while some journal prompts were individualized based on students' survey responses, others were more general. This provided a foundational understanding of how journal prompts could be used to explore SE, but also highlighted the potential for deeper integration between survey data and qualitative reflections. In addition, some prompts were written to further explore their SE in relation to their previous daily journal prompt responses if the researcher had more questions about their daily journal prompt response.

The Fall 2021 data collection maintained the structure of the ESM surveys and daily journal prompts, but introduced a key methodological refinement: the use of the student's own language from their open-ended survey responses to examine the ESM task-specific SE scores and was used to shape the journal prompts.

This added point of integration allowed for greater individualization of the journal prompts. In

Fall 2021, the prompts were not only tailored based on SE scores, but also crafted using the students' own wording from their open-ended survey responses, making them more personally relevant. By incorporating students' language, we bridged the gap between their real-time survey responses and their journal reflections. For example, if a student reported working on a physics assignment in a survey, we would craft a prompt specifically about their experience with that assignment, prompting them to reflect on their SE related to that task. This connection encouraged students to elaborate on specific moments they reported in their ESM surveys and explore how those moments influenced their confidence in a particular context. This methodological evolution reflects our ongoing effort to strengthen the alignment between ESM data and journal reflections, ultimately improving the mixed methods approach to exploring SE in real time.

6.3.4 Individualized Daily Journal Prompt Implementation and Design

Journals can take many formats, including video, audio, written responses, photos, drawings, or artwork; however, many studies overlook discussing the pros and cons of different journal formats, as well as providing justification for their choices (McCombie et al., 2024). This gap in the literature highlighted the importance of explaining our own decisions, not only to evaluate how the iterations using this mixed methods approach was conducted but also to ensure the study planning addressed the diverse needs of the participants.

6.3.4.1 Implementation

For our study, the individualized journal prompts were placed and delivered in Microsoft OneNote (REF), as it was a widely used application at the participants' home university. By selecting Microsoft OneNote for its accessibility and ease of use, we aimed to balance participant preferences with practical considerations, such as ease of access and mitigating the burden of participation. This approach aimed to ensure that the format aligned with both the technological context of the participants' university and their personal engagement with these iterations employing our mixed methods approach.

Each student had their own notebook within OneNote to keep their data private, shared only with the research team. OneNote was also selected for its accessibility on phones and computers,

allowing students to complete prompts without needing to carry additional equipment. This approach also ensured an electronic data collection, reducing the risk of lost paper diaries or forgotten entries. Research suggests that participants tend to engage more effectively with electronic data collection methods, particularly when the devices they already own are used, leading to higher levels of participation and completion (McCombie et al., 2024).

Across both iterations, participants were instructed to choose their preferred format for recording their journal responses. Specifically, at the top of each delivered daily journal prompt, we told the participants: "For each daily reflection you can type, write with a stylus, or record your voice. You may choose to reflect on any part of your day today that you wish. Below are prompts to inspire the kinds of things you might reflect on." Students normally chose to respond to the prompts by typing their responses to the prompts. We posit that students likely opted to type their responses because the individualized daily journal prompts were presented in typed form within their notebooks. Additionally, these directions gave the students autonomy in that they could choose to reflect on their day with or without using the prompts. Students normally choose to respond to the prompts we wrote.

An important consideration to our research design was the potential burden on participants. To minimize this burden, participants were only required to complete one journal reflection per day, rather than multiple. By structuring the study in this way, we were able to ensure that participants could engage meaningfully without facing excessive strain, while also preserving the temporal aspect needed to capture the dynamics of SE (2020 perc paper REF). Given that SE is understood to be dynamic, it was essential to explore how it fluctuates over time without overwhelming participants with excessive demands on their time and energy.

6.3.4.2 Design

Designing effective journal prompts for qualitative research can be challenging. If prompts are too open-ended, participants may provide responses that stray from the research focus, potentially offering irrelevant data (Graneheim and Astro, 2016). Conversely, overly restrictive prompts can limit the depth and authenticity of participants' reflections (Gill et al., 2016; Rudrum et al.,

2022). This latter scenario diminishes the exploratory potential of qualitative methods, which goes against our goal of using the individualized daily journal prompts to support the exploration of the ESM survey data and students' real-time experiences related to their SE. Therefore, it is crucial that the design of the journal prompts within our mixed methods approach is intentional and balanced. Failing to disclose how these prompts were designed or how participants were guided could undermine the trustworthiness of the study's findings (McCombie et al., 2024).

We adapted interview protocols to support writing some of the daily journal prompts; however, the majority of the individualized daily journal prompts are written to expand or confirm/disconfirm the ESM survey data; specifically, in response to events and emerging themes from their ESM data. However, this raises many questions about how to write individualized prompts, how to write prompts that complement the ESM survey data, and how to write prompts that elicit information about how their real-time experiences are influenced by or influencing SE. In the following, we discuss key considerations for intentionally writing prompts to meet the goal of tailoring prompts to the individual, complementing the ESM data, and eliciting reflections on SE.

Tailoring Prompts to the Individual: One key consideration was how to create individualized journal prompts that responded to each student's unique experiences. This personalized approach was essential for capturing specific moments in students' daily lives, making their reflections more meaningful and directly relevant to their SE.

In both data collection iterations, one approach involved customizing prompts based on students' task-specific SE scores from the ESM data. Depending on the presence of threads (see Chapter 5) from the ESM data, we also paid close attention to the language students used in their previous daily journal prompt responses, using their specific wording to shape subsequent prompts. In the Fall 2021 iteration, we also incorporated students' own language from their open-ended ESM survey responses when writing prompts, further aligning them with the participants' experiences and perspectives.

Complementing ESM Data: In addition to individualization, it was crucial that the journal prompts complemented the ESM data. This meant designing prompts that were both consistent

with the data being collected through the ESM surveys and responsive to the unique dynamics observed in students' responses. The goal was to ensure that the journal prompts provided an opportunity to expand upon or confirm/disconfirm the ESM data, creating a more holistic view of participants' SE.

For example, consider the following individualized daily journal prompt from a student in the Fall 2021 data collection:

On Friday, we noticed that you spent most of your day doing the [computational modeling course] homework. We noticed an upward trend in how skilled and successful you felt when you were working on that homework.

- *Can you take us back to Friday and discuss what was happening while you were working on your [computational modeling course] homework?*

- *How confident were you in doing this [computational modeling course] homework?*

- *How did that homework impact your ability to succeed in your [computational modeling course] course?*

This prompt was informed by the ESM data in several key ways. First, our awareness of the student's activity across the day—working on their computational modeling homework—came directly from their open-ended responses across multiple ESM surveys. Second, the observation of an upward trend in their task-specific SE (i.e., "skill" and "success" ratings) was derived from plotting their raw ESM scores and noticing a positive shift over time. Third, we intentionally echoed the language used in the ESM survey by referencing "skilled" and "successful" within the journal prompt. This intentional language mirroring further embedded the prompt in the theoretical and methodological framing of the ESM instrument.

Turning to the questions within the prompt, each serves a specific function in expanding upon the ESM data. The first question—"Can you take us back to Friday and discuss what was happening while you were working on your [computational modeling course] homework?"—invites

166

the participant to provide a rich, contextual narrative that extends beyond what is possible to capture through the ESM survey alone. Our use of the ESM did not include space for detailed, moment-by-moment descriptions, so this prompt question provides a valuable opportunity to explore those details qualitatively.

The second question—"How confident were you in doing this homework?"—offers the student space to affirm or complicate the upward trend we observed in their task-specific SE scores. While we might expect their response to align with the ESM data, the open-ended nature of the question preserves the participant's agency to disagree with our interpretation or nuance it in ways that are not visible through quantitative data alone. In this way, the prompt allows us to validate—or question—patterns seen in the ESM data.

Finally, the third question—"How did that homework impact your ability to succeed in your course?"—helps us begin to explore the relationship between task-level and course-level SE. This is significant, as it offers a pathway into understanding how localized experiences (e.g., working on a specific assignment) may shape, or be shaped by, broader perceptions of competence and success in the course as a whole. While such relationships could be investigated quantitatively by linking task-specific and course-level SE items within the ESM survey, the real-time nature of our data collection process limited our capacity to do so analytically during the study. Instead, the journal prompts served as a real-time tool for probing these theoretical connections qualitatively, as researchers were simultaneously analyzing ESM data and crafting individualized prompts for participants on a daily basis.

Eliciting Reflections on Self-Efficacy: Lastly, the journal prompts needed to elicit reflections on how students' real-time experiences were related to their SE. We aimed to craft prompts that encouraged students to connect their day-to-day experiences to their beliefs in their capabilities, ensuring that the prompts facilitated meaningful insights into the dynamics of SE over time. This was particularly important given the dynamic nature of SE and the need for prompts that helped capture how students' SE fluctuated in response to different academic challenges.

By addressing these considerations, we sought to balance structured prompts with the flexibility

needed for meaningful reflection. Since our primary aim was to elicit insights into students' SE, this particular work focuses on the development and use of a codebook to examine how the individualized daily journal prompts captured the nuances of SE.

6.4 The Creation of the Codebook

In this section, we detail the development of two distinct coding schemes: Contexts and Experiences. First, we outline the Context Codes, which were constructed based on predefined domains relevant to our study but required further refinement to develop detailed coding descriptions. Next, we describe the data-driven process used to develop the Experiences codes, which emerged from examining the daily journal prompts through the lens of SE. We then justify the overall codebook structure, demonstrating how both coding schemes align with the theoretical foundations and complexities of SE. Finally, we discuss the iterative nature of the codebook development, emphasizing how the inter-rater reliability (IRR) process informed the refinement and finalization of the codes.

6.4.1 Methodological Influences on Codebook Development

In developing the coding scheme for this work, I drew on the methodological literature that centers not only the mechanics of qualitative coding but also the interpretative and theoretical decisions embedded in the process. Scholars offered foundational insights into how codes serve as both analytic tools and meaning-making devices (Saldaña, 2016; Miles and Huberman, 1994; Strauss and Corbin, 1998). Saldaña's framing of coding as a cyclical, iterative process resonated particularly well with my approach, as I often returned to earlier coding decisions after new insights emerged through group discussions or data immersion. On the other hand, Strauss and Corbin's open coding and constant comparison was useful for considering how ensure that the codes were distinct from each other.

Merriam and Tisdell claimed that when considering qualitative coding, the categories, themes, and findings must be responsive to the research questions, guiding me to expand upon using key words, phrases or sentences when thinking about how the codes aligned with the theoretical foundations of SE (e.g., domains and sources). Additionally, they suggested that codes or categories should be mutually exclusive in that a relevant unit of data can be placed in only one category. In

practice for the purposes of this codebook, a unit of data may be coded with multiple codes; yet the reasoning behind why multiple codes can be applied to the same unit of data is distinct from one another. This tension pushed me to think more carefully about the interpretative logic behind the application of each code and encouraged me to engage in inter-rater reliability (IRR) as a way to surface and negotiate these distinctions.

In addition to the historic literature, I want to honor the more recent work of Laura Wood () whose approach to using "indicator words" influenced how I began thinking about the coding process. This encouraged me to start by analyzing the daily journal prompts for key words or phrases while also reflecting on how particular words or phrases might serve as meaningful entry points into students' experiences (Wood, 2022).

Additionally, I found resonance with the epistemological stance offered by researchers such as Stephen Secules and Angela Calabrese Barton, who emphasize the cultural and contextual framing of STEM education research (Barton et al., 2017). Their work reminded me that building a codebook is not a neutral act - it reflects which narratives are being legitimized and whose language is being prioritized. This shaped my decision to anchor codes in the students' own phrasing when possible, and to design categories (like "Interactions with Others") that left room for nuance rather than preimposing theory.

Importantly, I also made a deliberate effort to stay close to the data because these individualized daily journal prompts were informed by the students' ESM survey responses or previous prompt responses. Honoring the students' data meant recognizing that even the prompts themselves carried traces of student voice. However, it is equally important to acknowledge that these prompts were also designed through the lens of the researchers - shaped by our theoretical commitments, interpretations, and decisions about what to ask and how to ask it. This dual nature of the prompts - as both student-informed and researcher-framed - made it all the more essential to approach coding with care, striving to interpret them in ways that remained faithful to both the students' language, the lens of the researchers, and the broader intentions of the research design. In this way, the process of developing the codebook was both a theoretical and ethical endeavor — one that sought

to explore SE through individualized daily journal prompts while remaining attentive to the layered dynamics of data construction and interpretation.

The literature also made space for me to treat disagreements in coding as productive (REF). Rather than seeking rigid inter-rater reliability (IRR) as a measure of objectivity, I followed the guidance of a qualitative piece that discussed the advantages of treating coding discrepancies as opportunities to deepen interpretive alignment (REF). This thinking shaped our IRR process, where discussions about disagreements led to refinements in the code descriptions and strengthened the conceptual clarity of the codebook but also expanded on the ways in which we can use prompts to capture the theoretical foundations of SE.

These methodological perspectives ultimately shaped not only *what* I coded for, but *how* I approached coding as a design and theory-driven act. In terms of thinking about how to communicate my decisions and the codebook itself, I lean heavily on the work done by Little et al. (REF), where they did a beautiful job discussing the development of their codes and the decisions they made (REF). In the following sections, I describe how these influences materialized in the Context and Experiences coding schemes, and how they contributed to a codebook that reflects and extends key ideas from SE theory.

6.4.2   Examining Evidence for Domains

One of the overall aims of the mixed methods approach is to explore students' SE across different domains - task, course, degree, and career. The SE literature provided us with valuable insights into identifying the different contexts within the individualized daily journal prompts. Although domain-specific definitions of SE vary between studies, common themes emerged. For example, some studies focused on SE based on subject (e.g., math or science SE), while others examined discipline-specific SE (e.g., physics or biology SE). Some literature even addressed task-specific, course-level, degree-level, career-level, and transfer SE (REF).

While the ESM survey items directly measured SE in the task, course, and career domains, the individualized daily journal prompts were designed both to expand upon survey responses and to capture additional information. As such, I initially examined the prompts for references to these

domains to create codes that captured each of these settings. However, a key decision that was a result of collaborative discussion was to distinguish between 'Context' and 'Domains' to ensure that individualized daily journal promptss were coded in a way that aligned with our study's goals. As such, we defined 'Context' more broadly to capture relevant experiences, even if SE was not explicitly mentioned within the individualized daily journal prompts. This decision allowed us to systematically track moments related to SE without imposing a strict requirement for explicit SE language. By structuring the codes in this way, we ensure that the individualized daily journal prompts complemented the ESM survey responses while also allowing space to examine in-the-moment experiences that contribute to the students' SE. Additionally, while some individualized daily journal prompts used explicit SE language, others intentionally did not. Our goal was to strike a balance, carefully framing questions to capture students' experiences without unintentionally leading their responses.

6.4.2.1 An Example of the Context Codes Development Process

When examining individualized daily journal prompts, the most straightforward indicator was the explicit mention of a course label (e.g., PHY 123). However, a pause in my thinking led me to ask myself: *What about cases where an individualized daily journal prompt referenced coursework without specifying a particular course? If a prompt referenced a subject area (e.g., 'a physics course' or 'a mathematics course') without specifying a course number, should we still apply the course code?* These questions underscored the need for clear and consistent coding decisions to ensure alignment between our coding framework and the theoretical foundation of domain-specific SE.

To illustrate how the Contexts codes were designed to reflect domains, we present the development of the COURSE code as an example. Course-level SE is defined as one's confidence in their ability to pass a specific STEM course (Stewart et al., 2020?). Guided by this definition, the COURSE code was designed to capture individualized daily journal prompts that referenced students' experiences, abilities, or confidence related to passing a specific course. While the COURSE code aligns with course-level SE by requiring a reference to a specific course, it intentionally does

not require explicit statements about confidence in ability, distinguishing it from the course domain of SE. For example, a prompt identified from the Fall 2021 iteration is provided.

From your app responses today it sounded like you were able to make it to the office hours for [modern physics course label]. We're curious to hear about how it went?

- *Who was there? What were you doing?*

- *How did it impact your sense of your abilities for [modern physics course label]*

This prompt was coded with the COURSE code because of the explicit mention of the [modern physics] course. However, none of the prompt questions directly reference SE. For instance, the final question - "How did it impact your sense of your abilities for [modern physics]?" —indirectly engages with SE by prompting the participant to reflect on their abilities, but it stops short of using explicit SE language such as "confidence," "belief in your ability," or "ability to succeed in this course." It also avoids more direct framings, such as: "After leaving office hours, how confident are you that you can complete the required activities in [modern physics]?" or "Considering how office hours went, do you feel confident you can pass this course?"

This distinction is meaningful: while the prompt encourages reflection on ability within a specific course context, it leaves space for the student to define whether—and how—that experience influenced their SE. As such, it illustrates how a prompt can be categorized under COURSE without explicitly probing course-level SE, thus offering a more open-ended, student-centered opportunity for self-reflection.

Notably, this is not to say that the prompt is not exploring SE, as the final question in the prompt invites participants to reflect on how attending the office hours influenced their perceived capabilities in the course. Using a more open-ended and student-centered framing, the prompt creates space for participants to share whether and how the experience shaped their SE, without prescribing a particular response or using technical SE terminology. In this way, the prompt is intentionally constructed to explore SE — even if it does so through more implicit means. This

example demonstrates how a prompt can be coded with COURSE and still provide meaningful insight into SE, even without using the formal language typically associated with SE theory.

6.4.2.2   The Context Codes Descriptions

In Table 6.1, we present the Context codes alongside brief descriptions. A single individualized daily journal prompt can be assigned multiple Context codes; however, the rationale for each code assignment differs. For example, a prompt may be coded as both TASK and COURSE if it references "working on a homework assignment in PHY123."

Code	Brief Description
Task	A daily journal prompt that discusses a specific task. A task is any activity/assignment AND/OR the act of doing a particular activity/assignment. Within this description of task, we include the act of attending academic spaces. This code is utilized when a specific task or multiple specific tasks are focused on within the daily journal prompt.
Course	A daily journal prompt that discusses a particular course or identifies the particular course or courses. It can pinpoint a particular course using language as broad as "particular" or as specific as having a course label within the daily journal prompt. Ideally, the participant is being asked to discuss the particular course or each course provided in such a way that a distinction between which course the participant is focusing on is clear.
Degree	A daily journal prompt that discusses degree or major, or the requirements or pursuing a degree or academic plan. As an important note then, in this work we are using the words degree and major interchangeably. We also use these words interchangeably with one another: "requirements of pursuing a degree," "academic plan," "degree plan," and "4-year plan."
Career	A daily journal prompts that discusses one's intended/chosen/future profession/career/occupation, or post-graduation/career plans. As an important note then, in this work we are using the words profession, career, and occupation interchangeably.

Table 6.1 Overview of Context codes used in the analysis. Each code mirrors a distinct academic level or domain referenced in students' daily journal prompts—ranging from specific tasks to broader academic or career goals.

These brief descriptions guide how prompts are initially coded. The full codebook definitions, as well as corresponding examples of individualized daily journal prompts that do and do not fit each Context category, are included in the Appendix B. The examples that align with each Context code serve as evidence supporting the development of the codes themselves and clarify how these categories are applied in practice—including edge cases where multiple Contexts codes may be

assigned.

6.4.3 Examining Evidence for Experiences

When I began developing the codes to examine for evidence of experiences within the daily journal prompts, I began with characterizing the Fall 2021 data. The opportunity of having been a part of the data collection for Fall 2021, meant that my coding was influenced through the lens in which I had (co)written many of these daily journal prompts; for example, we had written them with the aim of further exploring the real time moments students had reported on the ESM surveys in relation to their SE.

Through this lens, there were two key components I wanted to characterize. First, some of these prompts were written to capture students' real-time experiences as they reported them on the ESM surveys. As such, I came in with familiarity toward the key words, phrases, and sentences we used to this work during the data collection process. For example, if we wanted to further explore a moment the student had experienced earlier in the day, then we asked them for more context about this moment such as "what were they working", "who they were working with," and "how they felt in that moment." I created a code to be able to capture this work that guided the students to share more about about a particular moment.

Another aspect of the daily journal prompts I wanted to capture with my coding was that the aim of the prompts was to further explore the students' SE. Again, from having participated in the Fall 2021 data collection, I had the knowledge about the various ways in which we wrote the prompts to elicit SE. For example, one way was through asking about their SE explicitly by providing them with questions around what capabilities/abilities they were confident in within certain contexts. In other cases, we explored the sources of SE through writing prompts using language around working with others (i.e., vicarious learning and social persuasion), comparing themselves to others (i.e., vicarious learning), and receiving encouraging or discouraging messages from others (i.e., social persuasion). As such, I wanted codes that captured the language that prompted the students to discuss the sources of their SE.

6.4.3.1 An Example of the Experiences Codes Development Process

In general, the initial phase of my process to detect evidence of SE experiences, I began by identifying language that could signal opportunities to engage with the sources of SE (REF). For example, being encouraged to pursue a physics major could represent a social persuasion experience—an interaction that may contribute to a student's SE related to their physics degree. However, the impact of this moment depends heavily on who delivered the encouragement; the source's credibility or relational importance may influence whether the experience functions as a meaningful source of SE for the student.

With this in mind, my first analytic step involved identifying any information that could be linked to the sources of SE. In the case of social persuasion and vicarious learning, a participant would, at a minimum, need to reference an interaction with another person. At a more specific level, I drew on indicator words from Wood (2022)—such as encouragement or discouragement—which could signal moments of social persuasion. When a daily journal prompt elicited or incorporated this kind of language, it was marked with a corresponding code to reflect that potential SE opportunity. Once I had done this, I began the process of going the codes into groups. For example, I highlighted prompts that included terms like "mentor", "network", "group work session with a friend on homework," or "conflict with mom." While identifying these words and phrases were useful in understanding who was being referred to in relation to the participant, I quickly realized that highlighting them alone wasn't sufficient for addressing the deeper goals of the project. These terms remained authentic to the language we used in the prompts, but on their own, they did not necessarily connect back to SE - the theoretical backbone of our work.

This led me to pause and ask myself: *Why am I highlighting these words? What do they actually tell me about the participant's SE?* As I revisited the theoretical grounding of our prompts, I recalled that we had intentionally included references to social relationships because were were interested in capturing social persuasion and vicarious learning - two sources of SE. However, not every mention of another person automatically conveyed an influence on the participant's SE.

Some prompts were more explicit, encouraging participants to reflect directly on how others

influenced their sense of ability. But even when prompts did not make that connection overt, I wanted to acknowledge that the participants may still give us meaningful insights into how participants thought about their interactions with others and how these interactions, in turn, shaped their perceptions of their own competence and abilities. For instance, we wrote a prompt about a "group work session with a friend on a homework" as a result of this being in their open-ended ESM survey response. In thinking about this kind of ESM response, if we write a broader prompt with no specific words around their confidence in relation to that experience and its' impact toward their SE toward the homework, the participant still may share with us that the session supported them in being able to complete the homework assignment, in which that phrase "being able to do the homework" - would signal a shift in their perceived capability in relation to the homework; thereby, their SE.

This analytical unpacking led to the creation of the "Interactions with Others" code. This decision allowed us to systematically capture prompts that explored the role of social dynamics on students' SE, which was aligned with our broader aim to explore how experiences in real time may support—or threaten—the SE as perceived by the student. For example, consider the following prompt from the Fall 2021 iteration of the mixed methods approach:

> *Reflecting over this past week, were there any people or experiences with people who encouraged or discouraged you in a particular course you are in right now? Who are they? What are the things they say to you that either encourage or discourage you in the course?*

This more general version of a daily journal prompt invited participants to reflect on the influence of others over the past week, specifically in terms of encouragement or discouragement they perceived receiving within a particular course. The structure and language of the prompt were intentionally designed to surface potential social persuasion moments—one of the core sources of SE identified by Bandura — by prompting participants to recall specific interactions in which others either encouraged or discouraged them in the context of that course.

6.4.3.2   The Experiences Codes Descriptions

In Table 6.2, we present the Experiences codes alongside their corresponding brief descriptions. These codes represent the types of experiences that each individualized daily journal prompt is designed to elicit. Fuller descriptions for the Experiences part of the codebook and corresponding examples of prompts that do or do not align with each of the Experiences codes are included in the Appendix C. These examples serve as evidence supporting the development and application of the codes. Again, a single prompt may receive multiple Experiences codes depending on the dimensions of experience it is structured to surface. For example, a prompt may be coded as both Influences on Context-Specific Confidence/Success and Interactions with others, if the prompt was designed to explore how a participant's network influenced their confidence in passing their modern physics course, obtaining their degree, or pursuing their future career. The Experiences codes reflect the theoretical foundation - source nature - of SE.

The Appendix C has fuller descriptions including inclusion and exclusion criteria to support distinguishing between these codes. Similarly to the Contexts codes, the short descriptions for the Experiences codes guide how prompts are initially coded, while the fuller descriptions and examples in the appendix clarify how these categories are applied in practice - including edge cases where multiple Experiences codes may be assigned.

6.5   Justifying the Codebook Structure

The codebook was developed through the lens of SE to categorize and analyze the daily journal prompts in relation to experiences and contexts - both of which reflect, but are not the same as, the sources and domains of SE. The Experiences codes allowed me to capture the kinds of moments students encountered - what happened, who was involved, and how those moments may have influenced their beliefs about their capabilities. Meanwhile, the Contexts codes helped me situate those moments within meaningful domains such as task, course, degree, or career, providing insight into where and in what ways SE was being shaped.

This dual-layered coding approach was rooted in both theory and data. Bandura's emphasis on the domain-specific nature of SE guided the need for contextual clarity, while the social and

Code	Brief Description
Evaluation of Confidence, Skills, and Success	A daily journal prompts written to elicit students' reflections on their perceived confidence, perceived skills, and perceived success. It focuses on self-assessments of abilities, capabilities, and achievement, often reflecting on past experiences and moments that shape these perceptions over time.
Influences on Context-Specific Confidence/Success	A daily journal prompts written to cue participants to discuss how particular factors (internal or external) may influence the participant's confidence or success within a particular context.
Comparison to Peers	A daily journal prompts written to cue the participant to share comparisons between themselves and their peers.
Interactions with Others	A daily journal prompt written to cue the participant to discuss their engagement with others, in which others here can be a single person or multiple people. These people can range in their roles in the participant's life.
Details and/or Feelings Regarding an Event	A daily journal prompt written to have a participant reflect over their lived experiences. These prompts are built as an invitation to the participants to share more details and feelings about a moment, event, activity/activities, opportunity or experience the person had.

Table 6.2 Overview of the Experiences codes used in the analysis, designed to align with the primary sources of SE as identified by Bandura (REF). Each code captures a distinct aspect of students' experiences related to their SE development.

emotional nuances of students' lived experiences required an approach to coding what those experiences entailed. Ultimately, this interplay between Experiences and Contexts codes enables a more complete and dynamic understanding of how SE was activated, challenged, and supported across different academic spaces and interactions.

Importantly, these codes were not designed to map directly onto the four sources or domains of SE. Rather, they were developed to reflect the unique, situated ways SE is explored within

individualized daily journal prompts. This distinction - between reflecting SE theory and being bound by it - became central in shaping the overall structure of the codebook. In this section, we offer a deeper rationale for how and why the codebook was organized in this way, including how we intentionally differentiated between source-like and domain-like constructs while still remaining grounded in the theoretical foundations of SE.

6.5.1 Contexts (Investigating the Domain-Specific Nature of Self-Efficacy)

The Context section of the codebook was designed to categorize daily journal prompts in relation to the domain-specific nature of SE. SE is inherently domain-specific because the definition of "task" can vary depending on the scope of the study. For example, SE toward completing a homework assignment is different than mastering a specific skill like mathematical integration than passing a course.

Similar to the Experiences codes, the Context codes were not strictly tied to the domains of SE (e.g., task, course, degree, and career). This is intentional, as it allows us to capture the wide range of contexts in which students may have SE, even if the daily journal prompts do not explicitly reference domain-specific SE. For example, a prompt might explore a student's abilities within a course without using language like "confidence in their abilities within that course." As such, the purpose of the Context codes is to identify the relevant context, but this does not mean that students were specifically led to discuss their domain-specific SE.

To better illustrate how the Context codes align with SE, we will walk through one example in detail: Task. The example will highlight how a specific code aligns with the nature of domain-specific SE. The remainder of the Context codes are discussed in Appendix D.

Task. One's confidence in their abilities to complete a particular task is task-specific SE (REF). Tasks within the context of our codebook include typical academic activities such as particular homework assignments, quizzes, or midterms, as well as reading assignments. To reflect this domain-specific aspect of SE, I designed the *TASK* code to capture prompts focused on a particular task or activity. A key difference between the Task code and Task-Specific SE here is that the *TASK* code is not dependent on the prompt explicitly referencing confidence or SE. Many of the journal

prompts were written to probe moments where students engaged in specific academic tasks, but they do not always ask directly about their confidence in completing those tasks. For example, consider this prompt from the Fall 2021 iteration:

"How did working on this [computational modeling course] assignment affect your confidence within the course?"

- Daily Journal Prompt, Mruczek, Week 4, Day 2

While this prompt clearly references a task — working on an assignment in the course — it also implicitly touches on course-level SE by asking about their confidence in the course in relation to this task. In the case of this prompt, task-specific SE is being explored because it directly links the student's perception of completing an assignment to their confidence within the course context. However, not all *TASK* prompts are structured in this way. Some may simply ask students to reflect on the experience of completing an assignment without probing into their confidence.

This distinction is important because the *TASK* code allows us to capture moments where students engage with specific academic activities, even if the prompt does not directly ask about their SE. This is a deliberate choice in the coding process, as many journal prompts were designed to elicit in-the-moment reflections on students' academic tasks, leaving room for them to interpret the impact of these tasks on their own sense of competence. The decision to separate task-focused prompts from direct inquiries into domain-specific SE provides a foundation for future work to investigate how these task-based experiences might be linked to SE in ways not immediately apparent from the data itself.

In this way, the *TASK* code is reflective of task-specific SE, but does not require that the task itself be explicitly linked to confidence. This flexibility allows for a broader exploration of how tasks influence students' academic experiences and sets the stage for future analysis that may directly investigate the relationship between these tasks and SE.

In addition, we have included "attending spaces" as part of the *Task* code - environments where students engage with content or support structures, such as attending office hours, class sessions,

180

or other academic spaces. For example, consider this prompt from the Fall 2021 iteration:

> *"From your app responses today it sounded like you were able to make it to the office*
> *hours for [modern physics]. We're curious to hear about how it went?"*
> - Daily Journal Prompt, Jane, Week 2, Day 3

This prompt invites the student to reflect on how attending an academic support space (office hours) may influence their confidence in completing a related academic task (e.g., a physics assignment). Although the prompt does not directly ask about the student's SE, it engages with the context of the task — the action of seeking academic support — that may influence their task-specific SE. In this way, the Task code helps capture how students' interactions with academic support structures contribute to their evolving confidence in completing academic tasks, even if confidence is not explicitly mentioned.

Additionally, in other cases, attending spaces on campus often involves students seeking specific resources or support, which aligns with the Task code, as these actions are directly tied to completing a task. However, this does not mean that attending these spaces automatically influences a student's task-specific SE. While these spaces could impact SE across different domains, the Task code focuses on the task-centered aspect of the activity itself. For the purposes of this codebook, the Task code captures these moments as part of engaging with academic tasks, without assuming a direct link to SE. This distinction allows for future research to investigate how these contexts might relate to SE, particularly in task-specific domains.

6.5.2 Experiences (Investigating Source Nature of Self-Efficacy)

The Experiences section of the codebook was designed to categorize the daily journal prompts in relation to Bandura's SE theory, which emphasizes the role of various sources in shaping beliefs about one's abilities. The codes related to experiences in this section are intentionally flexible and not strictly tied to the four traditional sources of SE. This flexibility in the codebook allows for a more nuanced understanding of how SE is shaped by real-world situation-dependent experiences. For example, the code "Interactions with Others" can encompass both social persuasion and vicarious

learning, depending on the situation. These two sources are distinct, and "Interactions with Others" cannot be simply replaced with "Social Persuasion," as they represent different processes within Bandura's framework.

To better illustrate how the Experience codes flexibly align with SE theory, we will walk through one example in detail: Influences on Context-Specific Confidence/Success.

The example will highlight how a specific code aligns with key elements of SE theory and serves as a tool for understanding how experiences shape SE beliefs. The remainder of the Experience codes are discussed in detail in Appendix E.

Influences on Context-Specific Confidence/Success. This code was designed to capture individualized daily journal prompts that explicitly explore the interrelationships between various sources and domains of SE. We were interested in understanding how different domains and sources interact and how students' experiences impact their SE across these contexts. For example, consider this prompt from the Fall 2021 data collection:

> *"When you think about your network, how have interactions with those folks impacted your confidence in...*
>
> - *passing your courses (MTH123, PHY123, AST123)*
>
> - *obtaining a physics/astronomy degree*
>
> - *or, pursuing a career?*
>
> - Daily Journal Prompt, Mruczek, Week 4, Day 2

This prompt aligns with SE theory by investigating the role of social persuasion in shaping SE beliefs in different domains. Bandura (1997) identified social persuasion as one of the primary sources of SE, emphasizing that encouragement or feedback from others can significantly influence one's confidence (REF). In this case, the participant's reflection on how their network influences their confidence in their course, obtaining their degree, and pursuing their future intended career speak diretly to how social interactions are related to the domains of SE.

182

Incorporating such prompts enables educators and researchers to understand which sources of SE students perceive as influential across different domains. For instance, students might view social persuasion as crucial for career-level SE, but less so for degree-related SE. Identifying these perceptions allows for the design of targeted experiences that boost students' SE at various levels. This example demonstrates how we used prompts to explore the relationships between the source nature of SE and the domain-specific nature of SE.

Similarly, some prompts were written to explore students' perceptions of the relationships between the different domains of SE. For example, consider this prompt from the Fall 2021 iteration:

> *"How did working on this [computational modeling course] assignment affect your confidence within the course?*
> - Daily Journal Prompt, Mruczek, Week 4, Day 2

This prompt explores how students perceive the task - "working on this [computational modeling course] assignment" - and its impact on their confidence within the course. Here we treat the [computational modeling course] assignment as a task and a mastery experience, where students may reflect on their success or failure in completing it. This prompt then extends this reflection by asking how this mastery experience on this particular task relates to their confidence in the course, thereby exploring the relationship between task-specific SE and course-level SE as well as how these domains are related to a mastery experience.

As such, this prompt would be coded as "Influences on Confidence/Success in Specific Contexts" because it investigates how students assess the impact of a specific task on their belief in their ability to succeed in the course. By reflecting on such tasks, students provide insights into how mastery experiences shape their SE beliefs in academic contexts, and potentially how their task-specific SE and course-level SE are related. In addition, the open-ended nature of the prompt allows students to consider other sources of SE, such as social interactions that may occur while completing the assignment.

Incorporating prompts focused on the relationships between sources and domains of SE, and the relationships between the different domains of SE are critical to understanding how the different theoretical underpinnings of SE are interacting and can influence one another. This insight can inform the development of instructional strategies aimed at enhancing students' SE in different academic settings.

6.6 Inter-Rater Reliability and the Iterative Nature of Codebook Development

To ensure consistent application of our codebook, we conducted an inter-rater reliability (IRR) check. Beyond methodological rigor, IRR also played an epistemological role in establishing a shared interpretive grounding across coders, allowing us to reliably trace how journal prompts were written in alignment with the theoretical foundations of SE. The codebook itself was developed to categorize journal prompts in a way that helps us understand (1) how the prompts are designed to elicit reflections on SE, and (2) what aspects of SE are being foregrounded or potentially underexplored. Drawing on the work of scholars such as Stephen Secules and others who foreground cultural and contextual considerations in education research, we view this process as more than classification—it is also a reflection of which narratives of SE are being amplified or marginalized through design choices (REF).

Given the mixed methods nature of our study, IRR also served a crucial role in supporting the trustworthiness of our data integration. Specifically, we sought to ensure that the journal prompts were exploring dimensions of SE that align with constructs measured in the ESM survey, thereby enabling meaningful convergence between qualitative and quantitative strands. In later sections, we demonstrate how the codebook not only confirms this alignment but also reveals that the individualized daily journal prompts offer opportunities to explore theoretical dimensions of SE — particularly the source nature of SE — that the survey alone is limited in addressing.

We chose to calculate the percent agreement between two coders based on the nature of the claims we were making in this study. Our aim was not to quantify the frequency of particular categories, but to demonstrate how prompts were written to reflect and expand on the theoretical foundations of SE. Because this work is inherently exploratory and formative, percent agreement

184

was an appropriate and widely accepted method for assessing IRR in early-stage qualitative research (Miles and Huberman, 1994). Although percent agreement does not account for chance agreement—making it a less conservative measure—it was sufficient for our purposes given our emphasis on theoretical alignment and design-based inquiry. Thus, percent agreement served as a practical tool for demonstrating the trustworthiness of the coding and supporting the conceptual claims related to the design and theoretical grounding of the prompts. An ideal threshold for percent agreement is typically considered to be 70% or higher (REF).

6.6.1 Coded Datasets and IRR Process

We began the formal coding process with data from Fall 2021, anticipating that these journal entries would be richer and more varied due to our increased capacity to write individualized daily journal prompts during that iteration. Seven students were selected from the Fall 2021 dataset. The rationale behind their selection included:

- One student was frequently given customized prompts during data collection due to the richness of his prior responses, making his dataset especially nuanced and valuable for codebook development.

- Two students had participated in both Fall 2020 and Fall 2021 data collection efforts, offering continuity across the dataset.

- The remaining four students were selected for their unique prompt content, strong narrative arcs, and varied experiences.

After selecting the Fall 2021 students, we developed a preliminary codebook based on their daily journal prompts. Once this initial codebook was established and refined, we applied it to the Fall 2020 data, which included 5 students. This step allowed us to test the codebook's applicability to a similar yet distinct dataset. Although both sets of prompts were written for the same mixed methods research project, the data sources used to inform the prompts differed between Fall 2020 and Fall 2021, allowing us to evaluate how transferable the codes were across contexts.

Once coding was completed across both iterations, we finalized the codes and their descriptions and began preparing for inter-rater reliability (IRR) testing. This included:

- Writing a detailed coding protocol that specified how to apply each code and how to define the unit of analysis.

- Meeting with an experienced secondary coder to review the codebook and protocol. This meeting allowed the coder to ask clarifying questions and helped us determine which subset of data would be jointly coded for IRR.

6.6.1.1 Selecting Data for IRR

To determine the appropriate number of daily journal prompt documents to code for our qualitative analysis, we used a combination of document counts (the daily journal prompts given to one student for one day) and existing code frequency data to ensure sufficient coverage of both frequently and infrequently applied codes. At the time, I had a file with all the data coded with codes and this was used to determine the code frequency for each code across the Contexts and Experiences codes.

To establish the total document pool, we began by identifying the total number of daily journal prompts documents available for coding. For the Fall 2021 iteration, each student was given a daily journal prompt document per day and this iteration ran across four weeks, which would equate to 20 documents per student. However, we did not give them a daily journal prompt document on the last day of the data collection, so students were only able to reply to 19 documents per student. We coded 7 students from the Fall 2021 data collection, which meant there were 133 documents from the Fall 2021 iteration. For the Fall 2020 document, students were given a daily journal prompt document per day and this iteration ran across two weeks, which would equate to 10 documents per student. However, we missed giving them a daily journal prompt document, so the students were only able to complete 9 documents. We coded 5 students from the Fall 2020 data collection, which meant there were 45 documents from this iteration. Accounting for all of the documents across Fall

2021 and Fall 2020 that were coded, there were a total of 178 documents. For ease of calculation, we rounded this number up to 180 documents as our full sample pool.

Next, to ensure the representation of low-frequency codes within the data for IRR, we examined the code frequency of the data I had completed coding (all the data) to be able to understand how often specific codes were applied. In the Contexts code category, the most frequently applied code appeared 325 times, and another relatively lower-frequency code appeared 69 times. Within the Experiences code category, the more frequently applied code appeared 196 times, while the least frequent code was applied 22 time and the next code was applied 46 times, respectively.

To ensure that even low-frequency codes would be sufficiently represented in the data we analyzed, we established a minimum threshold of 50 code applications based on the lower-frequency code counts across both the Contexts and Experiences codes. We reasoned that if a code was applied 50 times across the 180 available documents, this would represent approximately 28% of the total documents. This proportion gave us a rough estimate of how often a typical low-frequency code appears in the dataset.

To determine how many daily journal prompt documents we would need to code to observe at least 10 instances of a low-frequency code (our minimum target), we set up the following proportion below. We set a minimum threshold of 10 code applications as a reasonable target to ensure adequate representation of even the least frequently occurring codes in our sample. This decision was informed by a common qualitative research practice that suggests coding at least 10% of the available data to ensure analytic rigor and representativeness (Guest et al., 2006). Since 10 is 10% of 100, we used this as a simplified heuristic to guide our calculation. We reasoned that if we could observe a given code at least 10 times within our selected sample, we would have enough data to meaningfully analyze and interpret that code — even for those codes that appeared least often in the larger dataset.

This yielded approximately 36 daily journal prompt documents needed to reasonably expect to see at least 10 applications of lower-frequency codes. This served as our minimum document count for this phase of the analysis.

The next step was then to select which data to code for the IRR that met the goal of at least the 36 daily journal prompts documents threshold. Each student in the Fall 2021 dataset completed 19 daily journal entries. Based on our calculation, coding two students would yield 38 documents (19 entries × 2 students), which exceeds the 36-document threshold. This approach ensured that we were likely to capture enough variation and presence of low-frequency codes without needing to code the entire dataset. Only selecting students from the Fall 2021 data was determined to be sufficient because we anticipated that Fall 2021 had a greater diversity in the prompts written to explore SE than the Fall 2020 daily journal prompts due to the addition of additional ESM data to inform the writing of the daily journal prompts. In Chapter 5 we saw that this led to more contextualized prompts impacting the individualization of the prompts.

As such, we first selected Ted Brown as one of the students to be coded because he received many of the individualized daily journal prompts we had the capacity to write. This increased variation in prompt types provides a more conservative test of coder agreement, as the prompts are more likely to differ in content and complexity. The second student we selected was Jane, who had journal entries spanning both the Fall 2020 and Fall 2021 iterations of the course (though we only used her Fall 2021 entries for inter-rater reliability). Jane's experiences were representative of a typical physics undergraduate — including working in a research lab, taking a wide range of STEM courses, and serving as a learning assistant. This made her prompts more likely to reflect diverse academic experiences. Additionally, we had already begun building a case study around Jane, so coding her entries helped establish the credibility of our broader analysis. Coding both Ted's and Jane's data allowed us to meet the minimum document threshold needed for sufficient code coverage.

6.6.1.2 Inter-Rater Reliability Process

We conducted two rounds of IRR. The percent agreements scores are presented in Table 6.3 for the first round of IRR. The first round involved two coders: myself, who developed the codebook, and an experienced coder, who had only limited exposure to its development during an early-stage group meeting. This arrangement allowed us to assess the trustworthiness and usability of the

codebook when applied by someone relatively unfamiliar with its construction.

Code	Percent Agreement
Evaluation of Confidence, Skills, and Success	38.60 %
Influences on Context-Specific Confidence/Success	32.43 %
Comparison to Peers	80.00 %
Interactions with Others	52.63%
Details and/or Feelings Regarding an Event	34.19 %
Task	64.86 %
Course	88.14 %
Degree	69.23 %
Career	93.33 %

Table 6.3

Upon examining Table 6.3, the IRR results revealed notable disagreements between the two coders for several of the Experiences codes. Specifically, four out of the five codes—Evaluation of Confidence, Skills, and Success, Details and/or Feelings Regarding an Event, Influences on Context-Specific Confidence/Success, and Interactions with Others—had percent agreement scores that fell below commonly accepted thresholds (REF). While this might initially appear concerning, low agreement scores can stem from several distinct sources, including: (1) conceptual misalignment between the codes and the data, (2) unclear or ambiguous codebook definitions, and/or (3) challenges related to the unit of analysis (e.g., excerpt boundaries, overlapping ideas within segments).

Rather than attempt to revise all four codes, we intentionally focused our efforts on two: Evaluation of Confidence, Skills, and Success and Influences on Context-Specific Confidence/Success. These were prioritized based on their strong alignment with the theoretical foundations of SE and their importance to the broader goals of the mixed methods design. By focusing on the most theoretically significant codes, we aimed to ensure that the qualitative data would meaningfully contribute to our mixed methods interpretation and strengthen the validity of our findings.

In the case of the Details and/or Feelings Regarding an Event code, we did not conduct a second round of IRR. This decision was based on the recognition that this code serves two distinct functions that extend beyond the immediate aims of this chapter and may require future disentanglement. First, the inclusion of feelings within the code aims to capture physiological states, one of the four

sources of SE. However, the emotional content elicited by prompts asking students about their feelings may not always be tied directly to SE-related physiological states, as conceptualized by Bandura. Second, the code seeks to gain a deeper understanding of the moments in students' daily lives. Questions such as "What were you doing?" or "Who were you with or working with?" are aimed at uncovering the contexts of students' experiences. These prompts do not directly map to SE theory, but instead provide foundational insights into the experiences students report in their journals or their ESM surveys. The multifaceted nature of this code likely contributed to the low percent agreement observed in the first round, as coders may have interpreted the code's focus on feelings and moments differently. For a more detailed discussion of the limitations and future directions related to this code, see Section 6.7.

For the Interactions with Others code, we chose not to conduct a second round of IRR because, while this code is important in relation to SE, we decided to temporarily set this work aside in favor of focusing on the Evaluation of Confidence, Skills, and Success and Influences on Context-Specific Confidence/Success codes. While Interactions with Others could still benefit from further refinement, the Evaluation of Confidence, Skills and Success and Influences on Context-Specific Confidence/Success codes are critical for understanding how students evaluate their own experiences in relation to their SE and how different sources and domains of SE interrelate. Methodologically, capturing this process close to the moment when students reflect on and evaluate their experiences is essential for designing timely interventions that support students' SE in real-time. Therefore, we prioritized refining these codes for the immediate goals of the study, with the understanding that the Interactions with Others code will be revisited in future iterations of the work.

With respect to the Context Codes, they were all above 60% agreement, which we took as a sign of strong face validity and conceptual clarity for this category of the codebook. As such, we did not undertake a comprehensive revision of these codes. However, we did briefly explore the lower-than-expected agreement for the Task code (64.86%), which appeared to be due to whether simply attending a space - such as going to class, office hours, or a lab - should be considered a

task. This points to a subtle but meaningful interpretive distinction about what constitutes action versus presence, which warrants further refinement in future iterations. Nevertheless, the Context Codes as a whole were judged to be sufficiently robust for the descriptive purposes they serve in this analysis, allowing us to concentrate our efforts on revising codes more central to the study's theoretical contributions.

The next step was to perform a second round of IRR to assess whether the refinements to the descriptions of the two codes — Evaluation of Confidence/Skill/Success and Influences on Context-Specific Confidence/Success — led to greater consistency and improved agreement across coders. Given that first the first two coders had discussed the initial results and issues with the codes (which could introduce bias into the coding process), we decided to involve a third coder to evaluate the impact of the revised code descriptions. Coder 3, who had not participated in prior discussions, therefore, coded based on the revised coding descriptions and protocol, thus minimizing potential subconscious biases and providing a more objective assessment.

Coder 3, a postdoctoral researcher, was chosen for this round of IRR. Although Coder 3 had no previously experience with the codebook or the coding protocol, her familiarity with SE was beneficial. Her background in SE provided a deeper understanding of the broader context in which the codes operate, but we also recognized the potential for her knowledge to influence her coding. Nonetheless, since she had not bee involved in the development or prior discussions of the codes, her fresh perspective allowed us to test the clarity of the refined code descriptions without over-relying on the familiarity with the codebook.

Coder 3 and I had two meetings for coder training. In the first meeting, we discussed the logistics and broader context of the IRR process. We discussed the coding protocol documentation, the coding descriptions, and the coding software. Also, we agreed to use the same data from the first iteration of IRR for independent coding. This choice allowed for a direct comparison to the percent agreement calculated previously, helping us assess whether the refinements to the code descriptions led to increased consistency and higher percent agreement.

During the second meeting, I went through the process of coding data with Coder 3. This

allowed Coder 3 to become more familiar with the codebook and observe how I was applying the coding protocol and the two revised code descriptions. It also provided an opportunity for Coder 3 to ask clarifying questions and ensure a solid understanding of the codebook and the coding protocol. The coding protocol, which was not modified from the first to the second round is shared in Appendix F. After this meeting, Coder 1 and Coder 3 independently coded Ted Brown and Jane's Fall 2021 data, and the percent agreement was recalculated. This is shown in Table 6.4.

Code	Round 1 - Percent Agreement	Round 2 - Percent Agreement
Evaluation of Confidence, Skills, and Success	38.60 %	69.69 %
Influences on Context-Specific Confidence/Success	32.43 %	69.23 %

Table 6.4

Following the second round of IRR, we observed an increase in the percent agreement for both codes, signaling that the refinements to the code descriptions had a positive impact on coding consistency. Specifically, the percent agreement for Influences on Context-Specific Confidence/Success rose from 32.43 % to 69.23 %, and the percent agreement for Evaluation of Confidence, Skills, and Success increased from 38.60 % to 69.69 %. Although these values remain below the typically desired 70 % threshold, they represent meaningful progress.

Through follow-up discussions, Coder 3 and I focused on the potential conceptual areas of disagreement. These involved determining whether certain prompts should be coded as Evaluation of Confidence, Skills, and Success or Influences on Context-Specific Confidence/Success. We were able to reach consensus on these cases. A key area that was a challenge was the unit of analysis. Although addressing this would require further refinement, time constraints prevented me from exploring this in detail for the dissertation and will be an important area for future work. It is an important area for future work. For more information regarding the coding protocol used for this work, see Appendix F and G.

6.7 Discussion of the Results

This chapter focused on the development and validation of a codebook designed to analyze a set of daily journal prompts used as research instruments in a broader study of students' SE. Unlike typical coding approaches that center participant responses, this work turned the analytical lens toward the prompts themselves, treating them not as neutral or incidental tools, but as intentional instruments through which SE is explored. This shift acknowledges that the ways in which we construct qualitative tools have direct implications for the kinds of theoretical claims we can responsibly make.

This shift in focus is both methodological and conceptual. If we aim to make claims about students' SE based on their journal responses, we must first establish that the prompts themselves are theoretically sound and capable of eliciting reflection tied to the foundational constructs of SE. In other words, the integrity of our theoretical claims is inextricably linked to the design of the instruments we use to collect data.

Importantly, the codebook was not developed to map prompts directly onto Bandura's four sources or specific domains of SE. Rather, the codes were constructed to reflect the situated, nuanced ways SE is explored within the context of individualized, daily journal prompts. The codebook serves as a theory-informed—but not theory-reductive—framework for analyzing prompt design, offering a structured way to interrogate how the prompts may invite engagement with SE-related constructs. In this way, it strengthens both the methodological transparency and the theoretical integrity of the broader study.

6.7.1 Contexts and Experiences as Essential Elements in Prompt Design

The codebook developed in this study demonstrates that to write individualized daily journal prompts exploring the theoretical foundations of SE (domains and sources), the prompts must draw on both the contexts surrounding the moment and the experiences within that moment. By categorizing responses into distinct contextual and experiential codes, the codebook shows that prompts designed to explore SE cannot focus solely on the task at hand (e.g., the homework assignment). Instead, they must also account for the personal and subjective dimensions of the experience—how

193

the student reflects on their emotions, thoughts, and the surrounding circumstances that might shape their confidence. For example, prompts that ask about task-specific details (e.g., "Describe the assignment you worked on today") are paired with prompts that explore the experience of the task (e.g., "How did you feel about your ability to complete the task after working with a peer?"). This dual focus ensures that the prompts address both the contextual factors shaping SE (task, course, career) and the experiential factors (emotions, thoughts, social interactions, and self-assessments).

To intentionally explore SE, it is essential that the prompts not only account for the context but also explore the additional layers of the experience - how the student perceives themselves in the moment and the surrounding influences (e.g., their emotional state and social dynamics). The codebook demonstrates that, in order to intentionally explore SE via individualized daily journal prompts, the prompts need to draw on both the context (e.g., the task) and the experience of a moment. This careful and deliberate crafting of prompts allows researchers to shape which aspects of SE are explored, offering a more focused and nuanced examination of SE through the individualized daily journal prompts.

6.7.2 Alignment Between Prompt Design and Theoretical Foundations of Self-Efficacy

The development and application of the codebook illustrate that the individualized daily journal prompts align with the theoretical foundations of SE (domains and sources). Rather than mapping directly onto the four sources or domains of SE, the codebook captures the situated and multifaceted ways prompts are designed to explore these theoretical constructs. Through this coding process, it becomes evidence that the prompts vary in how they foreground the domain-specific (e.g., task, course, career, degree) nature of SE and how they up spaces for students to reflect on the source nature of SE. This alignment reinforces the idea that prompt design is not incidental to theoretical exploration - it is a critical site where SE is constructed, negotiated, and made visable through reflection.

While prompt design is often treated as a logistical or procedural step in qualitative research, this study positions it as a core site of theoretical engagement. The alignment between the codebook and the theoretical foundations of SE demonstrates that prompts are not merely neutral tools used

to collect data—they actively shape what aspects of SE become available for reflection. The language, focus, and framing of a prompt can guide students toward particular domains (e.g., task, course, career) and sources (e.g., mastery experience, affective state) of SE, thereby influencing what students attend to, how they make sense of their experiences, and what becomes visible in their reflections. In this way, prompt design is a generative space where SE is not just measured but constructed, negotiated, and surfaced. Treating prompt design as theoretically consequential allows us to be more intentional in how we elicit reflections and, critically, in the kinds of claims we can responsibly make about students' beliefs and experiences.

By illuminating the relationship between prompt design and the theoretical foundations of SE, the codebook underscores that the intentional crafting of prompts is not just a methodological choice, but a theoretical one. This alignment gives researchers the language and structure needed to identify how and where SE constructs are being engaged—whether at the level of domain, source, or both. Ultimately, the codebook affirms that prompts are more than passive containers for student reflection; they are active sites where theory meets practice, shaping both what students reflect on and what researchers can meaningfully claim.

6.7.3 Individualized Daily Journal Prompts Extend Beyond Domain-Specific Self-Efficacy: Including the Source-Nature of Self-Efficacy

In contrast to traditional instruments like ESM surveys—which are often used to capture task-specific or course-level SE in real time—individualized daily journal prompts offer a complementary approach that expands the theoretical terrain researchers can explore. While ESM excels at identifying fluctuations in domain-specific confidence within a given moment, it is typically constrained to surface-level indicators of task, course, or career-level SE. The individualized daily journal prompts, by comparison, enable reflection not only on these same domains but also on the sources of SE (e.g., mastery experiences, affective states, social persuasion, and vicarious experiences). This flexibility allows for a more holistic and theoretically grounded engagement with SE, offering insights into how students interpret their experiences and construct their confidence across multiple layers of academic life.

This expansion is important because it illustrates the added value of daily journal prompts as a methodological tool. While ESM captures the domain-specific dimensions of SE with high frequency and granularity, individualized journal prompts make it possible to also explore the source nature of SE—revealing the mechanisms behind students' reported confidence. This finding reinforces the value of prompt design not just as an act of reflection, but as a powerful way to deepen our theoretical understanding of SE.

The codebook developed in this study makes this distinction visible by including codes that capture both the domain (e.g., task, course, degree, career) and source (e.g., mastery experience, vicarious experience, affective state, social persuasion) nature of SE. While the ESM survey is limited to measuring the domain in which confidence is expressed, the codebook allows researchers to identify why students might feel confident or uncertain—what experiences, emotions, or inter-actions are shaping their beliefs. For example, a prompt asking a student to reflect on how they felt after presenting a project may surface evidence of social persuasion or mastery experience, even though the surface context remains task-level. In this way, the codebook operationalizes a theoretical nuance that other tools may miss, offering a more layered understanding of SE.

This capacity to explore both the domains and the sources of SE underscores the unique methodological and theoretical contribution of individualized daily journal prompts—offering researchers a more textured and explanatory view of how SE is shaped in real-world academic contexts.

6.7.4 Strengthening Theoretical Precision and Trustworthiness through Inter-Rater Reliability (IRR)

In this study, inter-rater reliability (IRR) was not simply a procedural step to ensure consistency in coding—it played a crucial role in shaping the theoretical application of SE constructs. Through ongoing discussion and clarification, IRR helped refine the codebook, ensuring that the categories were not only reliable but also theoretically precise. A key example of this came when the team worked to distinguish between the "Influences on Context-Specific Confidence/Success" and "Evaluation of Confidence, Skills, and Successes" codes, which revealed the different theoretical

196

functions these codes serve in exploring SE.

The Influences on Context-Specific Confidence/Success code was developed to explore how different domains of SE (e.g., task, course, career) relate to each other and influence students' confidence within specific contexts. This code also captures how sources of SE (e.g., mastery experience, affective states, social persuasion, vicarious experiences) might be prioritized or interact differently depending on the domain being explored. For example, students may rely more on mastery experiences when reflecting on academic tasks, but may draw more on social persuasion when thinking about their future career. This broadens the scope of the code to not only look at relationships between domains but also to explore the dynamic interplay between sources and their varying influence in different contexts.

On the other hand, the Evaluation of Confidence, Skills, and Successes code focuses on understanding how students evaluate their own SE by reflecting on their experiences and how they prioritize different sources of SE (e.g., mastery experiences, affective states, social persuasion). This code was built around the theoretical idea that different students may prioritize different sources of SE when evaluating their confidence or abilities. It allows us to examine how students process and interpret their own capabilities, based on how they assess information and use it to evaluate their SE.

The IRR process was instrumental in clarifying these distinctions, ensuring that each code was not only consistent in its application but also theoretically grounded. The discussions led to refinements that enhanced our understanding of how each code serves to explore distinct dimensions of SE—whether it's exploring how different domains influence each other, how sources of SE interact within those domains, or how students reflect on their sources of confidence and skills.

Rather than being a static task, IRR became an active tool for conceptual refinement, sharpening the focus of the study and ensuring that the prompts didn't just assess surface-level confidence but also captured the nuances of SE—its sources, dynamics, and relationships. By requiring the team to continuously revisit and refine the codebook, IRR ensured that the prompts were grounded in a rich, theoretically informed framework, actively shaping what is made visible in students'

197

reflections on SE. Thus, the IRR process in this study demonstrates that methodological rigor and theoretical precision are inextricably linked. Ensuring the trustworthiness of the coding process simultaneously strengthened the theoretical application of SE constructs, allowing for deeper and more valid conclusions about how SE is explored through individualized journal prompts.

6.7.5 Limitations of Codebook

Although the codebook currently includes separate coding schemes for Contexts and Experiences, it does not yet provide a framework for designing the individualized nature of the daily journal prompts. Tailoring prompts to students' ESM data and experiences requires thoughtful decisions during data collection and utilizing students' open-ended survey responses and previously daily journal prompt responses has been valuable in this process. However, the codebook does not currently outline specific tools or strategies for systematically tailoring prompts. Developing such a framework would enhance methodological transparency and support a more systematic approach to using daily journal prompts to explore students' experiences and SE in greater depth.

The codebook does not yet provide a framework for how daily journal prompts specifically complement the ESM survey data. In Chapter 5, our case study analysis outlined a framework for identifying threads from the ESM survey data to follow up on in the daily journal prompts, shaping how these prompts were written and tailored to students' experiences. However, this paper does not examine whether the daily journal prompts and their responses primarily serve to expand upon the ESM survey findings, confirm observations from the ESM data, or introduce counterarguments that challenge these findings.

Although the codebook includes the code 'Details and/or Feelings Regarding an Event', it does not fully capture the temporal aspect of the daily journal prompts. Additionally, while this code accounts for prompts that ask students to reflect on their feelings, the level of specificity is critical. Researchers often investigate the physiological states of participants by asking about their feelings in relation to their confidence or abilities. However, not all prompts designed to explore the feelings of participants are necessarily related to their physiological states. Similarly, I have previously discussed the limitations of 'Interactions with Others' and 'Comparison to Peers'.

Furthermore, our ability to explore transfer experiences through daily journal prompts was inherently limited by the scope of our participant pool. Because we only collected data from students at bachelor's degree-seeking institutions who had transferred from a two-year institution, the types of prompts written to explore transfer experiences were shaped by this specific population. Additionally, as discussed above, we excluded certain daily journal prompts from coding in this study because they appeared to focus on validating aspects of the ESM survey data - such as how students interpreted their task-specific SE scores or patterns in how they reported high versus low scores and what those patterns might indicate.

Although this work provides an initial framework for examining the structure and content of daily journal prompts, more work is needed to refine how prompts are designed, analyzed, and linked to the ESM survey data. Addressing these limitations will enhance the methodological transparency and depth of the SE exploration, ultimately strengthening our understanding of how students reflect on their academic experiences in real time.

6.8 Conclusion

This study has explored the development and application of a codebook for analyzing individualized daily journal prompts as a tool for examining SE in academic settings. Through this work, we have demonstrated how journal prompts, tailored to students' lived experiences, can reveal the multifaceted and dynamic nature of SE across various academic domains. By expanding the scope of inquiry to include both the domain-specific and source-specific aspects of SE, this study offers a more comprehensive framework for understanding how students' confidence and self-beliefs evolve throughout their academic journeys.

The key contribution of this research lies in its methodological innovation: the use of individualized journal prompts not only as a data collection tool but as an active mechanism for exploring SE in real time. This approach allows for a deeper, more nuanced understanding of how students interpret their experiences and construct their confidence in relation to specific academic tasks, courses, and broader career goals. Moreover, the development of the codebook enables the systematic analysis of these prompts, offering a novel way to map the complex interplay between students'

SE beliefs and their lived academic experiences.

As this work progresses, there are several key directions for future research. Expanding the codebook to incorporate a more structured approach for designing tailored prompts, linking journal responses with existing SE data, and further investigating the temporal dimensions of SE are critical next steps. These efforts will deepen our understanding of how SE develops and fluctuates in response to students' academic activities and interactions. Additionally, exploring the role of emotion and physiological state in these reflections, along with examining how these prompts align with other tools like the ESM survey, will strengthen the methodological rigor and theoretical precision of future SE research.

In conclusion, this study underscores the importance of methodological choices in shaping the insights we gain into students' SE. The iterative nature of the research process, the inclusion of students' voices, and the refinement of our approach to data collection all contribute to a more robust exploration of how SE is constructed, negotiated, and sustained. As we continue to refine and expand upon this work, we look forward to furthering our understanding of how personalized prompts can empower students to reflect meaningfully on their academic experiences and cultivate stronger beliefs in their abilities to succeed.

CHAPTER 7

THE ROLE OF EVIDENCE IN LINKING QUANTITATIVE AND QUALITATIVE DATA FOR VALID SELF-EFFICACY CLAIMS

This chapter extends the methodological and analytical work presented in Chapters 5 and 6 by elaborating on how connections were made between quantitative and qualitative data in the study. Specifically, it focuses on two core strategies for integration: the use of "following the thread" as a method for linking insights across data sources, and the analytic role of joint displays, which are introduced here for the first time to support the visualization of the linkages being made.

In answering these questions, I present an analytic framework for identifying and justifying linkages across datasets, grounded in three distinct types of evidence. I also explore how the direction of the analytic thread – whether beginning with survey data or individualized daily journal reflections – impacted the interpretation and credibility of those connections. Ultimately, this chapter gives transparency to the reasoning that shaped the integration of data and highlights the methodological choices that supported or potentially limited the building of cross-source linkages.

7.1 Introduction

Mixed methods research holds the promise of offering deeper, more nuanced understandings of complex phenomena by integrating quantitative and qualitative data (REF). Integration can be thought of as the dialogue between the quantitative and qualitative strands, where each informs and enriches the other [70]. This interplay is often regarded as the defining characteristic of mixed methods research studies [13, 70]. Yet, the practical work of constructing meaningful linkages, evidence that integration occurred, between data types is often hard to do in practice. While some empirical studies describe how integration was implemented in specific research contexts [112, 113], other works discuss integration in more abstract, theoretical terms [63, 69, 71]. For researchers new to mixed methods research, like myself, navigating this divide and determining how integration applies to one's own work remains a critical and often daunting task.

Within this work, we present a mixed methods research approach that first collects Experience Sampling Method (ESM) survey data – open- and closed-ended data gathered via survey – followed

by individualized daily journal prompts and reflections. This design was chosen to explore students' self-efficacy (SE) experiences in a manner that is closer to when those experiences are occurring. Central to this work are two key research questions:

1. What does it mean to connect quantitative and qualitative data in this particular mixed methods approach to exploring students' SE?

2. How does one build linkages between quantitative and qualitative data in practice?

To answer these questions, I explore the methodological and interpretative work involved in integrating these two data sources (i.e., data from ESM surveys and data from the individualized daily journal reflections). In previous work (Chapter 5), I discussed the integration of "following the thread", which links ESM data and individualized journal reflections as I used the observations from the ESM data sources together to form threads, and I followed up on these threads when writing the individualized daily journal prompts. Here, I build upon that discussion by presenting specific evidence from the ESM survey data and the journal reflections and explain why this evidence supports the claim that these data sources are meaningfully linked. This leads to third research question of this work:

3. What does it mean for our data to be meaningfully linked in the context of this mixed methods approach?

Rather than presenting the production of this integration, this paper focuses on the process itself – examining how we know the data can be linked to form an integrated "thread" and identifying the types of the evidence that validate this link. Specifically, I present a typology of three evidence types that were used to justify integration, along with examples of how each type shaped – or in some cases constrained – the analytic process. I will show that the strength and direction of the thread not only guided interpretation but also influenced the perceived credibility of cross-source linkages.

In doing so, this work contributes to the growing body of mixed methods literature that attends, not just by discussing research design, but also by attending to the lived and interpretative labor of data integration – highlighting the complexities of moving from distinct data sources to a coherent, meaningful whole.

7.2   Bridging Traditions: Triangulation, Integration and Mixed Methods Thinking

Discussions of integration in mixed methods research are often rooted in earlier methodological conversations about triangulation (Greene et al., 1989; Creswell and Plano Clark, 2018). Across both qualitative and quantitative traditions, triangulation has long served as a strategy for increasing credibility, validating findings, or uncovering deeper understandings of complex phenomena (Denzin, 1978; Mathison, 1988). In qualitative research, triangulation may involve drawing on multiple data sources, perspectives, or theoretical frameworks to enhance interpretive depth. In quantitative work, it is often linked to measurement reliability or design validity through the use of multiple indicators or instruments.

Greene, Caracelli, and Graham (1989) identified triangulation as one of five primary purposes for using mixed methods, emphasizing convergence and corroboration through the comparison of qualitative and quantitative findings [65]. While this framing positioned triangulation as a core motivation for mixing methods, subsequent scholarship has questioned whether it fully captures the breadth and potential of mixed methods inquiry. Some scholars argue that the term triangulation has become overly associated with validation, which can obscure the broader purposes of integration—such as complementarity, expansion, and initiation [67–69]. Others have shifted toward focusing on what researchers do with their data, emphasizing how the intent of the study (e.g., in a convergent design) shapes the role of integration beyond mere confirmation (REF).

Within mixed methods research, triangulation is increasingly reimagined not solely as a means of validation, but as a purposeful act of integration—bringing together distinct strands of data to offer more layered or generative insights (Creswell and Plano Clark, 2018). This shift reflects a move from viewing triangulation as a strategy for checking consistency to treating integration as an interpretive process—one that can reveal tensions, open new questions, and deepen analytic

insight.

The following section explores how integration has been conceptualized in the mixed methods literature, the frameworks that guide its implementation, and the challenges of making integration visible, transparent, and meaningful in practice.

7.3 Conceptualizing and Operationalizing Integration in Mixed Methods Research

While integration is often described as the defining feature of mixed methods research, what it means to integrate - and how integration is achieved - varies across the literature (). At a broad level, integration refers to the intentional process of bringing together qualitative and quantitative strands of a study to generate insights that are more comprehesive, complex, or nuanced than could be achieved by either approach alone (Creswell and Plano Clark, 2018; Fetters et al., 2015; Fetters et al., 2017; Tashakkori and Teddlie, 2010). However, beyond this shared goal, scholars offer a range of conceptual models and practical strategies to guide the work of integration in different research designs (Creamer, 2020; O'Cathain et al., 2010; Fetters et al., 2013; Fetters et al., 2017; McCrudden et al., 2021).

One of the earliest and most cited contributions to this conversation is Greene, Caracelli, and Graham's (1989) typology of five purposes for mixing methods: triangulation, complementarity, development, initiation, and expansion [65]. Each purpose reflects a different orientation to integration—for instance, using one data type to elaborate on or clarify another (complementarity), or seeking to provoke new lines of inquiry through the contradictions between strands (initiation). This framework has been foundational in shifting integration away from a singular focus on validation toward a broader set of interpretive and generative goals.

More recent scholarship has focused on *how* integration happens at the level of methods and analysis. Fetters, Curry, and Creswell (2013), for example, offer three techniques for achieving integration: *merging*, where qualitative and quantitative data are brought together during analysis; *connecting*, where data from one strand inform the sampling or data collection of the other; and *embedding*, where one type of data is nested within a larger design driven by the other [62]. These approaches provide language and structure for operationalizing integration across different phases

of the research process.

One strategy that exemplifies Fetters et al.'s (2013) concept of *connecting* is the method of "folloing a thread" - an approach in which a theme, concept, or case emerging from one data strand is used to inform the inquiry in another (Moran-Ellis et al., 2006; O'Cathain et al., 2010). Rather than treating datasets as parallel or static, this strategy treats integration as an evolving process of analytic responsiveness, where insights gained from one method shape the questions, prompts, or focus of the next. In doing so, "following the thread" highlights the dialogic and interpretive labor of integration, supporting not just continuity across data sources but also deeper contextualization of emergent patterns. "Following the thread" has been increasingly used in mixed methods studies to support deeper integration across strands [73], emphasizing how a concept can be traced and developed across time, method, or context. In my study, this strategy was central to how I connected the responses of the ESM survey of the students to their individualized journal prompts, using the initial quantitative and open-ended survey data to shape personalized reflection opportunities that extended and complicated previous responses.

Despite these conceptual tools, researchers continue to grapple with the practical challenges of integration. As Bryman (2006) and others have noted, integration is often discussed at the design level but remains undertheorized and under-documented at the level of actual practice [74]. The complexity of aligning data sources that differ in form, timing, and epistemological grounding means that integration often involves substantial interpretative labor—work that is difficult to capture in published write-ups but central to the credibility and coherence of a mixed methods study.

In the following section, I build on these foundational frameworks to explore the specific methodological and analytic work of integration in my own study. I focus particularly on how integration was enacted between the ESM surveys and the individualized journal reflections, and on the types of evidence that made those linkages visible, interpretable, and meaningful.

7.4 Challenges in Integration in Practice

Despite the robust frameworks and conceptual tools for integrating qualitative and quantitative data in mixed methods research, many researchers face significant challenges when attempting to make integration meaningful in practice [66]. While integration is often discussed as a theoretical ideal, the practical work of aligning, linking, and interpreting data from different methods can be fraught with complexity [66].

One of the primary challenges is the visibility of integration. In many mixed methods studies, the process of integration is implied but not always clearly documented. While the final product may demonstrate how the qualitative and quantitative strands inform one another, the steps taken to create this integration remain hidden in much of the literature [62]. This lack of transparency can make it difficult for others to replicate or critically assess how integration was achieved, raising concerns about the rigor and trustworthiness of the findings. Without explicit documentation of how integration occurs, researchers may struggle to convince others that their approach is valid or meaningful.

Another challenge is the timing and alignment of integration. At what stage in the research process should integration occur? Should it happen during data collection, analysis, or interpretation? In some designs, qualitative and quantitative data are collected concurrently, while in others, one data type may be collected first, followed by the other (Creswell and Plano Clark, 2018). The challenge lies in ensuring that the data collected at different stages are aligned in a way that allows for meaningful integration. For example, if qualitative data is collected after quantitative data, it may be difficult to adjust the qualitative instruments or data collection methods to directly respond to the patterns revealed in the quantitative phase. This misalignment can lead to difficulties in ensuring that the integration is both intentional and coherent.

Perhaps the most significant practical challenge in integration is the interpretative labor involved in synthesizing data from distinct methods and paradigms [66]. Quantitative data are typically analyzed through statistical tools and focused on measurement, whereas qualitative data are often more descriptive, seeking to capture rich, nuanced insights. The task of bringing these two forms of

data together in a way that preserves the integrity of each while generating new, combined insights is highly interpretive and requires careful analysis (Tashakkori and Teddlie, 2010). Researchers must make decisions about how to "merge" or "connect" data [62], and these decisions often involve subjective judgment calls about what constitutes meaningful integration.

The practical challenges of integration can significantly shape the way mixed methods research is designed and executed. In the next section, I examine how these challenges are addressed within the context of my study, particularly the integration of ESM data and individualized journal reflections, and I provide a detailed look at how these two data sources were meaningfully linked.

7.5   How Integration is Operationalized in This Study

Building on the conceptual frameworks described above, this section outlines how integration was enacted in this mixed methods study. As shown and discussed in previous chapters, the mixed methods approach we have developed consisted of two primary data strands: (1) open- and closed-ended responses collected through ESM surveys, and (2) individualized daily journal reflections written by students in response to prompts informed by those surveys. Integration was not treated as post-hoc interpretive exercise, but rather as a guiding methodological principle that informed the design, data collection, analysis, and interpretation across the study.

The integration strategy in this work most closely aligns with Fetters et al.'s (2013) notion of *connecting*, where one data strand directly informs the subsequent strand. Specifically, responses gathered through the ESM surveys - particularly the open-ended survey items and the task-specific SE ratings - were used to develop individualized daily journal prompts for each participant. These personalized follow-ups invited students to elaborate, reflect, or complicate their earlier ESM responses, thereby extending the inquiry in a wy that was both timely and tailored. In this way, integration began during data collection and continued through analysis, as insights were traced across moments, methods, and contexts.

A central tool in enacting this integration was the strategy of "following a thread" [69, 72, 73], which provided a flexible but focused way of linking data across sources. In this approach, a concept, theme, or question emerging from the ESM data was used to guide the content of a

student's subsequent journal entry, forming a kind of analytic throughline or "thread" across the dataset. This not only created a sequenced connection between data points, but also offered an interpretive bridge - making it possible to examine how students' SE experiences developed across time and contexts, and across quantitative and qualitative modes of expression.

To ensure that these connections were not simply assumed, and to make the logic of integration visible and credible to others, this study draws on Zhou and Wu's (2022) framing of integration as both a process and a product. In their discussion of mixed methods validity, Zhou and Wu emphasize that the quality of integration should not be treated as a static outcome, but rather as an evolving argument built through empirical evidence. Drawing parallels to Messick's (1989) and AERA/APA/NCME's (1999) definitions of validity in assessment, they argue that "integration is the property that indicates the quality of mixed methods research from an ontology perspective; whereas collecting evidence for integration is an ongoing process of developing a scientifically sound integration argument to support the value and quality of mixing different types of data in a single study" (Zhou and Wu, 2022). In this view, integration requires more than a conceptual intention to mix — it demands that researchers provide concrete, transparent, and theoretically grounded evidence that the integration achieved is meaningful, purposeful, and adds value to the understanding of the phenomenon under study.

In the next section, I walk through how I went about collecting and articulating this evidence of integration. Specifically, I introduce a typology of three distinct evidence types that supported the identification and validation of analytic threads across ESM surveys and individualized journal reflections. These evidence types did not simply justify integration in general terms—they were the means through which integration happened. That is, following a thread involved recognizing moments where data sources meaningfully intersected, and those intersections were made visible, interpretable, and credible through the accumulation of evidence. This typology therefore functioned as a methodological scaffold for determining when a connection between data sources was not only plausible, but also analytically productive. Through concrete examples, I demonstrate how different forms of evidence helped build and sustain specific threads, and how these threads

ultimately shaped the claims I make about students' SE experiences and the added insight that emerged through this integrated approach.

7.6 Methods: Operationalizing Integration in Practice

This section outlines the methodological decisions that supported the integration of the quantitative and qualitative data sources and data in this study. Specifically, this work describes the explanatory sequential mixed methods design, the collection and organization of the ESM survey and individualized daily journal reflections (i.e., the prompts and the responses), and the analytic and interpretative strategies used to develop meaningful linkages across these data sources and data. In doing so, this work lays the foundation for understanding how integration was made visible and analytically generative in the subsequent stages of the research.

7.6.1 Study Design Overview

This study used an explanatory sequential mixed methods design in which students' SE experiences were explored in real time through daily ESM surveys, followed by individualized daily journal reflections. The ESM surveys, which included both closed- and open-ended survey items, were designed to capture momentary fluctuations in students' SE throughout the day. Individualized daily journal prompts were written to support the ESM data as well as to explore the theoretical foundations of SE (as shown in Chapter 6), asking the participants to reflect deeply on key patterns, tensions, or moments observed from their ESM survey responses.

This design was chosen for its potential to balance immediacy and reflection, with the ESM providing snapshots of SE in action and the journals offering space for further exploration of their SE. The interplay between these two forms of data collection allowed the emergence of "analytic threads" that could be traced across sources – threads that would later serve as the basis for integration strategies detailed in the analysis.

7.6.2 Participants and Context

This work draws on the Fall 2021 iteration of using our mixed methods approach. Participants in this iteration were undergraduate students enrolled in STEM-related courses as at a large public research university. As a part of the selection criteria, students were chosen to participate in the

study if they had at least one or more credits transferred from a community college and/or identified as a first-generation college student. Approximately 18 students started the study and participated to some extent during the first two weeks of the data collection, with data from 15 students included by the end of the study after accounting for those who dropped out or had limited participation. They were asked to participate in this mixed methods study for four weeks, with a week-long break between the first two weeks of consecutive data collection and the last two weeks of consecutive data collection.

For the context of this work, we drew on data from five students—Cristina, Jane, Ted Brown, Yay, and Charlotte (pseudonyms chosen by the students). One key integration strategy in this study involved linking ESM data with individualized daily journal reflections, drawing on the fact that both data sources were tied to the same student. This approach enabled more cohesive and meaningful integration across data types. Cristina, Jane, and Ted Brown were selected due to their overlap with the work presented in Chapters 5 and 6, while Charlotte and Yay were included based on the broader mixed methods project's needs and the analytic focus developed in collaboration with the postdoctoral researcher. These students supported the first step in our analytic process of linking ESM survey data and individualized daily journal reflections.

In the next phase of the analytic process of examining in more depth the types of integration evidence, we narrowed our focus to Ted Brown and Jane, whose data offered particularly strong opportunities to examine integration in depth. Ted Brown's journal responses were especially rich, prompting the development of more tailored prompts during data collection. Jane's data had been used extensively in previous analyses (Chapters 4, 5, and 6), making her a natural continuation point for this phase of the study. Her case continues to offer promise for future, more focused case study work on SE using this integrated approach.

7.6.3 Data Collection Procedure

ESM Surveys. Participants received one survey four times per day across four designated time windows. The survey was designed to measure the domain-specific nature of SE. The full survey instrument can be found in Appendix **??**. This format generated high-frequency, real-time data that

captured both quantitative trends related to domain-specific measures of SE and qualitative cues about the unfolding experiences participants were engaged in when rating their task-specific SE.

Individualized Daily Journal Prompts. The research team—including myself—collected and reviewed ESM survey responses daily to inform and craft individualized daily journal prompts. Participants were signaled to complete their journal entries in the evening. These prompts were designed to probe emerging threads (as discussed in Chapter 5) observed in the ESM data. This approach supported responsive reflection, allowing participants to deepen and contextualize moments of significance that surfaced through their ESM responses.

7.6.4 Designing for Integration: Following the Thread

A key integrations strategy in this design was the method of "following thread" (Moran-Ellis et al., 2006; O'Cathain et al., 2010; Dupin and Borglin, 2020), in which observations from one data source (e.g., a spike in reported SE in the ESM) prompted targeted inquiry in the next (e.g., a journal prompt seeking explanation, elaboration, or confirmation). Rather than treating the ESM and journal reflection data as parallel or post hoc datasets, the approach actively used one to shape and contextualize the other (as seen in Chapter 5).

This process made integration iterative and dialogic - data were not only collected sequentially but interrogated across sources, with the goal of building threads of interpretation that could trace the evolution of SE experiences across time and format.

Linking Data Sources Linkages between ESM and journal data were not assumed, but were constructed and validated through analytical decisions. These included:

- Temporal proximity (e.g., journal entries directly responding to same-day ESM data),

- Thematic alignment (e.g., recurring language or concepts across data types),

- Participant sense-making (e.g., students explicitly referencing earlier ESM experiences in journal responses),

- Contradictions (e.g., moments where journals complicated or challenged what was reported in the ESM surveys).

As I built these threads from the ESM and journal data, I became increasingly aware of the specific kinds of evidence I was drawing on to justify integration. This process clarified which data sources—outlined in Table 7.1—I was using to claim that integration had occurred. That awareness contributed directly to the development of a typology of evidence types for integration, which is elaborated in the following section and operationalized in the analytic framework used for interpretation.

Data Source	Data
ESM Task Open-Ended Survey Item	ESM Task Open-Ended Survey Item Response
ESM Course Open-Ended Survey Item	ESM Course Open-Ended Survey Item Response
ESM Closed-Ended Task-Specific SE Survey Item Language	-
ESM Closed-Ended Task-Specific SE Survey Items	- ESM Closed-Ended Task-Specific SE Survey Item Responses
Individualized Daily Journal Prompt	Individualized Daily Journal Prompt Response
Individualized Daily Journal Prompt Language	Individualized Daily Journal Prompt Response

Table 7.1 Overview of the primary data sources and the corresponding participant data used to support integration across the ESM Surveys and the Individualized Daily Journal Reflections.

7.6.5 Analytic Process for Identifying the Typology of Evidence Types

The development of the typology of evidence types emerged from an earlier phase of analytic work, in collaboration with a postdoctoral researcher. At that time, we were exploring whether the direction in which we attempted to link the ESM survey data with the individualized daily journal reflections impacted the kinds of linkages we were building.

The postdoctoral researcher began with the individualized daily journal reflections—examining both the prompts and students' responses—before identifying relevant connections in the ESM data. I, on the other hand, began with the ESM survey responses and worked toward linking them to the journal reflections. We each recorded the linkages we identified in a shared Google Sheet, where each row represented a single instance of integration between data sources.

Conceptually, we did not expect the starting point to matter. Because the underlying data remained constant, we assumed that we would arrive at similar linkages regardless of which

source we began with. However, as I began examining the table more closely, I noticed subtle but meaningful differences in how linkages were constructed depending on the analytic direction. These differences were not only in the content of the connections but also in the strength, clarity, and interpretive distance between the two data points.

This observation led me to a deeper reflection on how connections between data sources were being justified. Some linkages felt grounded in clear thematic or temporal continuity, while others relied more heavily on inference or interpretation. To make these distinctions more explicit—and to create a framework that could guide and support future integration work—I began developing a typology of evidence types. This typology became a key methodological tool, offering both structure and transparency in the analytic process and allowing me to more rigorously evaluate the credibility of each linkage.

Thus, the next stage of my analytic process involved taking a deliberate pause to articular the kinds of "evidence" I was relying on to justify the linkages built between ESM survey responses and individualized daily journal reflections. I returned to the linkage table in the shared Google Sheet and began examining each entry – scrutinizing the specific components of both the ESM data and the individualized daily journal reflections that were being used to establish a connection.

As I worked through these linkages, I began to notice recurring patterns: certain types of information, phrases, or alignments appeared repeatedly in my reasoning. These patterns pointed to distinct forms of evidence that I was drawing on – sometimes implicitly – to claim that a meaningful relationship existed between the two data sources. Recognizing this, I began developing a typology to name and organizes these evidence types, with the goal of making my analytic decisions more transparent and consistent.

7.7 Approach to Integration: Constructing Linkages Between the ESM Surveys and Individualized Daily Journal Reflections

This section bridges the methodological and analytic components of the study by examining how quantitative and qualitative data were meaningfully integrated in the analysis of students' SE experiences. Rather than presenting findings from the ESM surveys and journal reflections

213

in parallel, I focus on the interpretive labor of integration—how threads of meaning were traced between these two data sources to construct coherent analytic insights.

Two key strategies guided this integrative process. First, I employed the technique of "following the thread" (Moran-Ellis et al., 2006), using participants' ESM responses to generate individualized journal prompts that deepened, clarified, or complicated their earlier responses. This approach enabled near real-time, sequential connections to emerge and supported the identification of thematic or conceptual continuities across data types. Secondly, while initial versions of this analysis employed joint displays (REF) to organize and visualize data integration, I present the examples below in narrative and side-by-side form to foreground the interpretive movement between sources. This approach emphasizes the analytic logic that guided integration, rather than relying solely on visual juxtaposition.

To ensure the credibility and transparency of these connections, I constructed a typology of three evidence types that justified when and how data sources were linked. These evidence types served as methodological scaffolding for the analytic process, helping to distinguish stronger, more grounded connections from those that were interpretively tentative or exploratory.

The Results section will be present the typology of the evidence types that supported data integration, in which around each typology we will outline its' description, discuss why this kind of evidence is a valid sign of integration between the ESM surveys and the individualized daily journal reflections, and then follow up with example of using the evidence to show evidence around whether to integrate the ESM surveys with the individualized daily journal reflections.

7.8   Results

The goal of this section is not to present an exhaustive list of all possible types of evidence that could support linkages between ESM surveys and individualized daily journal reflections. Rather, the focus is on demonstrating how specific elements—such as task-specific SE item responses, numeric scores, or the language embedded in those items—can function as evidence to support integration across sources.

To be clear, the three evidence types described below do not represent the full range of ways

these data sources might be linked. Instead, they reflect the types of evidence that most meaningfully supported the analytic work in Chapter 5. These forms of evidence were especially generative in helping to "follow the thread" across data points, and they frequently served as the foundation for identifying continuity, divergence, or elaboration in participants' experiences.

Importantly, the presence or absence of a particular type of evidence should not be read as a binary indicator of whether integration is possible. Rather than assuming that linkage depends solely on pre-defined types of evidence, the analysis emphasized alignment and non-alignment between the ESM and journal data. These terms - alignment and non-alignment - allow for interpretive nuance, acknowledging both convergences and dissonances as meaningful analytic moments.

7.8.1 Typology of Evidence Types Use for Integration

Evidence Type: Alignment Between the Temporal of the ESM Surveys and the Individualized Daily Journal Reflection

One key aspect of integration between the ESM surveys and individualized daily journal reflections lies in the temporal alignment between these data sources. This evidence type refers to the way journal prompts are designed to either probe ESM responses from the same day or revisit ESM data from a previous day. The timing of both the surveys and the reflections creates an opportunity to explore shifts in SE, either immediately after the ESM survey or through more delayed reflective prompts.

In some cases, journal prompts are written to directly probe the responses from the ESM survey completed earlier that same day. These prompts invite students to elaborate on their SE ratings, uncovering the underlying reasons behind their confidence or uncertainty at that specific moment. This alignment between the ESM survey and journal reflections on the same day allows for a deeper understanding of how students perceive and articulate their SE in real time.

In other cases, journal prompts may be written on the following day, encouraging students to reflect on their experiences from the previous day, including their ESM responses from that day. Here, the alignment is less about strict temporal overlap and more about allowing students to revisit or reconsider their SE ratings with the benefit of additional time and reflection. This

215

retrospective reflection provides a more nuanced understanding of SE as students may see their confidence differently after having had time to process the experience.

Additionally, some journal prompts were designed to encourage broader reflections on SE, such as experiences over the past week or month. These prompts might not align directly with any specific ESM survey response but still offer valuable insight into students' evolving SE over a longer period. While these broader reflections don't have a clear one-to-one correspondence with the ESM data, they offer a larger view of how SE develops across various time frames, helping to situate individual ESM responses within a broader context of the student's lived experience.

The integration of the ESM survey data with individualized journal reflections—whether temporally aligned or not—serves to deepen the understanding of students' SE. Temporal alignment and non-alignment both offer unique insights into how SE is experienced, interpreted, and reinterpreted over time. The ability to track shifts in SE through these reflective practices makes this integration a powerful tool for understanding how SE develops in real-world contexts, providing richer, more dynamic data than could be captured by either the quantitative or qualitative data sources alone.

This evidence type is a valid sign of integration because it allows us to track evolving perceptions of SE over time. The ability to reflect on a given day's SE data through journal prompts written immediately or the following day highlights both continuities and changes in students' beliefs about their abilities. By exploring both immediate reflections (same-day prompts) and more delayed retrospections (reflecting on experiences from the overall week or month), we gain insights into how students' SE may be shaped by ongoing experiences and changing perspectives. This form of integration provides a dynamic and evolving picture of SE, where non-alignments are just as informative as alignments, offering critical moments for analysis and interpretation.

Example of Temporal Alignment:

An example of identifying an alignment between the temporal aspect between an ESM survey response and an individualized daily journal reflection is given below.

Student's ESM Survey: Week 3, Day 4, Notification #55

Corresponding Daily Journal Reflection for Week 3, Day 5 Reflection:

- Prompt: "Yesterday, in the app, you told us you were working on [modern physics] home-work, and today, in the app, you told us you were working on [modern physics] homework. We noticed that you indicated you were feeling less successful on your [modern physics] homework. Can you tell us what caused this shift?"

- Response: When I was working on the homework the first day what made the biggest difference is that I only focused on the ones that I could do for sure, because we had not yet covered all of the content in the class yet. So where I struggled, I assumed it was because we had not covered it yet. So when I was working on it today, where my lass class for this lesson was yesterday, I had left a couple of problems yesterday that I did actually have the knowledge for in the past days but were just really trick to do.

In this example, the temporal alignment between the ESM survey response and the individual-ized daily journal reflection can be identified in several ways:

- Timing and Sequence: The ESM survey response was collected on Week 3, Day 4, and the corresponding daily journal reflection was written on Week 3, Day 5. This reflects the temporal proximity between the two sources of data, suggesting that the journal prompt is probing the student's experience shortly after completing the ESM survey. The sequential nature of the data collection process strengthens the connection between these two moments in time.

- Language in the Prompt and Response: In the journal prompt, the phrase "Yesterday, in the app, you told us. . . " is a clear indicator that the student is being asked to reflect on their experience from the day prior (Week 3, Day 4). This explicit reference to "yesterday" directly links the journal reflection to the ESM survey data from that day. Furthermore, the student's response, using the language "the first day," further grounds the experience in Week 3, Day 4, connecting it back to the ESM survey.

217

- Temporal Elements in the Student's Reflection: The student's response also provides additional temporal language, such as "the first day" and "today." These references support the alignment between the ESM survey data from Week 3, Day 4 and the journal reflection from the following day (Week 3, Day 5). The student's use of temporal language in their reflection helps reinforce that they are reflecting on their experience from the previous day while drawing on more recent insights gained the following day.

By examining both the timing of data collection and the explicit temporal language used in both the prompt and the student's response, we can identify a clear alignment between the ESM survey data and the individualized daily journal reflection. This alignment underscores the connection between the student's SE rating from the ESM survey and their reflective processing of that experience the next day.

This example demonstrates temporal alignment because the journal reflection is directly addressing the ESM survey response from the previous day. The alignment is marked by both the timing of the data collection (Week 3, Day 4 vs. Week 3, Day 5) and the temporal language used in both the prompt ("Yesterday, in the app, you told us...") and the student's response ("the first day"). These temporal markers help to link the ESM survey data with the journal reflection, underscoring the student's SE rating from the ESM survey while also offering insight into how that SE was processed and reinterpreted through reflection on the following day.

Example of Temporal Non-Alignment:

An example of identifying a non-alignment between the temporal aspect between an ESM survey response and an individualized daily journal reflection is given below.

Student's ESM Survey: Week 2, Day 2 Notification #27; Open-Ended Task Item Response: Reviewing how to model ODE for my [computational modeling course] in-class assignment for tomorrow; Open-Ended Course Item Response: [computational modeling course]

Corresponding Daily Journal Reflection for Week 2, Day 2 Reflection:

- Prompt: Tell me a story about the last week or today where someone said something that

encouraged or discouraged you. What were you doing? What did they say?

- Response: Last week in [computational modeling course] someone said something that encouraged me that I was doing a good job in class. We had recently gotten new groups to work with in class for the in-class assignments and I was just getting to know these people. A person who had missed the last couple of classes had now joined our group for the first time and we were catching up on things. We were talking about our coding skills and that no one had any past experiences. One person in the group questioned that I did not have any past experiences because he said that I seemed to already understand all the stuff that we were doing. This encouraged me that I was doing well on the in-class assignments and has been a piece of motivation for me when working on these projects and so what.

Although the ESM survey and the individualized daily journal reflection were both collected on Week 2, Day 2, this example illustrates temporal non-alignment due to the time frame the journal reflection is referencing. The prompt asks students to describe an event from "the last week or today," giving them the flexibility to reflect on experiences that occurred outside the immediate time frame of that day's ESM data.

The student explicitly chooses to reflect on a moment from "last week," as indicated by the temporal marker in their journal response: "Last week in [computational modeling course]..." This means that while both data sources were collected on the same day, the student's reflection does not correspond directly to the SE experience captured in the ESM response. Instead, it draws from a past experience for which no ESM data may exist, creating a non-alignment in temporal focus.

This type of non-alignment is analytically meaningful. It highlights how journal prompts designed to elicit broader reflection may surface experiences and meaning-making not captured by time-anchored survey instruments like ESM. These reflections still contribute valuable context and insight into a student's evolving SE but do so without offering a direct link to the corresponding ESM data point.

These examples illustrate how temporal alignment and non-alignment can function as mean-

ingful evidence for linking ESM survey responses with individualized journal reflections. While alignment provides opportunities to track immediate changes in SE, non-alignment surfaces broader reflective insights that may not be captured in real time but still inform our understanding of students' evolving SE experiences.

Alignment Between Open-Ended Task Item Response and Individualized Daily Journal Prompt Response

Evidence Description: This evidence type focuses on whether the content of the open-ended task item response on the ESM survey aligns conceptually or thematically with the student's individualized daily journal prompt response. The alignment here is identified when both sources describe or refer to the same academic task, activity, or experience, either by name (e.g., "physics quiz") or by context (e.g., studying, working on an assignment).

Analytic Reason: Why is this a valid sign of integration? When a student reflects on the same academic task across both the ESM survey and the journal response, it allows for a deeper interpretive connection to be made between what they were doing (task) and how they were experiencing it (SE, emotion, confidence, etc.). This type of alignment is valuable because it lets the analysis trace how students' real-time experiences (ESM) are expanded upon, complicated, or reinterpreted through reflective writing, creating a richer, more integrated understanding of SE in context.

Example of Alignment:

Student's ESM Survey Open-Ended Task Response: Studying for my physics quiz

Corresponding Daily Journal Prompt Response: When I was taking my quiz today I went through a range of emotions. Starting all the way back to when I was studying this morning and walking to class to take the quiz I did not feel very worried at all. I felt calm and confident in myself that I would do well on this quiz, like I have for the past 5 weekly quizzes. Yet, when I had sat down in the lecture hall and overheard people discussing around me...

In this example, both the ESM task response and the journal reflection clearly center around the same academic experience — preparing for and taking a physics quiz. The student uses their

journal to revisit and narrate the same moment referenced in the ESM data, elaborating on how their SE experience evolved from studying to sitting for the quiz. This direct alignment enables the researcher to track shifts or consistencies in the student's SE within a continuous learning episode.

Example of Non-Alignment:

Student's ESM Survey Open-Ended Task Response: My [computational modeling course label] preclass assignment for tomorrow.

Corresponding Daily Journal Prompt Response: Last week in [computational modeling course label] someone said something that encouraged me that I was doing a good job in class. We had recently gotten new groups to work with in class for the in-class assignments and I was just getting to know these people. A person who had missed the last couple of classes had now joined our group for the first time and we were catching them up on things. We were talking about our coding skills and that no one had any past experiences. One person in the group questioned that I did not have any past experiences he said that I seemed to already understand all the stuff that we were doing. This encourage me that I was doing well on the in-class assignments and has been a piece of motivation for me when working on these projects and so what.

Although both sources reference the same course, the task content is not aligned. The ESM response focuses on a preclass assignment due tomorrow, while the journal reflection recounts a social interaction from the previous week during in-class group work, focusing on the in-class assignments instead. While both entries relate to the broader experience of learning in the course, they do not describe or interpret the same task or academic moment. This lack of alignment suggests that the student's journal reflection is drawing on a different memory or experience than the one captured in the ESM response.

Alignment Between Task-Specific Self-Efficacy Item Language and Individualized Daily Journal Reflection

Evidence Description: This evidence type examines whether the language and constructs used in the ESM survey's task-specific SE items (e.g., perceived skill, control, success) are mirrored or directly referenced in the individualized daily journal reflections. This alignment is about

conceptual echoing: students use language or ideas from the ESM questions in their reflective writing, creating a bridge between the survey-based constructs and their own narrative expressions of SE.

Analytic Reason: Why is this a valid sign of integration? This type of alignment reflects conceptual integration across data sources. When students draw on the phrasing, structure, or logic of the ESM SE items in their journal responses, they are actively engaging with the constructs the ESM seeks to measure. It suggests that the prompts are not only eliciting surface-level responses, but are shaping how students think and talk about their experiences—creating a shared interpretive frame. This supports the validity of using both sources together to build a cohesive understanding of students' SE development.

Example of Alignment

Task-Specific Self-Efficacy Survey Items: (1) How skilled are you in the activity?, (2) Do you feel in control of the situation?, (3) Are you succeeding at what you are doing?

Corresponding Daily Journal Reflection:

- Prompt: Looking back on today could you give us a specific example of a success that you were thinking of when answering the questions on the app?

- Response: One example of a success today was, going back to the [computational modeling course label] honors project, figuring out a full plan of what topic I want to do my project on and the steps/data I am going to use to successfully do it. I have been a bit anxious about the project because when I have been looking through the examples, no unique idea stood out to me of what I wanted to do. Yet, this morning I was able to take a topic I brainstormed and make it both much more of my own and applicable to the project's goals. This was an important success to me today because I have been worried that I would not be able to find a solid topic and that would up hurting me when I will start to actually code my project.

In this example, the prompt itself begins by drawing on the language of success, mirroring the third task-specific SE item ("Are you succeeding at what you are doing?") through the phrasing:

"example of a success that you were thinking of when answering the questions on the app." The student then picks up this language in their response: "One example of a success today was, going back to the [computational modeling course] honors project..." They go on to describe this success in detail, emphasizing its personal importance and weaving in SE-relevant language (e.g., "not able to find a solid topic") to contextualize their growth. This alignment demonstrates how the journal reflection builds directly from the construct of task-specific SE introduced by the ESM item, allowing the student to expand on the idea of "success" through reflective elaboration.

Example of Non-Alignment

Task-Specific Self-Efficacy Items: (1) How skilled are you in the activity?, (2) Do you feel in control fo the situation?, (3) Are you succeeding at what you are doing?

Corresponding Daily Journal Reflection:

- Prompt: Yesterday it stood out to us that your confidence in your abilities in [honors research course] was notably higher. Could you tell us what was happening yesterday? What were you doing? Who were you working with? What impacted your sense of confidence?

- Response: What made my confidence get much higher in that course was a combination of two things. One of them was that for an assignment that I was given for the research project I went above and beyond of what was asked of me and did more than just look at some of the data that I was given but I made a full power point presentation about my analysis and conclusion off of this. The second point that gained me confidence from this was that my research mentor was very impressed and made a point to compliment many specific details about both me even putting together a presentation and also how I set it up overall. This all made me gain a lot of confidence in myself and in my work.

Although the reflection is rich in SE content — particularly around confidence—the language and framing are not directly linked to the task-specific item structure from the ESM survey. The student reflects on their success and recognition from others, but the concepts of skill, control, or success as defined by the ESM are not explicitly echoed. This reflects non-alignment at the

223

construct level, even though both sources relate to similar themes. The journal stands more as an independent narrative rather than a direct continuation of the ESM items' framing.

7.9 Discussion

This study aimed to explore the integration of ESM surveys and individualized daily journal reflections in understanding students' SE experiences. Specifically, the research sought to address three key questions:

(RQ1): What does it mean to connect quantitative and qualitative data in this particular mixed methods approach to exploring students' SE?

(RQ2): How does one build linkages between quantitative and qualitative data in practice?

(RQ3): What does it mean for our data to be meaningfully linked in the context of this mixed methods approach?

The results of this study demonstrate that these questions were addressed through showing a detailed examination of the integration process between the ESM survey data and the individualized daily journal reflections.

To answer the first question, I examined how the quantitative data from the ESM surveys and the qualitative data from the individualized journal reflections could be meaningfully connected. This was achieved by identifying several typologies of evidence, which can be more broadly summarized as temporal alignment, content alignment, and language alignment, that served as the backbone for linking these data sources.

In response to the second research question, the process of building linkages between data sources was originally articulated through the "following the thread" approach described in Chapter 5. In that approach, specific ESM responses were used to generate individualized journal prompts, allowing for a direct, participant-centered connection between quantitative and qualitative data points. This study extends that work by not only tracing these threads between specific data moments, but also by examining how the nature of the data produced—through both the ESM responses and the journal reflections—can support different kinds of interpretive linkages. In doing so, the analysis moves beyond simply connecting one data source to another and begins to

articulate how the quality and character of the evidence across sources shapes the way those threads are followed and understood.

The third research question was addressed by demonstrating that the data could indeed be meaningfully linked, albeit not always in a straightforward or binary manner. Instead of assuming a direct connection between the data sources, I highlighted the nuanced moments of alignment and non-alignment, showing that both convergences and dissonances provided rich analytical insights into students' SE experiences. With these research questions addressed, the discussion now turns to the implications of these findings and suggestions for future work.

7.9.1 Connecting the Process of "Following the Thread" to Evidence Types

One of the central contributions of this work is the novel approach of linking the "following the thread" method (Moran-Ellis et al., 2006) with a typology of evidence types. This dual approach offers a structured yet flexible framework for integrating quantitative and qualitative data. By outlining three distinct types of evidence that support integration—such as temporal alignment, alignment with the task-opened survey item, and alignment with the language of the task-specific SE items—this study adds clarity to the often murky process of determining when and how to meaningfully connect data from different sources.

This contribution is significant because it extends the idea of "following the thread" beyond a heuristic method to a methodological tool that can be applied in other mixed methods research contexts. The typology provides a systematic way of understanding how these different strands of data can interact, enriching our analysis and ensuring that connections are not only intuitive but also well-justified with specific evidence.

In doing so, this work also provides practical guidance on how to track, evaluate, and substantiate cross-source connections, addressing a critical gap in the mixed methods literature where integration is often discussed in theory but not always operationalized in practice.

7.9.2 Clarifying the Relationship Between Threads and Linkages

As I reflect on the analytic process and the methodological choices made in this study, I recognize an important conceptual distinction that remains in development: the difference between

225

a thread and a linkage. While the two are often used interchangeably in this work, I believe they represent different stages or elements of the integration process.

A thread, as conceptualized here, begins during data collection – particularly when following the process outlined in Chapter 5, where individualized daily journal prompts are constructed based on students' earlier ESM responses. A well-constructed thread is intended to elicit a meaningful continuation or complication of an earlier experience, ideally resulting in a strong foundation for these data sources and data to be integrated in the analysis state of the study.

In contrast, a linkage refers to the product of the integration process – the analytical connection made between an ESM data point and a daily journal reflection. In this sense, a thread is a potential pathway to a linkage, but not all threads may result in successful or justifiable linkages. A thread becomes a linkage when it is successfully followed through and supported by evidence that integrated was produced during the analysis.

At this stage in the work, we were not always able to clearly distinguish between these concepts in practice, in part because this study represents an early iteration of our efforts to trace and evaluate the connections systematically. Future work should further interrogate the relationship between threads and linkages, both conceptually and methodologically, to more fully articulate the mechanics of integration and to refine how we design and analyze studies using mixed methods in this way.

7.9.3 Investigating the Interaction of Evidence Types in Self-Efficacy Research

This study provided valuable insights by examining individual evidence types, but several important areas for future research remain. Specifically, while the current work explored each evidence type in isolation, future studies could focus on how different evidence types interact and build upon one another when integrating ESM surveys and individualized daily journal reflections.

One key avenue for future exploration would involve considering how various evidence types—such as temporal alignment, conceptual congruence, or task-specific SE language—work together to form a more nuanced understanding of how SE is experienced and represented. It would be useful to investigate whether certain combinations of evidence types more strongly suggest a valid link

between ESM surveys and journal reflections, thus informing researchers about which types of evidence are most effective in constructing a coherent narrative of SE.

Furthermore, the interaction between evidence types could be studied in relation to alignment and non-alignment between the data sources. How do alignments and non-alignments between different typologies of evidence—whether temporal, conceptual, or task-based—shape the interpretations we make about SE? For instance, an alignment between task-specific language in both the ESM and journal reflections might strengthen claims about how students experience SE in academic contexts. Conversely, non-alignments—where discrepancies or contradictions between the data sources emerge—may highlight complex, evolving experiences of SE that deserve closer attention.

The way these evidence types work together, or do not, has significant implications for the kinds of claims that can be made about SE in future studies. If certain patterns of evidence consistently suggest a more integrated view of SE, researchers may be able to develop stronger, more grounded theories of SE that account for the dynamic and multifaceted nature of student experiences. Moreover, exploring the impact of these different types of evidence on the integration process could further illuminate how mixed methods approaches can be optimally employed in SE research, ultimately contributing to more robust and context-sensitive understandings of SE development.

In sum, while this study laid the groundwork for understanding how to integrate different data sources to explore SE, future work should focus on investigating how combinations of evidence types may better support or complicate the integration process. By addressing these questions, researchers can develop more comprehensive strategies for linking qualitative and quantitative data, and gain deeper insights into how SE unfolds across multiple contexts.

7.9.4 Limitations

While this study provides important insights into the integration of ESM surveys and individualized journal reflections in exploring SE, one of the primary limitations is that the analysis focused primarily on a specific set of evidence types, without exhaustively exploring all potential

types that could arise from the combination of ESM surveys and daily journal reflections. Each of the evidence types explored in this study performs distinct conceptual work in examining SE. For example, the alignment or non-alignment between task-specific SE items in the ESM survey and their mirrored language in the daily journal reflection (prompt and response) serves to investigate the domain-specific nature of SE through the ability to explore task-specific aspects of SE. In contrast, the temporal evidence highlights the dynamic nature of SE over time. This distinction suggests that different types of evidence may align with various theoretical foundations of SE. By focusing solely on the three evidence types discussed here, the study may have limited its ability to comprehensively explore the broader theoretical foundations of SE. Future research that expands on these evidence types and their interactions could provide a more nuanced understanding of the multiple dimensions of SE.

7.10 Conclusion

Taken together, this study contributes a novel methodological framework for integrating ESM and individualized journal reflections, grounded in a typology of evidence that enables richer, more nuanced interpretations of SE. While limited in scope, the findings offer a starting point for future mixed methods research seeking to operationalize integration in ways that are both systematic and theoretically meaningful. As researchers continue to explore the dynamic nature of SE, approaches like the one presented here can help foreground the complexity of student experiences while offering rigor in how those experiences are understood across data sources.

CHAPTER 8

CONCLUSIONS AND FUTURE WORK

This dissertation focused on how to design a mixed methods research approach to explore the real-time experiences shaping students' self-efficacy. Specifically, the way I approached this was through focusing on how to integrate the Experience Sampling Method with individualized daily journal reflections as a mean to be able to make visible how students' experiences within physics and more broadly STEM are shaping their SE within particular academic moments.

The process began with a preliminary analysis to understand how ESM and individualized daily journal reflections were communicating with each other to explore students' SE. This early work revealed that to better capture the richness of students' lived experiences and their SE, I needed to modify the original mixed methods approach. This led to the addition of an integration point, where I intentionally set up communication between the open-ended ESM survey items (qualitative) and the closed-ended ESM survey items (quantitative). This modification opened new possibilities for understanding how these data sources could inform one another and, ultimately, how they could be more meaningfully linked to enhance our understanding of SE.

These larger questions, framed in Chapter 1, set the stage for the methodological exploration that follows. They concern the very foundations of how knowledge is constructed and validated within the field of physics education research (PER), and they speak directly to the epistemological and methodological commitments that underlie this dissertation. In exploring the integration of ESM surveys and individualized journal reflections, I have not only built a case for how mixed methods can provide richer insights into SE, but also contributed to answering these broader questions by demonstrating how we might bridge the gap between different kinds of evidence in a way that meaningfully advances our understanding.

8.0.1 Contributions and Insights

The results of this study show that integrating qualitative and quantitative data is not just a technical endeavor, but one that has significant implications for how we theorize and understand the constructs we study — in this case, self-efficacy. In Chapter 7, I demonstrated how the process

of "following the thread" between the ESM surveys and the journal responses is not merely a method for linking two distinct types of data, but a process that generates new insights into the dynamic, temporal, and context-specific nature of SE. This methodological contribution advances our understanding of how mixed methods can be used to capture complex phenomena in a way that honors both the richness of lived experience (through qualitative data) and the precision of repeated measurement (through quantitative data).

This work also interrogates the theoretical underpinnings of SE. By tracing the connections between the evidence types—temporal alignment, content alignment, and language alignment—this dissertation highlights the different theoretical work each evidence type performs, thus offering a more nuanced and comprehensive understanding of SE. By focusing not only on what the data tells us, but how we link and interpret these data sources, this study underscores the importance of methodological reflexivity in the construction of knowledge in physics education research.

8.0.2 Reflecting on Methodological Innovation

A central contribution of this dissertation is the development of a "following the thread" approach to integrating data sources. This technique offers a practical, systematic tool for mixed methods research that can be applied across other areas of educational research, not just in the study of SE. This innovation aligns with the broader question of how we know what we know in physics education research—by making clear how methodological rigor and reflexivity can lead to more grounded and meaningful interpretations.

The study also makes important strides in refining the application of mixed methods approaches in educational research. Specifically, it addresses a key gap in the literature: the operationalization of data integration in mixed methods designs. The typology of evidence types introduced here provides a framework for understanding the different ways in which data from distinct sources can be linked and interpreted. This framework is not only an empirical contribution, but a methodological one, advancing how researchers design and analyze complex educational data.

8.0.3 Looking Ahead: Open Questions and Future Research

While this study offers important insights, it is far from exhaustive. As discussed in Chapter 7, future research should focus on how different evidence types interact, and how their integration can support a more nuanced and layered understanding of self-efficacy (SE). This work also invites further conceptual development of the distinction between threads and linkages—two analytic constructs that need more precise definition and testing through continued empirical and methodological exploration.

There are several promising directions for future research that build upon the findings and methodological innovations of this study. These include refining the analytical tools developed here, expanding the temporal scope of inquiry, and investigating the potential of reflective journaling as an intervention. Specifically:

Enhancing the Codebook and Analytical Framework: While the codebook provided a strong foundation for analyzing journal prompts and students' reflections, future iterations could expand its use to explore how students actually interpret and respond to prompts in practice. This includes refining how we code for emotional and physiological expressions, and investigating how such expressions map—or fail to map—onto Bandura's conception of physiological state as a source of SE. Additionally, developing a framework for tailoring prompts even more precisely to individual students' experiences could deepen the work of personalization and contextual sensitivity.

Temporal Aspects of Self-Efficacy: Given the daily nature of the journal reflections and the real-time capture of ESM, there is an opportunity to investigate SE as a dynamic, fluctuating phenomenon. Future longitudinal research could examine how students' SE evolves across days, weeks, or even semesters, providing insight into how academic experiences build or erode confidence over time—and when interventions might be most needed.

Further Integration of ESM and Journal Prompts: Another promising direction involves exploring additional forms of integration between ESM data and individualized daily journal reflections. In this study, I brought together established ideas from the mixed methods community—such as points of integration and the "following a thread" approach—and operationalized the notion that

integration is not only a process but also a product: a process of gathering evidence to demonstrate that integration has meaningfully occurred. Future work could investigate other forms of integration and their analytic value—whether they serve to confirm, extend, or complicate emerging narratives around constructs like self-efficacy. Such explorations would not only sharpen the interpretive tools available to researchers, but also help to further legitimize integration as a methodological contribution in its own right within mixed methods research.

Broader Applications of This Approach: The design and methodological framework developed in this dissertation hold potential for application in other contexts. Future studies might extend this mixed methods approach to graduate education, or to explore other constructs such as identity, belonging, or motivation—each of which shares theoretical terrain with SE but presents unique analytical challenges and opportunities.

Legitimizing Mixed Methods Research: Finally, this work contributes to ongoing efforts to legitimize and refine mixed methods research in education. It underscores the value of being explicit and intentional about integration—not as an afterthought or superficial combination of data types, but as a meaningful, generative, and theoretically motivated process. Continued work in this direction can help elevate the standards of methodological rigor and creativity across the field.

Moreover, the broader questions raised in Chapter 1 continue to guide this work: What does it mean to do physics education research? and How do we know what we know as a field? These are not merely philosophical provocations—they are central to the work of knowledge-making. This dissertation contributes to these conversations by modeling an integrative, reflexive, and methodological approach to studying student experience. By bringing qualitative and quantitative data into closer conversation, I offer one possible way forward: an approach that not only studies self-efficacy but also challenges and reimagines how we do research in physics education more broadly.

Final Thoughts In sum, this dissertation has demonstrated that integrating qualitative and quantitative methods—through the careful design and thoughtful application of a mixed methods approach—can offer rich, meaningful insights into self-efficacy and beyond. It has contributed

both empirically and methodologically, offering a new way to think about and engage with the data we collect. More importantly, it has shown that answering the fundamental questions of what it means to do physics education research and how we know what we know requires not only the right theoretical grounding but also the right tools to connect the dots across different kinds of evidence. As the field of physics education research continues to grow, this work lays the foundation for further exploration into how we can meaningfully integrate different data sources to build a more nuanced, robust understanding of student learning and development.

8.1 Final Reflections/Closing Thoughts

In conducting mixed methods research, it is crucial for researchers to be intentional about both the why and how they integrate quantitative and qualitative data. Mixed methods research is not simply about collecting two types of data and then synthesizing the results—it requires a thoughtful and strategic approach to integration at every stage of the study.

Integration should not be an afterthought or a mere appendage to the research design. Instead, it should be embedded throughout the entire methodology, from the research questions and design to the data collection and analysis processes. Researchers must be deliberate about where integration occurs, why integration is occurring at each step, and what this integration looks like in practice.

In this study, we aimed to demonstrate that a clear, coherent integration of both qualitative and quantitative methods can provide deeper insights into students' SE experiences. The ability to integrate findings from different data sources—and to understand the connections between them—allowed for a more nuanced understanding of how SE manifests and evolves in the context of students' academic experiences.

By emphasizing intentional integration, we not only enrich the findings but also contribute to the broader discussion of how mixed methods research can be applied effectively in educational settings. As researchers continue to explore complex constructs like SE with mixed methods research, it is essential to think critically about the ways in which multiple quantitative and qualitative data sources can complement one another, rather than simply existing side by side.

Furthermore, as researchers delve into constructs as dynamic and context-dependent as SE, it

is important to reflect on how methodological choices impact not just the findings, but also the interpretations and implications that arise from those findings. Our work showcases that thoughtful, intentional integration is not merely a methodological choice, but a fundamental aspect of how we generate knowledge in educational research.

BIBLIOGRAPHY

[1] S. Secules, *On the importance of (white) humility: Epistemological decentering as a positional orientation toward research*, Apr. 2023. DOI: 10.1002/jee.20508.

[2] K. M. T. Collins, "Validity in multimethod and mixed research," in *Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*, S. N. Hesse-Biber and B. Johnson, Eds. Oxford University Press, Incorporated, Jul. 2015, pp. 240–256, In Collins 2015 work (i.e., Validity in Multimethod and Mixed Research), it is unclear to me how she identified these dimensions of high-quality research in MMMR. My gut instinct or thought would be that she read many papers and noticed that high quality research papers included the following criteria.

[3] J. C. Greene, "Preserving distinctions within the multimethod and mixed methods research merger," in *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry*, S. N. Hesse-Biber and R. B. Johnson, Eds. Oxford University Press, 2015. [Online]. Available: http://iom.edu/Reports..

[4] D. Hammer, "The necessarily, wonderfully unsettled state of methodology in per: A reflection," in *The International Handbook of Physics Education Research: Special Topics*, M. R. F. T. P. L. Heron, Ed. AIP Publishing, 2023, pp. 1–12.

[5] N. Mackenzie and S. Knipe, "Research dilemmas: Paradigms, methods and methodology," *Issues in educational research*, vol. 16, pp. 193–205, 2 2006.

[6] S. Irwin, "Data analysis and interpretation emergent issues in linking qualitative and quantitative evidence," in *Handbook of Emergent Methods*, S. N. Hesse-Biber and P. Leavy, Eds. Guilford Publications, 2008, pp. 415–436, ISBN: 9781593851477.

[7] A. V. Knaub and L. Ding, "Research design concerning equity in physics education research (per)," in *The International Handbook of Physics Education Research: Special Topics*, M. F. Taşar and P. R. L. Heron, Eds. AIP Publishing, 2023, pp. 1–28.

[8] M. Borrego, A. Chasen, H. C. Tripp, E. Landgren, and E. Koolman, *A scoping review on u.s. undergraduate students with disabilities in stem courses and stem majors*, Dec. 2025. DOI: 10.1186/s40594-024-00522-2.

[9] A. Chasen and M. A. Pfeifer, "Empowering disabled voices: A practical guide for methodological shifts in biology education research," *CBE Life Sciences Education*, vol. 23, 3 Sep. 2024, ISSN: 19317913. DOI: 10.1187/cbe.24-02-0076.

[10] S. Kanim and X. C. Cid, "Demographics of physics education research," *Physical Review Physics Education Research*, vol. 16, 2 Dec. 2020, ISSN: 24699896. DOI: 10.1103/PhysRevPhysEducRes.16.020106.

[11] E. L. Usher, "Acknowledging the whiteness of motivation research: Seeking cultural relevance," *Educational Psychologist*, vol. 53, pp. 131–144, 2 Apr. 2018, ISSN: 00461520. DOI: 10.1080/00461520.2018.1442220.

[12] L. A. Wood and V. Sawtelle, "Narrative analysis of a woman's experience transferring from a tyc reveals impact of supporting characters," *CBE Life Sciences Education*, vol. 21, 3 2022, ISSN: 19317913. DOI: 10.1187/cbe.21-09-0253.

[13] J. H. Hitchcock and A. J. Onwuegbuzie, "The routledge handbook for advancing integration in mixed methods research an introduction," in *THE ROUTLEDGE HANDBOOK FOR ADVANCING INTEGRATION IN MIXED METHODS RESEARCH*, J. H. Hitchcock and A. J. Onwuegbuzie, Eds. Taylor Francis, 2022, pp. 3–27.

[14] M. Swirtz and R. Barthelemy, "Queering methodologies in physics education research," in *Physics Education Research Conference Proceedings*, American Association of Physics Teachers, 2022, pp. 457–462, ISBN: 9781931024389. DOI: 10.1119/perc.2022.pr.Swirtz.

[15] M. Boveda and S. A. Annamma, *Beyond making a statement: An intersectional framing of the power and possibilities of positioning*, Jun. 2023. DOI: 10.3102/0013189X231167149.

[16] K. F. Jackson *et al.*, *Positionality in critical feminist scholarship: Situating social locations and power within knowledge production*, Feb. 2024. DOI: 10.1177/08861099231219848.

[17] S. Secules *et al.*, "Positionality practices and dimensions of impact on equity research: A collaborative inquiry and call to the community," *Journal of Engineering Education*, vol. 110, pp. 19–43, 1 Jan. 2021, ISSN: 21689830. DOI: 10.1002/jee.20377.

[18] A. Bandura, *Self-Efficacy: The Exercise of Control*. W.H.Freeman and Company, 1997.

[19] V. Sawtelle, E. Brewe, and L. H. Kramer, "Exploring the relationship between self-efficacy and retention in introductory physics," *Journal of Research in Science Teaching*, vol. 49, pp. 1096–1121, 9 Nov. 2012, ISSN: 00224308. DOI: 10.1002/tea.21050.

[20] E. L. Usher and F. Pajares, *Sources of self-efficacy in school: Critical review of the literature and future directions*, 2008. DOI: 10.3102/0034654308321456.

[21] A. L. Zeldin and F. Pajares, "Against the odds: Self-efficacy beliefs of women in mathematical, scientific, and technological careers," Tech. Rep., 2000, pp. 215–246.

[22] S. Hood *et al.*, ""I Like and Prefer to Work Alone": Social Anxiety, Academic Self-Efficacy, and Students' Perceptions of Active Learning," *CBE—Life Sciences Education*, vol. 20, no. 1, ar12, Mar. 2021, Publisher: American Society for Cell Biology (lse). DOI: 10.1187/cbe.19-12-0271. [Online]. Available: https://www.lifescied.org/doi/full/10.1187/cbe.19-12-0271 (visited on 11/03/2021).

[23] T. Luo, W. W. M. So, W. C. Li, and J. Yao, "The Development and Validation of a Survey for Evaluating Primary Students' Self-efficacy in STEM Activities," en, *Journal of Science Education and Technology*, vol. 30, no. 3, pp. 408–419, Jun. 2021, ISSN: 1059-0145, 1573-1839. DOI: 10.1007/s10956-020-09882-0. [Online]. Available: https://link.springer.com/10.1007/s10956-020-09882-0 (visited on 10/21/2021).

[24] D. H. Schunk and M. K. Dibenedetto, "Academic self-efficacy," in *Handbook of Positive Psychology in Schools*, M. J. Furlong, R. Gilman, and E. S. Huebner, Eds., 2nd ed. Taylor Francis, 2014, pp. 115–130. [Online]. Available: https://www.ebsco.com/terms-of-use.

[25] J. M. Nissen and J. T. Shemwell, "Gender, experience, and self-efficacy in introductory physics," *Physical Review Physics Education Research*, vol. 12, 2 2016, ISSN: 24699896. DOI: 10.1103/PhysRevPhysEducRes.12.020105.

[26] J. M. Nissen, "Gender differences in self-efficacy states in high school physics," *Physical Review Physics Education Research*, vol. 15, 1 Mar. 2019, ISSN: 24699896. DOI: 10.1103/PhysRevPhysEducRes.15.013102.

[27] S. J. Sedberry, I. D. Beatty, W. J. Gerace, M. A. Elobeid, and J. E. Strickhouser, "Three critical issues that shape and complicate stem self-efficacy intervention research: Reflections and analysis from an interdisciplinary research team," in *Physics Education Research Conference Proceedings*, American Association of Physics Teachers, 2019, pp. 541–547, ISBN: 9781931024365. DOI: 10.1119/perc.2019.pr.Sedberry.

[28] B. J. Zimmerman, "Self-efficacy: An essential motive to learn," *Contemporary Educational Psychology*, vol. 25, pp. 82–91, 1 2000, ISSN: 0361476X. DOI: 10.1006/ceps.1999.1016.

[29] F. Pajares, "Self-efficacy beliefs in academic settings," Tech. Rep., 1996, pp. 543–578.

[30] S. Cwik and C. Singh, "Not feeling recognized as a physics person by instructors and teaching assistants is correlated with female students' lower grades," *Physical Review Physics Education Research*, vol. 18, 1 Jun. 2022, ISSN: 24699896. DOI: 10.1103/PhysRevPhysEducRes.18.010138.

[31] S. Cwik and C. Singh, "Damage caused by societal stereotypes: Women have lower physics self-efficacy controlling for grade even in courses in which they outnumber men," *Physical Review Physics Education Research*, vol. 17, 2 Dec. 2021, ISSN: 24699896. DOI: 10.1103/PhysRevPhysEducRes.17.020138.

[32] J. Durk, A. Davies, R. Hughes, and L. Jardine-Wright, "Impact of an active learning physics workshop on secondary school students' self-efficacy and ability," *Physical Review Physics Education Research*, vol. 16, 2 Oct. 2020, ISSN: 24699896. DOI: 10.1103/PhysRevPhysEducRes.16.020126.

[33] T. Espinosa, K. Miller, I. Araujo, and E. Mazur, "Reducing the gender gap in students'

physics self-efficacy in a team- and project-based introductory physics class," *Physical Review Physics Education Research*, vol. 15, 1 May 2019, ISSN: 24699896. DOI: 10.1103/PhysRevPhysEducRes.15.010132.

[34] H. Fend and K. Scheel, "Engaging students an examination of the effects of teaching strategies on self-efficacy and course climate in a nonmajors physics course," *Journal of College Science Teaching*, vol. 35, pp. 20–24, 2005.

[35] C. Lindstrøm and M. D. Sharma, "Self-efficacy of first year university physics students: Do gender and prior formal instruction in physics matter?" Tech. Rep., 2011, pp. 1–19.

[36] R. Freed, D. McKinnon, M. Fitzgerald, and C. M. Norris, "Development and validation of an astronomy self-efficacy instrument for understanding and doing," *Physical Review Physics Education Research*, vol. 18, 1 Jun. 2022, ISSN: 24699896. DOI: 10.1103/PhysRevPhysEducRes.18.010117.

[37] E. Palmgren, K. Tuominen, and I. Kontro, "Self-efficacy and conceptual knowledge in quantum mechanics during teaching reforms and the covid-19 pandemic," *Physical Review Physics Education Research*, vol. 18, 2 Jul. 2022, ISSN: 24699896. DOI: 10.1103/PhysRevPhysEducRes.18.020122.

[38] R. Dou, E. Brewe, J. P. Zwolak, G. Potvin, E. A. Williams, and L. H. Kramer, "Beyond performance metrics: Examining a decrease in students' physics self-efficacy through a social networks lens," *Physical Review Physics Education Research*, vol. 12, 2 2016, ISSN: 24699896. DOI: 10.1103/PhysRevPhysEducRes.12.020124.

[39] V. Sawtelle, E. Brewe, and L. H. Kramer, "Positive impacts of modeling instruction on self-efficacy," in *AIP Conference Proceedings*, vol. 1289, 2010, pp. 289–292, ISBN: 9780735408449. DOI: 10.1063/1.3515225.

[40] B. J. Zimmerman and A. Kitsantas, "Homework practices and academic achievement: The mediating role of self-efficacy and perceived responsibility beliefs," *Contemporary Educational Psychology*, vol. 30, pp. 397–417, 4 Oct. 2005, ISSN: 0361476X. DOI: 10.1016/j.cedpsych.2005.05.003.

[41] N. Naibert, K. D. Duck, M. M. Phillips, and J. Barbera, "Multi-institutional study of self-efficacy within flipped chemistry courses," *Journal of Chemical Education*, vol. 98, pp. 1489–1502, 5 May 2021, ISSN: 19381328. DOI: 10.1021/acs.jchemed.0c01361.

[42] P. R. Pintrich, "A manual for the use of the motivated strategies for learning questionnaire (mslq). institution national center for research to improve postsecondary teaching and learning," Tech. Rep., 1991, This piece is essentially essential because this is where they publish the MSLQ, and one can find the entire assessment to provide to participants including the demographics and MSLQ items along with the scales.

[43] C. S. Hung and H. K. Wu, "Tenth graders' problem-solving performance, self-efficacy, and perceptions of physics problems with different representational formats," *Physical Review Physics Education Research*, vol. 14, 2 Nov. 2018, ISSN: 24699896. DOI: 10.1103/PhysRe vPhysEducRes.14.020114.

[44] K. Miller, N. Lasry, B. Lukoff, J. Schell, and E. Mazur, "Conceptual question response times in peer instruction classrooms," *Physical Review Special Topics - Physics Education Research*, vol. 10, 2 Aug. 2014, ISSN: 15549178. DOI: 10.1103/PhysRevSTPER.10.020113.

[45] K. Miller, J. Schell, A. Ho, B. Lukoff, and E. Mazur, "Response switching and self-efficacy in peer instruction classrooms," *Physical Review Special Topics - Physics Education Research*, vol. 11, 1 Feb. 2015, ISSN: 15549178. DOI: 10.1103/PhysRevSTPER.11.010104.

[46] K. Miller, T. Espinosa, I. Araujo, and I. Gallegos, "Response-shift bias in student self-efficacy during an actively taught physics course," *Physical Review Physics Education Research*, vol. 19, 2 Dec. 2023, ISSN: 24699896. DOI: 10.1103/physrevphyseducres.19.02 0167.

[47] D. J. Rosen and A. M. Kelly, "Working together or alone, near, or far: Social connections and communities of practice in in-person and remote physics laboratories," *Physical Review Physics Education Research*, vol. 18, 1 Jun. 2022, ISSN: 24699896. DOI: 10.1103/PhysRe vPhysEducRes.18.010105.

[48] A. M. Cavallo, W. H. Potter, and M. Rozman, "Gender differences in learning constructs, shifts in learning constructs, and their relationship to course achievement in a structured inquiry, yearlong college physics course for life science majors," *School Science and Mathematics*, vol. 104, pp. 288–300, 6 Oct. 2004, ISSN: 19498594. DOI: 10.1111/j.1949-8594.2004.tb18000.x.

[49] 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," *Journal of Science Education and Technology*, vol. 29, pp. 758–773, 6 Dec. 2020, ISSN: 15731839. DOI: 10.1007/s10956-020-09853-5.

[50] L. Ainscough *et al.*, "Changes in biology self-efficacy during a first-year university course," *CBE Life Sciences Education*, vol. 15, 2 Jun. 2016, ISSN: 19317913. DOI: 10.1187/cbe.15 -04-0092.

[51] R. Henderson, V. Sawtelle, and J. M. Nissen, "Gender self-efficacy: A call to physics educators," *The Physics Teacher*, vol. 58, pp. 345–348, 5 2020.

[52] J. M. Majer, "Self-efficacy and academic success among ethnically diverse first-generation community college students," *Journal of Diversity in Higher Education*, vol. 2, pp. 243–250, 4 Dec. 2009, ISSN: 19388926. DOI: 10.1037/a0017852.

[53] N. E. Betz and G. Hackett, "The relationship of career-related self-efficacy expectations to perceived career options in college women and men," *Journal of Counseling Psychology*, vol. 28, pp. 399–410, 5 1981, ISSN: 00220167. DOI: 10.1037/0022-0167.28.5.399.

[54] D. H. Schunk and M. K. Dibenedetto, "Self-efficacy theory in education," in *Handbook of Motivation at School*, K. R. Wentzel and D. B. Miele, Eds., 2nd ed. Taylor Francis Group, 2016, pp. 34–54. [Online]. Available: http://ebookcentral.proquest.com/lib/michstate-ebooks/detail.action?docID=4415668..

[55] 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*, vol. 45, pp. 955–970, 8 Oct. 2008, ISSN: 00224308. DOI: 10.1002/tea.20249.

[56] R. M. Klassen, *A cross-cultural investigation of the efficacy beliefs of south asian immigrant and anglo canadian nonimmigrant early adolescents*, Dec. 2004. DOI: 10.1037/0022-0663.96.4.731.

[57] E. L. Usher and F. Pajares, "Sources of academic and self-regulatory efficacy beliefs of entering middle school students," *Contemporary Educational Psychology*, vol. 31, pp. 125–141, 2 Apr. 2006, ISSN: 0361476X. DOI: 10.1016/j.cedpsych.2005.03.002.

[58] E. L. Usher and F. Pajares, "Sources of self-efficacy in mathematics: A validation study," *Contemporary Educational Psychology*, vol. 34, pp. 89–101, 1 Jan. 2009, ISSN: 0361476X. DOI: 10.1016/j.cedpsych.2008.09.002.

[59] A. Bandura, "Self-efficacy: Toward a unifying theory of behavioral change," Tech. Rep., 1977, pp. 191–215.

[60] A. Tashakkori and C. Teddlie, "Integrating qualitative and quantitative approaches to research," in *The SAGE Handbook of Applied Social Research Methods*. SAGE Publications, Inc., Feb. 2014, pp. 283–317. DOI: 10.4135/9781483348858.n9.

[61] M. Sandelowski, "Tables or tableaux? the challenges of writing and reading mixed methods studies," in *Handbook of Mixed Methods in social and behavioral research*. SAGE Publications, 2003, pp. 321–350, ISBN: 9780761920731.

[62] M. D. Fetters, L. A. Curry, and J. W. Creswell, "Achieving integration in mixed methods designs - principles and practices," *Health Services Research*, vol. 48, pp. 2134–2156, 6 PART2 2013, ISSN: 14756773. DOI: 10.1111/1475-6773.12117.

[63] M. D. Fetters and J. F. Molina-Azorin, *The journal of mixed methods research starts a new decade: The mixed methods research integration trilogy and its dimensions*, Jul. 2017. DOI: 10.1177/1558689817714066.

[64] J. W. Creswell and V. L. P. Clark, "Core mixed methods designs," in *Designing and*

Conducting Mixed Methods Research, 3rd ed. Sage, 2018, pp. 77–84.

[65]   J. C. Greene, V. J. Caracelli, and W. F. Graham, "Toward a conceptual framework for mixed-method evaluation designs," Tech. Rep., 1989, pp. 255–274.

[66]   Y. Zhou and M. L. Wu, "Reported methodological challenges in empirical mixed methods articles: A review on jmmr and ijmra," *Journal of Mixed Methods Research*, vol. 16, pp. 47–63, 1 Jan. 2022, ISSN: 15586901. DOI: 10.1177/1558689820980212.

[67]   S. Mathison, "Why triangulate?" *Educational researcher*, vol. 17, pp. 569–574, 2 1988.

[68]   M. Sandelowski, *Triangles and crystals: On the geometry of qualitative research.* 1995. DOI: 10.1002/nur.4770180612.

[69]   A. O'Cathain, E. Murphy, and J. Nicholl, "Ocathain2010$_three techniques for integrating data in mixed me$ *The BMJ*, 2010. [Online]. Available: https://doi.org/10.1136/bmj.c4587.

[70]   V. L. P. Clark, *Meaningful integration within mixed methods studies: Identifying why, what, when, and how*, Apr. 2019. DOI: 10.1016/j.cedpsych.2019.01.007.

[71]   M. D. Fetters and D. Freshwater, *The 1 + 1 = 3 integration challenge*, Apr. 2015. DOI: 10.1177/1558689815581222.

[72]   J. Moran-Ellis *et al.*, "Triangulation and integration: Processes, claims and implications," *Qualitative Research*, vol. 6, pp. 45–59, 1 Feb. 2006, ISSN: 14687941. DOI: 10.1177/1468 794106058870.

[73]   C. M. Dupin and G. Borglin, *Usability and application of a data integration technique (following the thread) for multi- and mixed methods research: A systematic review*, Aug. 2020. DOI: 10.1016/j.ijnurstu.2020.103608.

[74]   A. Bryman, "Integrating quantitative and qualitative research: How is it done?" *Qualitative Research*, vol. 6, no. 1, pp. 97–113, Feb. 2006, Publisher: SAGE Publications, ISSN: 1468-7941. DOI: 10.1177/1468794106058877. [Online]. Available: https://doi.org/10.1177/14 68794106058877 (visited on 01/19/2022).

[75]   J. Dalgety and R. K. Coll, "Exploring first-year science students' chemistry self-efficacy," *International journal of science and mathematics education*, vol. 4, pp. 97–116, 1 2006.

[76]   Y. Li and C. Singh, "Do female and male students' physics motivational beliefs change in a two-semester introductory physics course sequence?" *Physical Review Physics Education Research*, vol. 18, 1 Jun. 2022, ISSN: 24699896. DOI: 10.1103/PhysRevPhysEducRes.18.010142.

[77]   J. A. Maxwell, *Qualitative Research Design: An Interactive Approach*, 3rd ed. Sage, 2013.

[78] J. W. Creswell, *Qualitative Inquiry Research Design: Choosing Among Five Approaches*, 2nd ed. Sage Publications, 2007.

[79] M. Csikszentmihalyi and R. Larson, "Validity and reliability of the experience-sampling method," *The Journal of Nervous and Mental Disease*, vol. 175, 9 1987.

[80] J. M. Hektner, J. A. Schmidt, and M. Csikszentmihalyi, *Experience Sampling Method Measuring the Quality of Everyday Life*. Sage Publications, 2007.

[81] M. Mehl and T. Conner, Eds., *Handbook of Research Methods for Studying Daily Life*. The GUILDFORD PRESS, 2012.

[82] S. Zirkel, J. A. Garcia, and M. C. Murphy, "Experience-sampling research methods and their potential for education research," *Educational Researcher*, vol. 44, pp. 7–16, 1 Jan. 2015, ISSN: 1935102X. DOI: 10.3102/0013189X14566879.

[83] S. Shiffman, A. A. Stone, and M. R. Hufford, *Ecological momentary assessment*, 2008. DOI: 10.1146/annurev.clinpsy.3.022806.091415.

[84] N. V. Berkel, D. Ferreira, and V. Kostakos, "The experience sampling method on mobile devices," *ACM Computing Surveys*, vol. 50, 6 Nov. 2017, ISSN: 15577341. DOI: 10.1145/3123988.

[85] T. Steenbergh, J. Runyan, D. Daugherty, and L. Oke, *Lifedata*.

[86] L. Tay, *Expiwell*, 2022.

[87] R. Henderson and V. Sawtelle, "Implementing a mixed-methods approach to understand students' self-efficacy: A pilot study," in *Physics Education Research Conference Proceedings*, American Association of Physics Teachers, 2020, pp. 204–209, ISBN: 9781931024372. DOI: 10.1119/perc.2020.pr.Henderson.

[88] 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*, vol. 23, pp. 399–409, 3 Aug. 2015, ISSN: 15524590. DOI: 10.1177/1069072714547322.

[89] S. M. Glynn, G. Taasoobshirazi, and P. Brickman, "Science motivation questionnaire: Construct validation with nonscience majors," *Journal of Research in Science Teaching*, vol. 46, pp. 127–146, 2 Feb. 2009, ISSN: 00224308. DOI: 10.1002/tea.20267.

[90] 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*, vol. 47, pp. 564–582, 5 May 2010, ISSN: 00224308. DOI: 10.1002/tea.20353.

[91] M. Richardson, C. Abraham, and R. Bond, "Psychological correlates of university students' academic performance: A systematic review and meta-analysis.," en, *Psychological Bulletin*, vol. 138, no. 2, pp. 353–387, 2012, ISSN: 1939-1455, 0033-2909. DOI: 10.1037/a002 6838. [Online]. Available: http://doi.apa.org/getdoi.cfm?doi=10.1037/a0026838 (visited on 10/30/2021).

[92] S. Lau and R. W. Roeser, "Cognitive abilities and motivational processes in high school students' situational engagement and achievement in science," *International Journal of Phytoremediation*, vol. 21, pp. 139–162, 1 2002, ISSN: 15497879. DOI: 10.1207/S1532697 7EA0802_04.

[93] R. R. Bryan, S. M. Glynn, and J. M. Kittleson, "Motivation, achievement, and advanced placement intent of high school students learning science," en, *Science Education*, vol. 95, no. 6, pp. 1049–1065, 2011, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/sce.20462, ISSN: 1098-237X. DOI: 10.1002/sce.20462. [Online]. Available: https://onlinelibrary.wile y.com/doi/abs/10.1002/sce.20462 (visited on 06/15/2021).

[94] R. W. Lent, S. D. Brown, and K. C. Larkin, "Comparison of three theoretically derived variables in predicting career and academic behavior: Self-efficacy, interest congruence, and consequence thinking," Tech. Rep., 1987, pp. 293–298.

[95] D. A. Luzzo, P. Hasper, K. A. Albert, M. A. Bibby, and E. A. Martinelli, "Effects of self-efficacy-enhancing interventions on the math/science self-efficacy and career interests, goals, and actions of career undecided college students," Tech. Rep., 1999, pp. 233–243.

[96] R. M. Marra, K. A. Rodgers, D. Shen, and B. Bogue, "Women Engineering Students and Self-Efficacy: A Multi-Year, Multi-Institution Study of Women Engineering Student Self-Efficacy," en, *Journal of Engineering Education*, vol. 98, no. 1, pp. 27–38, 2009, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/j.2168-9830.2009.tb01003.x, ISSN: 2168-9830. DOI: 10.1002/j.2168-9830.2009.tb01003.x. [Online]. Available: https://onlinelibrar y.wiley.com/doi/abs/10.1002/j.2168-9830.2009.tb01003.x (visited on 01/12/2022).

[97] N. Choi, "Self-efficacy and self-concept as predictors of college students' academic performance," *Psychology in the Schools*, vol. 42, pp. 197–205, 2 Feb. 2005, ISSN: 00333085. DOI: 10.1002/pits.20048.

[98] H. L. Rotberg, F. D. Brown, and W. B. Ware, "Career self-efficacy expectations and perceived range of career options in community college students," *Journal of Counseling Psychology*, vol. 34, 2 1987.

[99] M. A. Hutchison, D. K. Follman, M. Sumpter, and G. M. Bodner, *Factors influencing the self-efficacy beliefs of first-year engineering students*, 2006. DOI: 10.1002/j.2168-9830.20 06.tb00876.x.

[100] E. L. Usher, "The roots and fruits of self-efficacy in diverse academic contexts," in *Moti-*

vation and Emotion in Learning and Teaching across Educational Contexts: Theoretical and Methodological Perspectives and Empirical Insights. Taylor and Francis, Nov. 2023, pp. 69–84, ISBN: 9781003303473. DOI: 10.4324/9781003303473-6.

[101] J. A. Chen and E. L. Usher, "Profiles of the sources of science self-efficacy," en, *Learning and Individual Differences*, vol. 24, pp. 11–21, Apr. 2013, ISSN: 10416080. DOI: 10.1016/j.lindif.2012.11.002. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1041608012001707 (visited on 06/11/2021).

[102] Y. Li, K. M. Whitcomb, and C. Singh, "How learning environment predicts male and female students' physics motivational beliefs in introductory physics courses," American Association of Physics Teachers (AAPT), Sep. 2020, pp. 284–290. DOI: 10.1119/perc.2020.pr.li.

[103] 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?" *Physical Review Physics Education Research*, vol. 14, 2 Dec. 2018, ISSN: 24699896. DOI: 10.1103/PhysRevPhysEducRes.14.020123.

[104] G. M. Quan and A. Elby, "Connecting self-efficacy and views about the nature of science in undergraduate research experiences," *Physical Review Physics Education Research*, vol. 12, 2 2016, ISSN: 24699896. DOI: 10.1103/PhysRevPhysEducRes.12.020140.

[105] J. M. Bailey, D. Lombardi, J. R. Cordova, and G. M. Sinatra, "Meeting students halfway: Increasing self-efficacy and promoting knowledge change in astronomy," *Physical Review Physics Education Research*, vol. 13, 2 Dec. 2017, ISSN: 24699896. DOI: 10.1103/PhysRevPhysEducRes.13.020140.

[106] E. G. Creamer, "Strategies for mixing prior to analysis," in *An Introduction to Fully Integrated Mixed Methods Research*. SAGE Publications, Inc., Jan. 2020, pp. 80–98. DOI: 10.4135/9781071802823.n8.

[107] T. C. Guetterman, M. D. Fetters, and J. W. Creswell, "Integrating quantitative and qualitative results in health science mixed methods research through joint displays," *Annals of Family Medicine*, vol. 13, pp. 554–561, 6 2015, ISSN: 15441717. DOI: 10.1370/afm.1865.

[108] J. M. Morse and L. Niehaus, *Mixed method design : principles and procedures*. Left Coast Press, 2009, p. 194, ISBN: 9781598742985.

[109] E. Hewitt, "Building bridges: The use of reflective oral diaries as a qualitative research tool," *International Journal of Research and Method in Education*, vol. 40, pp. 345–359, 4 Aug. 2017, ISSN: 17437288. DOI: 10.1080/1743727X.2015.1114601.

[110] S. K. T. Cheng and K. K. Ho, "A reflective learning taxonomy for an educational tour," *Educational Research for Policy and Practice*, vol. 11, pp. 243–260, 3 Oct. 2012, ISSN:

15731723. DOI: 10.1007/s10671-011-9124-7.

[111]   B. D. Lutz and M. C. Parettie, "Development and implementation of a reflective journaling method for qual-itative research development and implementation of a reflective journaling method for qualitative research," in *2019 American Society of Engineering Education (ASEE) Annual Conference Exposition*, 2019.

[112]   E. L. Usher, C. J. Ford, C. R. Li, and B. L. Weidner, "Sources of math and science self-efficacy in rural appalachia: A convergent mixed methods study," *Contemporary Educational Psychology*, vol. 57, pp. 32–53, Apr. 2019, ISSN: 10902384. DOI: 10.1016/j. cedpsych.2018.10.003.

[113]   A. M. White, J. T. DeCuir-Gunby, and S. Kim, "A mixed methods exploration of the relationships between the racial identity, science identity, science self-efficacy, and science achievement of african american students at hbcus," *Contemporary Educational Psychology*, vol. 57, 2019, ISSN: 10902384. DOI: 10.1016/j.cedpsych.2018.11.006.

Survey Block	ESM Questions/Items
Open-Ended Task	(Q1) What is the main thing you are doing right now? (Q2) Where are you doing this activity?
Task-Specific Self-Efficacy	(Q3) How skilled are you in the activity? (Q4) Do you feel in control of the situation? (Q5) Are you succeeding at what you are doing?
Open-Ended Course	(Q6) Is the activity you are doing for a course you are currently enrolled in? (Yes, No) (Q7) What is the name of the course? Please write the course label and number (e.g. PHY123) in answering this question.
Course Level Self-Efficacy *or* Career-Level Self-Efficacy	(Q7a) Considering the difficulty of this course, the instructor, and my skills, I think I will do well in this course. (Q8a) I'm confident I can do an excellent job on assignments and tests in this course. (Q9a) I'm certain I can master the skills being taught in this course. (Q7b) Considering the difficulty of my planned profession and my skills, I think I will do well once I begin working in the professsion. (Q8b) I'm confident I can do an excellent job on professional assignments once I am working in my chosen occupation. (Q9b) Once I am working in my chosen career, I am confident that I will master the skills that will make me successful.

Table A.1 The order of the ESM survey questions and items given to the research participants via the Expiwell application. Q1 and Q2 are open-response task questions, Q3 through Q5 are the task-specific self-efficacy items measured on a scale from 0 - 100, and Q6 and Q7 represent the open-ended course questions. Then, depending on whether they answer yes or no to if the activity is related to a course or not, the participants will either receive the 3 course-level self-efficacy items or the 3 career-level self-efficacy items.

APPENDIX B

CODEBOOK FOR JOURNAL PROMPTS - CONTEXTS CODES: ELICITING THE THEORETICAL FOUNDATION OF THE DOMAIN-SPECIFIC NATURE OF SELF-EFFICACY

This appendix section presents the Contexts section of the codebook; thus, only shows the Contexts codes. To support readability and clarity, the codebook is structured as a list rather than a table.

Within each of these two main sections, the individuals codes are listed alongside the following elements:

- A brief description providing a concise summary of the code.

- An expansive description detailing the boundaries of the code, including when and how it should be applied.

- Examples of prompts that are coded with this code (i.e., Example - Does Count).

- Examples of prompts that are not coded with this code (i.e., Example - Does Not Count), including borderline or easily confusable cases.

In the examples of what counts in coding a prompt or sub-prompt with a code, some words or phrases will be bolded. This is to indicate to the audience that while I shared the entire prompt, there may have been words or phrases that signaled me to code this prompt in a certain way. Additionally, When applicable, the expansive description includes explicit inclusion and exclusion criteria to guide coders. These criteria references examples that are color-coded to visually connect them with the relevant examples in the Example - Does Count and Example - Does Not Count.

- Purple text is used to link inclusion criteria to examples of what does *count* for this code.

- Violet text is used to link exclusion criteria to examples of what does *not count* for this code.

This format is intended to promote transparency in the coding process and facilitate replicability or adaption for future research.

B.1   Contexts Codes

B.1.1   Code: `Task`

Brief Description: A task is any activity/assignment AND/OR the act of doing a particular activity/assignment. Within this description of task, we include the act of attending of academic spaces. This code is utilized when a specific task or multiple specific tasks are focused on within the daily journal prompt.

Expansive Description: A task is any activity/assignment AND/OR the act of doing a particular activity/assignment. Examples of daily journal prompts that were tagged as TASK included, but not limited to, the following:

Inclusion Criteria:

- A singular or specific assignment (i.e., homework assignment - Example (1), in-class assignment - Example (2), course assignment, professional assignment, etc.)

- An exam, midterm, quiz, or assessment

- A project

- Researching or looking into a specific topic (i.e., materials online, REU programs, etc.)

- Reading a textbook

- Memorizing information

- Getting a document verifying accommodations on campus (VISA)

- Setting up appointments

- Learning about a topic or how to do something (i.e., office, calling financial aid, completing the FASFA, etc.)

- Receiving or findings resources

- Building an academic or degree plan or 4-year plan

- Performing fieldwork

Within the description of task, we include the act of attending academic spaces. Academic spaces are spaces, where students may interact with other actors to pursue academic business. These kind of tasks look like, but are not limited to, the following:

- being in/attending a class

- attending/going to an office hour/help room

- attending/going to a specific resource center on campus

- being in/attending/going to a particular research lab

- being in/attending an ULA training

- attending student organization led events - Example (3)

Exclusion Criteria: Daily journal prompts were not tagged as TASK if they did not provide a specific task. Example (1) shows a daily journal prompt that falls under this criterion.

The daily journal prompt may involve more than one task but has to be focusing on the tasks at hand. If the daily journal prompt asks about tasks or activities in a plural way, then the daily journal prompt does not receive a TASK code. To distinguish between these two sentences, I am providing two examples. The first example is to demonstrate how we have asked about multiple tasks within the daily journal prompt while focusing on specific tasks as shown in Example (4) found in the corresponding list under "Example - Does Count". The second example is to demonstrate a daily journal prompt that focuses on many tasks without focusing on a specific task as shown in Example (2) found in the corresponding list, "Example - Does Not Count."

Example - Does Count:

- Example (1): On Friday, we noticed that you spent most of the day **doing [computational modeling course] homework**. We noticed an upward trend in how skilled and successful you felt when you were working on that homework.

249

– Can you take us back to Friday and discuss what was happening while you were **working on your [computaitonal modeling course] homework**?

– How confident were you in **doing this [computational modeling course] homework**?

– How did **that homework** impact your confidence in your ability to succeed in your [computational modeling course] course?

- Example (2): Yesterday you told us about how you feel about your ability in [computational modeling course] in-class assignments. On the app today you noted that you were **working on one of those assignments**.

    – Take me back to that moment. What were you working on? Who [were] you working with? How did you feel about your abilities?

- Example (3): Today you indicated you **attended the SPS coffee break [student organization led event]** again. Can you tell us a bit about how that meeting impacted your sense of confidence for your future career plans?

- Example (4): Questions that probe your confidence in your abilities could vary quite a bit depending on what happens in the next year or what happens several years from now. Think about **an assignment that you've had within the last week. And a "professional assignment"** that you can imagine having in your future career.

    – **Compare those two assignments** for us. What do they look like? What are you doing? What skills do you need?

Example - Does Not Count:

- Example (1): Look back at the feelings you listed in the second prompt. How do those feelings impact your sense of confidence in performing tasks in your classes or in the classes?

- Example (2): Yesterday most of your day seemed to be focused on [modern physics]. Tell me a little more about how you are feeling in that class. When you think about the various

kinds of activities that are part of the class (for example attending lecture, doing homework, preparing for exams) are there some that you feel more or less capable while you're doing them?

– Was there a moment in that last two days when you felt particularly confident in your abilities in [modern physics]? What were you doing?

B.1.2 Code: `Course`

Brief Description: A daily journal prompt that focuses on a particular course or identifies the particular course or courses. The daily journal prompt can pinpoint a particular course using language as broad as "particular" or as specific as having a course label within the daily journal prompt. Ideally, the participant is being asked to discuss the particular course or each course provided in such a way that a distinction between which course the participant is focusing on is clear.

Expansive Description: The COURSE code captures coding daily journal prompts pertaining to a particular course or identified which course(s) we were investigating.

Inclusion Criteria:

• A daily journal prompt that asks about a particular course

– We may provide a full course label (e.g., PHY 471; please read what not tagged at COURSE below) to ask about a particular course - Example (1)

– We may ask them to tell us what course they are referencing as part of the daily journal prompt

– We may ask them to consider a particular course without asking them to necessarily tell us which course they are thinking of - Example (2)

• A daily journal prompt can ask about multiple courses but we must specify the particular courses we are discussing with the daily journal prompt.

251

- The prompt could have identified two particular courses by referencing two course labels - Example (3)

- The prompt could provide a list of the courses for the participants to discuss - Example (4)

Exclusion Criteria: Daily journal prompts with the following criteria were not coded as COURSE:

- Asking about courses or classes in general, or a prompt that does not ask about a particular course - Example (1)

- If we do not provide a full course label (e.g., MTH) within the daily journal prompt, then the participant may reflect upon all courses with this label (e.g., math courses); hence this is not specific enough to be counted as COURSE.

- Daily journal prompts asking about discipline-based courses is not specific enough to count. For example, if we ask about math courses, participants may reflect upon all math courses, which is not specific enough - Example (2)

- Asking about coursework alone in a prompt would not count unless this was explicitly linked to some kind of identification of a course(s). Prompts discussing coursework by itself does not provide enough context to code as COURSE because participants may think of their overall coursework that they had completed in academia or are pursuing rather than thinking of a specific course.

Example - Does Count:

- Example (1): Yesterday most of your day seemed to be focused on **[modern physics]**. Tell me a little more about how you are feeling in that class. When you think about the various kinds of activities that are **part of the class** (for example attending lecture, doing homework,

252

preparing for exams) are there some that you feel more or less capable while you're doing them?

- – Was there a moment in the last two days when you felt particularly confidence in your abilities **in [modern physics]**? What were you doing?

- Example (2): Think of a time in the past month where you felt really confident about your performance **in a particular course** you're taking now. What about it made/makes you feel confident?

  - – Take me back to that moment; what were you feeling and experiencing?

- Example (3): Over the past 3 days, your responses in the app indicated you were working on both **Cell and Molecular Biology Course and General Chemistry**. However, you indicated you were less skilled and felt less successful at **Cell and Molecular Biology Course** than **General Chemistry**. Can you tell us a little bit more about why you feel differently in these two classes?

  - – Looking back over the past 3 days, could you give us a specific example from one of these courses when you felt extremely skilled? What were you doing? How did you feel?

  - – Looking back over the past 3 days, could you give us a specific example from one of these courses when you felt not at all skilled? What were you doing? How did you feel?

- Example (4): When you think about your network, how has interactions with those folks impacted your confidence...

  - – Passing your courses (e.g., [MTH Course - Analysis I], [PHY Course - Quantum Mechanics I], [Astronomy Course - Galaxies and Cosmology], [Sociology, Behavioral, and Educational Science Course - Government and the Individual])

  - – **Obtaining a physics/astronomy degree**

253

– Or pursuing a career?

Example - Does Not Count:

- Example (1): Look back at the feelings you listed in the second prompt. How do those feelings impact your sense of confidence in performing tasks in your classes or in the classes?

- Example (2):

    – We'd like to ask a little more about your experiences building up a peer network.

    – Yesterday you noted that you wish you had more of a peer network in your physics and math classes because you said, "I'm struggling extra hard because I am alone". Do you have an example of a course experience where you did have this peer network? What course was that? How did having a peer group support you?

B.1.3 Code: `Degree`

Brief Description: DEGREE captures coding daily journal prompts discussing degree or major, or the requirements of pursuing a degree or academic plan. As an important note then, in this work we are using the words degree and major interchangeably. We also use these words interchangeably with one another: "requirements of pursuing a degree," "academic plan," "degree plan," and "4-year plan."

Expansive Description: This code captures coding daily journal prompts discussing degree or major, or the requirements of pursuing a degree or major. As an important note then, in this work we are using the words degree and major interchangeably. We also use these words interchangeably with one another: "requirements of pursuing a degree," "academic plan," "degree plan," and "4-year plan." Examples of daily journal prompts that were tagged as DEGREE included, but not limited to, the following:

Inclusion Criteria:

- Daily journal prompts discussing the pursuit of or encouragement of academic plans OR a bachelor's degree - Example (1)

254

- Daily journal prompts discussing earning a bachelor's degree/major, changing of one's major, decisions regarding one's pursuit of their major/degree plan - Example (2)

- Daily journal prompts discussing the requirements of a degree or the major (e.g., classes required for your major, future classes required for your major, etc.) - Example (3)

- Daily journal prompts discussing one's experiences or trajectory in their academic career - Example 4

Exclusion Criteria: Daily journal prompts were not coded as DEGREE if they were unable to meet the criteria above (e.g., discussing degree or the requirements of obtaining a degree). Daily journal prompts based on topics below provide some insight around prompts that were unable to meet this criteria:

- Asking about a meeting with an academic advisor would not count unless this was explicitly linked to degree and/or the academic plan. Students may meet with academic advisors for other reasons outside of discussing their degree. For example, they may ask their academic advisor about professional development opportunities that do not impact their degree attainment. - Example (1)

- Asking about coursework alone in a prompt would not count unless this was explicitly linked to degree and/or the academic plan. Prompts discussing coursework by itself does not provide enough context to code as DEGREE because students are able to take courses, in which these courses may not be related to the requirements for their degree. - Example (2)

Example - Does Count:

- Example (1): When you think about your network, how has interactions with those folks impacted your confidence...

– Passing your courses (e.g., [MTH Course - Analysis I], [PHY Course - Quantum Mechanics I], [Astronomy Course - Galaxies and Cosmology], [Sociology, Behavioral, and Educational Science Course - Government and the Individual])

– **Obtaining a physics/astronomy degree**

– Or pursuing a career?

- Example (2): What do people you know (family/teachers/peers) say to you about your decision **to pursue this major plan (or change your major)**?

- Example (3): When you think about the **classes required for your major**, what feelings do you experience? Did you feel any of these feelings distinctly today? If so, when? For example: comfortable, nervous, confused, excited, at ease, anxious, sad, determined, attentive, frustrated, inspired, detached, active, alert, etc.

- Example (4): When you look back at your **academic career** (including your experience prior to MSU), is there anything you would do differently if you had the chance?

  – Tell me a story about how you think changing that past experience would have impacted your experiences today. What would be different?

Example - Does Not Count:

- Example (1): You also told us that you were planning to meet with an academic advisor soon. Has that happened? If yes, how did it go? If not, how are you feeling?

- Example (2): When you think about your peer network (or perhaps lack thereof), how do you see it impacting your coursework?

  – How (if at all) does this impact your confidence in your abilities in your coursework?

B.1.4 Code: `Career`

Brief Description: CAREER captures daily journal prompts discussing one's intended/chosen/future profession/career/occupation, or post-graduation/career plans. As an important note then, in this work we are using the words profession, career, and occupation interchangeably.

Expansive Description: This code captures daily journal prompts discussing one's intended/chosen/future profession/career/occupation, or post-graduation/career plans. As an important note then, in this work we are using the words profession, career, and occupation interchangeably. Examples of daily journal prompts that were tagged as CAREER included, but not limited to, the following:

Inclusion Criteria:

- Confidence in their future career - Example (1)

- Requirements to move onto their future profession OR pursuing their future career - Example (2)

- Daily journal prompts probing into what their future career looks like or ideas about their future career - Example (3)

- Daily journal prompts discussing their professional assignments - Example (4)

- Daily journal prompts asking about their career plans - Example (5)

Exclusion Criteria: Daily journal prompts were not tagged as CAREER if the daily journal prompts were discussing professional development opportunities but had not explicitly linked these to their intended career. Not all professional development opportunities may impact a students' career-level self-efficacy; thus, I have decided when prompts only investigate the professional development opportunity without linking to the students' intended career in some way, these do not receive a CAREER code.

- Prompts investigating Research Experiences for Undergraduates (REUs)

- Prompts investigating research positions students may have or are pursuing - Example (1)

- Prompts investigating internship opportunities students may have or are pursuing - Example (2)

- Prompts investigating jobs students may have while in academia

Example - Does Count:

- Example (1): How do your experiences in the high energy physics lab or as a ULA affect your **confidence in pursuing your career**?

- Example (2): When you think about the **requirements to move on to your future profession**? What comes to mind? What does that look like for you?

- Example (3): We noticed in the app today you were looking at REU programs. Could you tell us a little more about how that experience went? Who were you working with? How did it go? How did it make you feel?

  - How did that experience impact your ideas **about your future career**?

- Example (4): Questions that probe your confidence in your abilities could vary quite a bit depending on what happens in the next year or what happens several years from now. Think about an assignment that you've had within the last week. And a **"professional assignment" that you can imagine having in your future career.**

  - Compare those two assignments for us. What do they look like? What are you doing? What skills do you need?

- Example (5): Today you indicated you attended the SPS coffee break again. Can you tell us a bit about how that meeting impacted your sense of confidence **for your future career plans**?

Example - Does Not Count:

258

- Example (1): When you think about your classmates, how prepared do you feel, relative to them, to get a research position, internship opportunity, or job in the upcoming year?

APPENDIX C

CODEBOOK FOR JOURNAL PROMPTS - EXPERIENCES CODES: ELICITING THE THEORETICAL FOUNDATION OF THE SOURCE NATURE OF SELF-EFFICACY

This appendix section presents the Experiences section of the codebook. The structure of this section of the codebook is similar to that of the Contexts section of the codebook previously presented, but only shows the Experiences codes. To support readability and clarity, the codebook is structured as a list rather than a table.

Within each of these two main sections, the individuals codes are listed alongside the following elements:

- A brief description providing a concise summary of the code.

- An expansive description detailing the boundaries of the code, including when and how it should be applied.

- Examples of prompts that are coded with this code (i.e., Example - Does Count).

- Examples of prompts that are not coded with this code (i.e., Example - Does Not Count), including borderline or easily confusable cases.

In the examples of what counts in coding a prompt or sub-prompt with a code, some words or phrases will be bolded. This is to indicate to the audience that while I shared the entire prompt, there may have been words or phrases that signaled me to code this prompt in a certain way. Additionally, When applicable, the expansive description includes explicit inclusion and exclusion criteria to guide coders. These criteria references examples that are color-coded to visually connect them with the relevant examples in the Example - Does Count and Example - Does Not Count.

- Purple text is used to link inclusion criteria to examples of what does *count* for this code.

- Violet text is used to link exclusion criteria to examples of what does *not count* for this code.

This format is intended to promote transparency in the coding process and facilitate replicability or adaption for future research.

C.1   `Experiences Codes`

C.1.1   Code: `Evaluation of Confidence, Skills, and Success`

Brief Description: Prompts written to elicit students' reflections on their perceived confidence, perceived skills, and perceived success. It focuses on self-assessments of abilities, capabilities, and achievement, often reflecting on past experiences and moments that shape these perceptions over time.

Expansive Description: This code is used to capture student reflections on their own confidence, skills, or perceived success in relation to a task, situation, or broader context. Prompts categorized under this code typically ask students to self-assess or evaluate their abilities, confidence levels, or sense of achievement. In some cases, students are asked to identify and rank the experiences or factors influencing their confidence or capability towards a specific task or goal. The focus is on the student's personal evaluation of their competence, confidence, and readiness, often reflecting on moments or experiences that shaped these perceptions over time. However, this code does not involve the direct connection between external or internal factors and a specific context, which is the focus of the Influences on Context-Specific Confidence/Success code. The "Evaluation of Confidence, Skills, and Success" code emphasizes general self-assessment, without explicitly linking confidence to particular influences within a context (i.e., how their homework grades influence their confidence within the course).

Inclusion Criteria:

- Evaluate what makes the participants feel confident OR how confident they feel - Example (1)

- Evaluate how skilled or successful they felt towards a specific task, course, degree, intended career, transfer OR information they use to determine their skill or success - Example (2)

- Evaluate their abilities - Example (2) and Example (3)

- Evaluate their capabilities - Example (4)

- Evaluate your competence - Example (3) and Example (5)

Exclusion Criteria:

- DOES NOT INCLUDE statements that ask the participants how a particular factors influence their self-efficacy – that belongs in the category "Influences Impacting Confidence or Success towards Domain(s)". The Evaluation code category is meant to capture DJPs that allow the participant to identify or describe what factors may be influencing their self-efficacy or allows the DJP is open enough that they choose that are relevant for providing an evaluation of their confidence in their skills, abilities, success, or purely how they know they improved in their skills, abilities, or success, whereas in the category "Influences impacting confidence or success towards domain(s)" is the researcher choosing a factor that we have posited influences self-efficacy. - Example (1)

- DOES NOT INCLUDE daily journal prompts are asking for an evaluation of an experience more broadly. In this example, there is no ask for the participant to share how this experience impacted their perception of their confidence in their abilities or a sense of their success or skill in their abilities. - Example (2)

- DOES NOT INCLUDE daily journal prompts asking for how they did something or how they made a decision. In this example, there is not ask for the participant to share how their perception of their confidence in their abilities impacted their decision or something along those lines in which they would need to share how the student is weighting, evaluating or interpreting their abilities, confidence levels, skills or success. - Example (3)

Example - Does Count:

- Example (1): If you were to **rank conversations with folks, who has the ability to impact your sense of confidence more?**

  – Family

- Faculty

- TAs/LAs

- Peers in your major courses

- Peers in your non-major courses

- Roommate

- Tell us others we have not listed above

• Example (2):

- **How skilled** did you feel taking your exam?

- **How successful** did you feel taking the exam?

- Take me back to that moment. What were you working on? Who [were] you working with? **How did you feel about your abilities?**

• Example (3): You've given us two examples (PHY471 and CMSE202) of when your own performance in comparison to peers matters **when you judge your own competence.** But yesterday you told us that **evaluate your abilities** in classes based on the grades you receive. **When you think about these two sources of information - which do you think you rely on most?**

- Does the course matter?

- Does the activity matter?

• Example (4): Yesterday most of your day seemed to be focused on [modern physics]. Tell me a little more about how you are feeling in that class. When you think about the various kinds of activities that are part of the class (for example attending lecture, doing homework, preparing for exams) are **there some that you feel more or less capable** while you're doing them?

- **Example (5):** Can you imagine a moment where your answer to the previous question would change? For example, if you usually use grades to **evaluate your confidence** can you imagine a moment where you would switch to other information? If yes, what would that imagined moment look like?

Example - Does Not Count:

- **Example (1):** Today you indicated you attended the SPS coffee break again. Can you tell us a bit about how that meeting impacted your sense of confidence for your future career plans?

- **Example (2):** When you look back at your academic career (including your experience prior to MSU), is there anything you would do differently if you had the chance?

    - Tell me a story about how you think changing that past experience would have impacted your experiences today. What would be different?

- **Example (3):** Tell us more about your decision to switch from Advanced Math to Data Science. How did you eventually make that decision?

C.1.2 Code: `Influences on Context-Specific Confidence/Success`

Brief Description: Prompts built to cue participants to discuss how particular factors (internal or external) may influence the participant's confidence or success within a particular context.

Expansive Description: Identifies references to how various influences (denoted as "X") affect a student's confidence, abilities, or success within a specific context. "X" may include internal factors (e.g., emotions, feelings, physiological states like stress) or external factors (e.g., resources, social support, training, feedback). The relationship between "X" and the context must be explicitly established in the prompt. This code is distinct from the "Evaluation of Confidence, Skill, and Success" code, which focuses on students' general assessment of their confidence or feelings towards a context or in other words for Evaluation of Confidence, Skills, and Success, there would be NO SPECIFIC "X" (i.e., group work, conversation with mom, etc.) identified.

Inclusion Criteria: Examples 1-4

264

- Mentions of specific influences (external or internal) that are linked to a students' confidence or ability or success within a particular context.

- Descriptions of how internal or external factors shape confidence or success in a given context.

- Statements where the influence is directly tied to the student's ability to succeed in or feel confident about a particular situation, task, or environment.

Exclusion Criteria: Examples 1-3

- Mentions of confidence or success that are general or abstract without linking them to a specific influence **and** specific context

- Descriptions that focus purely on self-assessments of confidence or feelings of success without identifying a specific influencing factor that is directly affecting their self-assessment within a given context.

- Evaluations that do NOT explore the relationship between an external or internal factor and a context, such as general statements about confidence without explaining the impact of specific influences.

- Mentions of experiences that are not tied to a specific influence or context related to the student's confidence, abilities, success, or skill. For example, financial troubles or future career ideas, unless these are explicitly connected to a student's confidence or success in a given context.

Example - Does Count:

- Example (1): When you think about your network, how has interactions with those folks impacted your confidence in...

- Passing your courses (e.g., [MTH Course - Analysis I], [PHY Course - Quantum Mechanics I], [Astronomy Course - Galaxies and Cosmology], [Sociology, Behavioral, and Educational Sciences Course- Government and the Individual])

- Obtaining a physics/astronomy degree

- Or pursuing a career?

- Example (2): How did this group work session with your friend on homework affect your confidence in your abilities in [PHY Course - Quantum Mechanics I]?

- Example (3): Today you indicated you attended the SPS coffee break again. Can you tell us a bit about how that meeting impacted your sense of confidence for your future career plans?

- Example (4): Can you tell us a little bit more about the lack of college knowledge? What do you think you are lacking versus your peers? How [does] this impact your success to earn your bachelor's degree?

Example - Does Not Count:

- Example (1): What troubles have you experienced financially? How (if at all) does this relate to being a first generation college student?

- Example (2): How did that experience impact your ideas about your future career?

- Example (3): Yesterday most of your day seemed to be focused on [modern physics]. Tell me a little more about how you are feeling in that class. When you think about the various kinds of activities that are part of the class (for example attending lecture, doing homework, preparing for exams) are there some you feel more or less capable while you're doing them?

C.1.3   Code: `Comparison to Peers`

Brief Description: Daily journal prompts queuing the participant to share comparisons between themselves and their peers.

Expansive Description: Asking the participant to share similarities and/or differences between themselves and their peers.

Inclusion Criteria:

- Asking the participant to share when they have compared themselves to their peers in terms of performance, knowledge, experiences, abilities or capabilities, etc. - Example (1)

- Using their judgment of themselves versus their peers to evaluate their abilities OR confidence - Example (2)

- Asking the participant to share what their peers think about their abilities or capabilities in comparison to one another - Example (3)

Example - Does Count:

- Example (1): Tell me a story about today where you noticed your performance in comparison to your peers.

    – What were you doing?

    – What were they doing?

    – How did you notice the differences in performance?

- Example (2): When you think about your classmates, how prepared do you feel, relative to them, to get a research position, internship opportunity, or job in the upcoming year?

- Example (3): Yesterday sounded like a rough day from your reflection. Today we'd like to ask a little more about the roles that other people play for you in making you feel more or less confident in future classes required for your major. When you think about the messages you get from other people, what do you hear from:

    – Your mom about what she hopes you'll do in college?

    – Your peers about how you're doing in relation to them?

C.1.4 Code: `Interactions with Others`

Brief Description: Daily journal prompts queuing the participant to discuss their engagement with others, in which others here can be a single person or multiple people. These people can range in their roles in the participant's life.

Expansive Description: Any daily journal prompt(s) cueing the participant to discuss who the participant interacts with for the purposes of mentorship/support, and sharing what this interaction or multiple interactions with a singular individual or multiple individuals entailed as well as possibly probing into the impacts of these interactions.

For transparency, interaction is being used here to mean communicating with or to (i.e., talking, looking, sharing or engaging in any kind of action), responding to or reacting to, and can even encompass an experience with someone or more broadly others.

For transparency, in the daily journal prompts, we may ask about a single person or multiple people, and these people can range in their roles in the participant's life meaning we may have referred to a single person or group of people as mentor(s), people that support them, family, instructor(s), teaching assistant(s) – TAs, learning assistant(s) – LAs, academic advisor(s), peers, folks in their network, people. . . in your academics and career plans, etc.

Inclusion Criteria:

- Identifying folks in their lives or networks that they interact with for support - Example (1)

- Discussing about a particular interaction or multiple interactions with either an individual or a group of people - Example (2)

- Sharing about interactions with others OR how these interactions. . .

  - Were encouraging or discouraging to the participant - Example (3)

  - Impacted their confidence in a singular or multiple contexts

  - Influence whether the participant thinks they are capable or have the ability to do something

– Meetings with others

Exclusion Criteria: DOES NOT INCLUDE daily journal prompts that do not include discussion around interactions. Example (1) demonstrates a prompt that does not receive the Interaction with Others code because we are merely asking who the participant spoke with, in which the participant could solely provide a name or role (i.e., a teaching assistant) without discussing the interaction with whom the person spoke.

Example - Does Count:

- Example (1): When you started at MSU **did you know anyone or have any points of contact you could go to for mentorship or support?**

  – Has that changed? If yes, **how did you find your mentors/support nextwork?**

- Example (2): Yesterday you also **mentioned Dr. McPadden is an informal mentor for you. Could you tell me a story about a moment where you felt that support.** What were you talking about? How did it go? How did you feel?

- Example (3): **When you think about your network, how has interactions with those folks** impacted your confidence...

  – Passing your courses (e.g., [MTH Course - Analysis I], [PHY Course - Quantum Mechanics I], [Astronomy Course - Galaxies and Cosmology], [Sociology, Behavioral, and Educational Science Course - Government and the Individual])

  – Obtaining a physics/astronomy degree

  – Or pursuing a career?

Example - Does Not Count:

- Example (1): Yesterday, in your reflection, you told us you went to help room. Who did you talk to in help room?

269

C.1.5 Code: `Details and/or Feelings Regarding an Event`

Brief Description: Daily journal prompts built to have a participant reflect over their lived experiences. These prompts are built as an invitation to the participants to share more details and feelings about a moment, event, activity/activities, opportunity or experience the person had.

Expansive Description:

Inclusion Criteria:

- Descriptions of what participants are feeling or experiencing during a moment - Example (1)

- Descriptions of what is happening during a moment - Example (2)

- Descriptions of how something went - "how did it go" - Example (3)

- Descriptions of how that moment came to be (e.g., questions that prompt the participant to discuss why they were participating in a moment like "who initiated it?" or "why did you meet with this group" etc.) - Example (4)

- Descriptions of your role within a moment or how the participant participated or showed up within a moment

- Descriptions of moments, opportunities, events, or experiences - Example (5). We ask these kinds of questions to understand more about a person's lived experiences

- Descriptions of a course, session, assignment(s) etc. - Example (6). We ask these kinds of questions because us, as researchers, are sometimes just trying to gain an understanding of what something even is or looks like.

- Descriptions of one thing they're proud of, frustrated by or a question they are still working through.

- Understanding what they mean or think of when they use particular language - Example (7). We may also be just trying to understand what they mean by a statement.

270

Example - Does Count:

- Example (1): Think of a time in the past month where you noticed yourself feeling like a first-generation college student. Take me back to that moment; **what were you feeling and experiencing?**

- Example (2): **Tell me a story about a moment** when an interaction with your network impacted your confidence in one of those areas. **Who were you talking with? What did they say? How did you feel?**

- Example (3): Yesterday you told us you were working on [Quantum Physics I] homework with a friend.

  – **How did the group work session on the homework go?**

- Example (4): We're also curious to hear about how this group session on Monday for [Quantum Physics I] came to happen. **Who initiated it? How did you decided to participate?**

- Example (5): Take me back to a moment today (if there wasn't one from today, was there one from last week - please let us know which you are talking about!) where you saw an opportunity to develop a skill, but that the course didn't necessarily support that development. **What was the opportunity you saw? How might that have looked different?**

- Example (6): We've noticed a few times that you reference working on MTH293. We don't see that course in the MSU course catalog. So that we're keeping ourselves straight **could you tell us more about what that course is, and who you are taking it with?**

- Example (7): **What are you thinking of when you say, "more confident of all aspects of school"?**

APPENDIX D

CONTEXT CODES - ALIGNMENT WITH THE DOMAIN-SPECIFIC NATURE OF SELF-EFFICACY

In this appendix, for each Context code, I'll discuss how it aligns with the domain-specific nature of self-efficacy.

Self-efficacy is not a general belief, but rather a domain-specific judgment of capability (Bandura, 1997). These codes reflect the situational and environmental contexts in which self-efficacy was enacted, constrained, or enhanced. Here, enacted means a person drew on or used their sense of self-efficacy, constrained means something in the context threatened their sense of capability, and enhanced means something in the context boosted their self-efficacy. This acknowledges that in some cases, the context supported the participants in feeling more capable; in others, it made them doubt their abilities or prevented them from acting on their confidence. These codes underscore the domain-specific nature of self-efficacy, illustrating how participants' confidence varied in response to the unique demands, expectations, and goals associated with the contextualizing the experience (e.g., task, course, degree, and career).

The four context codes - task, course, degree, and career - were chosen to reflect key domains in which participants formed and evaluated their self-efficacy beliefs. They represent increasing levels of abstraction and temporal distance, from immediate academic actions to future professional aspirations.

Code: Task. This code focuses on the specific activities, assignments, or actions that participants engage with within the academic setting. It includes tangible academic tasks, such as homework assignments, projects, exams, and presentations, as well as less formal academic actions like attending office hours, group study sessions, or reaching out for help.

For example, the following prompt from Fall 2021 is an example of a prompt that would receive the Task code.

Yesterday, in the app, you told us you were working on [modern physics] homework, and today, in the app, you told us you were working on [modern physics] homework.

272

We noticed that you indicated you were feeling less successful on your [modern physics] homework. Can you tell us what caused this shift?

This prompt directly addresses a specific academic task — homework in modern physics — and links it to a fluctuation in the participant's SE. While the prompt does not explicitly ask about confidence, it encourages the participant to reflect on the reasons behind the change in their perception of success, thus revealing how their confidence might fluctuate based on their engagement with a task.

In this way, the Task code serves as a starting point for exploring task-specific SE. By capturing moments tied to particular academic tasks, it sets the stage for investigating how students' confidence may shift over time based on their experiences with those tasks. However, the Task code does not directly code for task-specific SE because it does not explicitly require SE-related language as shown in the example prompt. This distinction enables a more nuanced exploration of task-related experiences and provides flexibility for future analysis, where the relationship between tasks and task-specific self-efficacy can be explored in more depth. By separating the task description from direct assessments of SE, this approach opens up possibilities for future research that can more directly examine the interaction between tasks and SE.

Code: Degree. The Degree code captures daily journal prompts exploring degree or major experiences, or the requirement or pursuit of a degree or academic plan.

For example, the following prompt from Fall 2021 is an example of a prompt that would receive the Task code.

When you think about your classmates, how prepared do you feel, relative to them to successfully earn your bachelor's degree?

This prompt is a clear example of a prompt that would receive the Degree code. However, it's important to note that while the prompt refers to the degree pursuit, it does not explicitly reference self-efficacy. The Degree code was designed to reflect degree-level self-efficacy by focusing on the pursuit of a degree, the requirements involved, and students' experiences related to their academic

273

plan. But because the prompt doesn't use explicit language about self-efficacy, it falls under the Degree code rather than being classified as degree-level self-efficacy.

Like the Task and Course codes, the Degree code was created to reflect degree-level self-efficacy by capturing moments where students engage with the concept of their degree. Degree-level self-efficacy is defined as one's confidence in their ability to complete the requirements of their planned degree program (CITE). However, the Degree code doesn't require prompts to include direct language about self-efficacy. It simply captures moments where the student reflects on their degree pursuit or related experiences, without necessarily focusing on their confidence in completing it.

Code: Career. The Career code captures the daily journal prompts discussing one's intended/chosen/future profession/career/occupation, or post-graduate/career plans.

> *When you think about the people (e.g. faculty, peers, mentors, family) who encourage*
> *you in your academics and career plans, who are they and what are the things they say*
> *to you that encourage you?*

This prompt is an example of one that would be coded with the Career code due to the reference to "...in your academics and career plans...". Although this prompt is exploring social persuasion in relation to the student's pursuit of their intended career, it does not explicitly use self-efficacy (SE) language.

The Career code was designed to be reflective of career-level self-efficacy. Career-level self-efficacy is defined as one's confidence in their ability to meet the educational requirements and perform the job duties of their planned career (CITE). As such, the Career code focuses on the educational requirements and job duties associated with students' intended careers or professions, but it does not require the inclusion of explicit self-efficacy language. Therefore, the Career code aligns with the concept of career-level self-efficacy by addressing students' reflections on their future career plans and aspirations.

APPENDIX E

EXPERIENCES CODES - ALIGNMENT WITH THE SOURCE NATURE OF SELF-EFFICACY

In this appendix, for each Experiences code, I'll discuss how it aligns with the domain-specific nature of self-efficacy.

These codes reflect the types of experiences that serve as inputs into beliefs, in line with Bandura's (1997) four primary sources: mastery experiences, vicarious experiences, verbal persuasion, and physiological/affective states. Each code is analyzed through this theoretical lens to illustrate how participants' experiences contributed to the development, reinforcement, or threat to their efficacy beliefs.

Code: Evaluation of Confidence/Success. The Evaluation of Confidence/Success code highlights daily journal prompts designed to broadly elicit reflections from students about how they interpret or assess their self-efficacy. This can include prompts that are broadly asking them to interpret how confident they feel about about a task or ranking their confidence in relation to different domains.

For example, the following prompt is an example of a prompt that receives an Evaluation of Confidence/Success code.

Yesterday most of your day seemed to be focused on [modern physics course]. Tell me a little more about how you are feeling in that class. When you think about the various kinds of activities that are part of the class (for example attending lecture, doing homework, preparing for exams) are there some that you feel more or less capable while you're doing them?

This prompt invites students to reflect on their general sense of confidence in various course-related activities. It doesn't require students to rank these activities or explicitly evaluate their confidence in each one, but instead prompts them to explore which tasks they feel more or less capable of performing. The open-ended nature of the prompt allows students to articular their perceptions of competence and confidence across different tasks, providing insight into their SE

without constraining them to predefined categories. This makes it a valuable tool for capturing students' internal evaluations of their abilities in relation to the course as a whole, even if those evaluations are explicitly ranked or quantified.

Another aspect of the Evaluation of Confidence/Success code is asking students to consider what experiences they are using to evaluate their SE within a given context or toward a certain task. Again, we intentionally write these questions in an open way so as to allow students to tell us which sources of SE they are using and prioritizing in evaluating their SE. This is different then us giving them specific sources and asking them directly how that source impacts their SE, as this is more guided and may lead them to tell us how this experience impacts their SE as well as prioritizes this experience in a way that may not be authentic to the student.

For example, consider this prompt from the Fall 2021 iteration.

Yesterday it stood to us that your confidence in your abilities in UGS200H was notably higher. Could you tell us what was happening yesterday? What were you doing? Who were you working with? What impacted your sense of confidence?

The last question—"What impacted your sense of confidence?"—was what led me to code this prompt as Evaluation of Confidence/Success. This question invites students to reflect on the factors that influenced their confidence, providing insight into how they are evaluating their SE. There could be many factors at play that influenced their confidence in their research course, and the open-ended nature of the question gives students the freedom to identify which experiences or interactions were most salient in shaping their perceptions.

Code: Interactions with Others. The Interactions with Others code highlights the daily journal prompts written to elicit reflections from students about how their interactions with others (e.g., peers, mentors, instructors) have shaped their self-efficacy. Because this code relates to students' interactions with others, the ways in which students may perceive or interpret their interactions can be different. For example, the prompt below shows how we used a daily journal prompt to explore a specific moment:

"Today, you told us you were meeting with [research faculty advisor] for [research project]. Can you tell us a little bit more about what is [research project]? How did your meeting with [research faculty advisor] go? How are you feeling?"

- Daily Journal Prompt, Rose, Week 4, Day 3

This prompt was coded as Interactions with Others because it encourages Rose to reflect on her meeting with her research faculty advisor. The open-ended nature of the prompt was intentional: we wanted to avoid leading Rose toward discussing a specific source of self-efficacy or making her feel pressured to link the interaction directly to her self-efficacy. By sharing how the meeting went, Rose might reference social persuasion, vicarious experiences, or other ways the advisor might impact her self-efficacy—or she may not make a direct link at all. This example demonstrates how Interactions with Others is connected to the experiences of self-efficacy, without necessarily being linked to a specific source of self-efficacy. As such, the Interactions with Others code cannot be directly equated with one of Bandura's sources of self-efficacy. However, it can offer valuable information that may cue or be linked to one of these sources. In particular, this code may likely investigate social persuasion or vicarious learning experiences as a result of exploring the interaction with another individual.

Some daily journal prompts identified as Interactions with Others were written to be more closely aligned with social persuasion. An example from the Fall 2021 data is provided:

Reflecting over this past week, were there any people or experiences with people who encouraged or discouraged you in a particular course you are in right now? Who are they? What are the things they say to you that either encourage or discourage you in the course?

- Daily Journal Prompt, Rose, Week 3, Day 5

This prompt was coded as Interactions with Others because it asks participants to reflect on how their interactions with others — such as professors, peers, or mentors — have influenced their self-efficacy in a specific academic situation. The question is designed to explore how these

277

interactions may either encourage or discourage students in their belief about their ability to succeed in the course.

In this case, the prompt is strongly linked to social persuasion, one of Bandura's primary sources of self-efficacy (1997). When participants describe receiving encouragement or discouragement from others, they reflect on how these external verbal influences, whether positive or negative, affect their confidence in their academic abilities. For example, if a student shares that a professor's praise boosted their confidence, that would clearly be an example of social persuasion. Similarly, if a student mentions that a peer's criticism made them question their abilities, that also reflects social persuasion shaping their self-efficacy.

This example demonstrates how Interactions with Others can lead to experiences of social persuasion, showing how external encouragement or discouragement may enhance or hinder a student's belief in their capabilities. However, it is important to note that while this experience is tied to social persuasion, the Interactions with Others code itself is not synonymous with the source of social persuasion. Rather, it highlights the type of experience that could influence self-efficacy through social persuasion.

Code: Comparison to Peers. This code captures moments where participants reflect on how they view their abilities in relation to others, particularly their peers in similar academic or professional situations. This type of experience often involves students measuring their performances, achievements, or progress against others, which may shape their confidence in their own capabilities. A prompt given in Fall 2021 is provided below.

> *When you think about your classmates, how prepared do you feel, relative to them to successfully earn your bachelor's degree?*
>
> - *What are the specific experiences you've had or things you've heard from other people that make you answer the above question the way you did?*
>
> - Daily Journal Prompt, Charlotte, Week 4, Day 4

278

This prompt was coded as Comparison to Peers because it asks participants to reflect on their level of preparation relative to their peers to successfully earn their bachelor's degree. The first question is designed to explore how they perceive their abilities to earn their degree compared to others, while the second question asks participants to share the experiences or insights that led to their evaluation.

The first question of the prompt strongly invites vicarious learning, one of Bandura's primary sources of self-efficacy (1997). When participants compare their abilities to those of their peers and reflect on how this comparison influences their confidence, it can lead to vicarious experiences. For example, students may perceive their peers' successes as motivating, while seeing others struggle may lower their confidence.

This example demonstrates how the Comparison to Peers code often prompts students to reflect on their peers' experiences, which can influence their own self-efficacy. When participants observe others overcoming challenges, it can enhance their belief in their ability to succeed. In contrast, if they feel their peers are performing better, it can undermine their own self-efficacy by highlighting perceived shortcomings.

Although the Comparison to Peers code does not directly represent vicarious experiences, it often acts as a catalyst for them. By reflecting on comparisons, students may implicitly consider others' actions and outcomes, which ultimately shape their self-efficacy.

Notably, Comparison to Peers is strongly linked to vicarious learning, but it's important to remember that it focuses exclusively on comparisons with peers. However, comparisons can also be made with other significant figures, such as mentors or faculty, and this influence should be explored in future prompts. To gain a fuller understanding, we can design broader prompts that allow students to reflect on any comparison, peer-related or otherwise, that could impact their self-efficacy. This could help capture a more holistic view of how comparison shapes their self-efficacy beliefs.

Code: Details and/or Feelings Regarding the Event. The Details and/or Feelings Regarding the Event code captures individualized daily journal prompts designed to invite participants to

279

reflect on specific moments, experiences, or events in their daily lives—eliciting rich, emotional, or descriptive information. While these prompts do not always directly explore participants' self-efficacy, they surface the context and lived experiences that may be shaping it over time. These types of prompts are particularly valuable in mixed-methods journaling, as they often serve as entry points: either by providing necessary background before deeper exploration of self-efficacy, or by introducing new, participant-driven insights that warrant follow-up through additional questions or codes.

For example, consider this daily journal prompt from the Fall 2021 iteration:

When you think about the requirements for completing your major, what do you think has been your biggest challenge?

- *Tell me a story that illustrates this challenge. What were you doing? Who were you talking to? How did you feel?*

This prompt was coded as Details and/or Feelings Regarding an Event because it asks the participant to reflect on a significant challenge related to completing their major, and to share a story that illustrates that challenge. While the experience described could involve a mastery experience or social persuasion (both sources of self-efficacy), the prompt is intentionally broad—designed to elicit the participant's own framing and interpretation of the event. Because of this openness, it does not cleanly align with a specific source of self-efficacy.

At the same time, the prompt clearly engages the Degree context—aligning with the domain-specific nature of self-efficacy—but without explicitly probing any particular source, such as mastery experiences or social persuasion. This is precisely why the Details and/or Feelings code exists: to capture foundational narratives and emotional reflections that may inform a student's self-efficacy, even when the prompt is not structured around a specific source. In this way, it fills an important gap in the Experiences part of the codebook.

While the prompt analyzed earlier focuses on the emotional aspects of an experience ("How did you feel?"), it is worth noting that the Details and/or Feelings Regarding an Event code also

captures emotional reflections that could be linked to physiological states. However, it is important to clarify that we cannot definitively claim that these emotional reflections are directly tied to physiological states such as heart rate or muscle tension. Emotions and physiological responses are complex, and while emotional experiences may sometimes reflect underlying physiological states, this relationship is not always straightforward. As such, this code allows for the capture of emotional and narrative reflections without explicitly asserting a connection to physiological states, leaving open the possibility that they could overlap or be distinct.

APPENDIX F

CODING PROTOCOL: UNIT OF ANALYSIS RULES

This appendix outlines the coding protocol used for the codebook, which outlines the unit of analysis guidelines.

Purpose of Document: Within our data, there are many prompts and sub-prompts symbolized by bullets and sub-bullets. To code the various daily journal prompts consistently, we developed a set of rules describing the "unit of analysis" – decisions around coding the prompts and sub-prompts depending on the portion of the content that receives that specific code. To describe our coding protocol, I wrote a section titled "Unit of Analysis Rules" to detail our rules and then I wrote a section titled "Walking Through Unit of Analysis Rules" to walk people through how we apply those rules to examples.

General Steps for Coding:

1. Focus on the entire daily journal prompt at a time. An entire daily journal prompt will be broken down into main prompts versus sub-prompts.

2. Look for and highlight referential phrases. Please find the purpose, description and further examples of referential phrases in Appendix G. Once you have done this, pull up the Unit of Analysis Rules provided in the next section of this document, and move onto (a) and (b).

   a) Code the appropriate prompts and/or sub-prompts with the Context codes. – *Note: When coding, the entire prompt or sub-prompt receives the appropriate codes (i.e., do not code each sentence or group of sentences based on theme).

   b) Code the appropriate prompts and/or sub-prompts with the Experiences codes. – *Note: When coding, the entire prompt or sub-prompt receives the appropriate codes (i.e., do not each sentence or group of sentences based on theme).

Unit of Analysis Rules:

1. Code the main prompt and sub-prompt separately if there are no referential phrases. For example, the sub-prompt counts as "course" if it directly names a course (e.g., PHY471) OR if it refers back to an earlier prompt that named a course (e.g., a referential phrase, in which case here is the course) OR if it asks a student to name a particular course in their response.

2. Main prompts and sub-prompts may get coded together whenever there is a referential phrase – when there is a phrase linking one prompt to another prompt. Note that the referential phrase may be in the main prompt or in the sub-prompt or in both prompts.

    a) The referential phrase does not only link the main prompt and the sub-prompt. Follow the rules accordingly for any mixing of levels between prompts. The referential phrase may link across prompt types:

        i. Main Prompt to Main Prompt

        ii. Main Prompt to Sub-Prompt

        iii. Sub-Prompt to Main-Prompt

        iv. Sub-Prompt to Sub-Prompt

    b) When the referential phrase, appearing within either/both the main prompts and/or sub-prompts, is a phrase based on **Context Codes**, then the main prompt and sub-prompt receive individual separate, same codes for domains.

        i. In the case, where we have referential phrase(s) within the main prompt and following sub-prompts, the prompts will only get separate, SAME code for Context Codes, if the referential phrases are across both the prompt and sub-prompts.

    c) When the referential phrase, appearing within either/both the main prompts and/or sub-prompts, is a phrase based on **Experiences Codes**, then we code the next prompt together with the earlier prompt resulting in a singular code across these prompts.

CODING PROTOCOL: REFERENTIAL PHRASES

This section serves the following purposes:

- Document why we created Referential Phrases - Motivation for Referential Phrases

- Present a general overview of Referential Phrases (e.g., description of the highlighted code and how to apply the code

Modification for Referential Phrases

When coding the daily journal prompts, there were a couple of daily journal prompts that were challenging to code because there was a nuance in them that I had observed, which was that it felt inauthentic to code prompts and/or sub-prompts separately when they were written about the same topic, event or contexts. To illustrate this, I have provided an example daily journal prompt below, where we have written a daily journal prompt regarding a participant's experience of working on a physics homework with a friend one day. The entire daily journal prompt consists of a main prompt and three sub-prompts. The main prompt situates the participant by telling them that the following questions will be regarding their experience of working with a friend on this physics homework. Then in each sub-prompt we ask a different question regarding this experience. If we were thinking about Context codes, we know the course (i.e., PHY 471) and the task (i.e., working on homework), but if I only apply the code based on the words within this prompt, we dropped the course context in the first two follow-up questions presented in the first two sub-prompts. This information still applies to those two sub-prompt though. Thus, we would have missed capturing this nuance in our coding, which felt inauthentic to the daily journal prompt(s) that were written, so I wanted to create a way to account for this complexity within the daily journal prompts. To do this work, I would need a way to link prompts and/or sub-prompts together based on shared topics and/or contexts, if present within the daily journal prompt, so the "Referential Phrases" code was built to do this work.

Yesterday you told us you were working on [Quantum Mechanics I] homework with a friend.

- *How did this group work session on the homework go?*

- *How [did] you feel about the homework after that session?*

- *How did this group work session with your friend on the homework affect your confidence in your abilities in [Quantum Mechanics I]?*

General Overview of Referential Phrases

- Code Description: A referential phrase is any phrase that may appear in a prompt and/or sub-prompt that serves to link prompts and/or sub-prompts together based on the sharing of a topic or context. A link between prompts and/or sub-prompts means that the prompt or sub-prompt is referring back to an earlier prompt or sub-prompt via the shared topic or context(s).

  - This is not a traditional code like the Contexts codes or the Experiences codes. This code is meant to support a coder in recognizing when prompts and/or sub-prompts should share Context and Experiences codes. Because Referential Phrases are meant to support when prompts and/or sub-prompts should share Context and Experiences code, in MAXQDA, I recommend using the Color Coding (Highlighting) tool. This will help draw the eye to Referential Phrases that exist within a prompt.

- Unit of Analysis:

  - Referential phrases are applied to phrases within prompt or sub-prompt. Unlike the other codes that exist, the Referential Phrase code does not get applied at the prompt-level (i.e., code the entire main prompt or sub-prompt).

  - Referential phrases may be in the main prompt or in the sub-prompt or in both the main prompt and sub-prompt.