Bridging the Data Divide: An Experiential Learning Approach to Preparing Students for Artificial Intelligence-Driven Careers

Note: This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. The final article will be available, upon publication, via its DOI (TBD).

Thacker, I., Jin, Y., Koyuncu, I., Rios, A. & Niu, J. (June, 2026). Bridging the Data Divide: An Experiential Learning Approach to Preparing Students for Artificial Intelligence-Driven Careers. Paper to be presented at the American Society for Engineering Education (ASEE), Computing and Information Technology (CIT) Division. Charlotte, NC.

Abstract
We designed and evaluated a 15-week experiential learning course designed to promote data literacy, academic motivation, and interdisciplinary applications in emerging technologies. Over 15 weeks, 23 undergraduate students at a Hispanic Serving Institution completed the course, revealing significant gains in data literacy efficacy (Cohen's $d = 2.19$) and interest in pursuing professions in emerging technologies ($d = .772$) as measured by a pretest to posttest questionnaire. Qualitative evidence suggested that students identified barriers to pursuing careers in emerging technology (e.g., a competitive job market and gender and racial bias in STEM) as well as the benefits (e.g., opportunities to develop skills and potential to improve the lives of others). Findings demonstrate that experiential data analytics curriculum can enhance students' motivation, confidence, and career readiness for careers in emerging technology.

Keywords: Experiential Learning, Data-Literacy, Data Science, Artificial Intelligence Education

Introduction

Preparing the next-generation workforce to thrive in a digital and artificial intelligence (AI)-driven world is essential for both individual success and for sustaining U.S. leadership in global innovation (Acemoglu et al., 2021; Acemoglu & Autor, 2011; Acemoglu & Restrepo, 2020b; Autor et al., 2003; Bessen, 2019; Brynjolfsson & McAfee, 2014; Frank et al., 2019; Mäkelä & Stephany, 2024). Yet, despite the rising demand for data and AI expertise, many higher education institutions lack interdisciplinary, hands-on learning opportunities in physics, engineering, and computer science—particularly affecting first-generation college students, women, and students of color (Ceci et al., 2009; Cheryan et al., 2017; Lohfink & Paulsen, 2005; Martin et al., 2020; Pascarella et al., 2004; Portes, 1998). Classroom culture and experiences in fields such as Physics, Engineering, and Computer Science tend to feel more welcoming to White and Male students from higher SES (Cheryan et al., 2009; McGee & Martin, 2011), which partly explains why women and students of color have lower college persistence in these fields, and lower representation in the STEM workforce (Ceci et al., 2009; Cheryan et al., 2017; NSF, 2020; NCSES, 2023). This underscores an urgent need to attract underrepresented students to the technology workforce who are dramatically underserved in STEM-related fields.

One useful educational strategy that has potential to provide students with meaningful and engaging learning experiences, while also preparing them for the workforce is to create authentic *experiential learning* opportunities. Experiential learning happens when people have real-world interactions that transform how they view the world. By engaging students in a curriculum explicitly intended to prepare them for internships with industry partners in emerging technology, we aimed to illustrate the relevance of experiential learning content as it relates to future work settings.

To address STEM disparities and provide meaningful learning experiences for students, we designed and evaluated a 15-week experiential learning course intended to promote data literacy, academic motivation, and interdisciplinary applications in emerging technologies. The goal of this project was to address these gaps by studying how experiential learning and mentorship experiences in emerging technology fields affects underrepresented student interest and motivation in these fields. The project focuses on developing practices that empower underrepresented students and establish a practical pipeline for recruiting, training, and retaining a diverse workforce in emerging technology. We collaborated with local industry partners to provide impactful hands-on experiential learning opportunities that enable participants to explore three fields related to emerging technology: digital twins, smart construction, and retail innovation.

In what follows, we summarize the educational theories of experiential learning, expectancy value, and data literacy that inform our curricular design, provide an overview of content sequenced in our 15-week course, describe the pre-test-posttest research design for assessing its efficacy in promoting key motivational outcomes, our findings, and conclude with reflection on the lessons learned from taking on this approach.

Theoretical Framework

The aim of this project is to promote students' interest and career readiness in fields of emerging technologies through transformative experiences. To frame our evidence-based approaches to curriculum design, we draw from Experiential Learning Theory (Kolb, 2014; Pugh, 2011) and academic motivation.

Experiential Learning Theory

The role of experience in education has a long tradition of scholarship (Dewey, 1938; Lewin, 1951; Piaget, 1971). While there is no consensus on a definition, experiential learning can be defined as a "process whereby knowledge is created through the *transformation of experience*. Knowledge results from the combination of grasping and transforming experience." (Kolb, 1984, p. 41). Or in other words, experiential learning happens when people have real-world interactions that transform how they view the world. Indeed, within the framework of *Experiential Learning Theory* (Kolb, 1984; 2014), learning through concrete, real-world experiences, such as internships, externships, capstone projects, and other workforce experiences, fundamentally transform how students see the world through iterative cycles of "grasping experiences"—concrete experiences and abstraction of those experiences—followed by "transforming experiences"—reflective observation and active experimentation (see also Pugh, 2011). These transformative cycles are thought to culminate in increasing levels of knowledge acquisition, specialization, and self-identification (Kolb 2014). As such, real-world experiences can be useful for students to gain practical knowledge and skills for solving real-world challenges related to emerging technologies.

Though we were unable to find empirical research on the impact of experiential learning opportunities related specifically to emerging technology, there is some research investigating the related field of data-science education. For example, evidence suggests that undergraduate engineering, computer science, and data-science curriculum that includes experiential learning opportunities with industry partners, consulting workshops, and real-world capstone projects have the potential to lead to fruitful outcomes such as improved data-analysis skills, preparation for internships, problem solving and critical thinking skills (Wu, 2022; Hassan et al., 2021) though some caution that larger real-world projects may not be successful without including structure around projects (Allen, 2021). However, this research is limited in that it makes use of pre-experimental research designs that have limited internal validity (Cambell & Stanley, 1963) and often use researcher-created measures with limited construct validity. Furthermore, this prior research does not investigate whether student *motivation* is altered through experiential learning curriculum.

Expectancy Value Theory

A wealth of research in education shows that student motivation is a key component in fueling learning and successful academic performance (e.g., Gottfried et al., 2013; Vu et al., 2022). To illustrate why motivation is important to consider when designing and testing experiential learning curriculum, we summarize theory and principles from *Expectancy Value Theory.* According to Expectancy Value Theory (EVT), students are motivated when they *expect* to be successful on a task and if they place *value* on the content that they are learning (Eccles et al., 1983; Wigfield et al., 2017). EVT posits that there are four different types of task value: *intrinsic value* is when a learner values a task because they find the activity enjoyable for its own sake, *attainment value* is perceived personal importance of a task as it relates to one's identity, *utility value* refers to perceptions that a task may be useful to a learner to achieve their present or future goals, and *cost* is the extent of time and effort that is perceived to complete a task. Interventions can be designed to promote these different forms of task value; we draw primarily from those that promote utility value.

Drawing from EVT, several researchers have created classroom interventions to promote utility value and learning by supporting students in making connections between their lives and the content they are learning (Eccles et al., 1983; Hulleman & Harackiewicz, 2021; Thacker et al., 2025). For example, students across disciplines often wonder "Why are we learning this?", and addressing this question with information about why content is useful in students' lives can promote utility value (Eccles et al., 1983; Wigfield et al., 2017). Namely, when interventions are personal, specific, and emphasize how content can be relevant to students' lives, utility value interventions can significantly increase student learning and interest (Hulleman et al., 2018) and may be particularly effective for underrepresented students in higher education (e.g., Harackiewicz et al., 2016).

Experimental evidence suggests that helping undergraduate students see how content is relevant to their lives can promote positive motivational and academic outcomes (e.g., STEM interest, utility value, grades, test scores, and passing rates), with stronger effects for marginalized groups of students in STEM (Hulleman et al., 2010; 2018; Harackiewicz et al., 2016; Rosenzweig et al., 2019; Kossovich et al., 2019; Seyranian et al., 2023). However, to date, no research has connected EVT outcomes with experiential learning opportunities, nor has it focused specifically on the acquisition of learning, motivational, and career outcomes as they relate to emerging technology careers. As such, we intend to select experiences that emphasize real-world applications, illustrate why it's relevant, and valuable in the field of emerging technology.

Data Literacy

A set of skills that are crucial for pursuing careers in emerging technology, and thus merit elaboration, is data literacy. Though there is no consensus on a definition, the term "data literacy” can be defined as the statistical competencies, methods, and techniques that facilitate decision-making (Gould, 2017). These include: *acquiring*, *managing*, *visualizing*, *interpreting*, *evaluating*, *reading*, and *using* data (Börner et al., 2019; Carlson & Johnston, 2014; Kim et al., 2023; Prado & Marzal, 2013; Ridsdale et al., 2015); dimensions that are also critical for driving insight and advancement in fields of science more broadly (Qiao et al., 2024; Thacker, 2023, 2024; Thacker et al., 2024, 2025).

Indeed, data literacy is a multifaceted competency and is widely considered an essential skill that holds significance across STEM domains (Bowler et al., 2017), and thus individuals' *perceptions* of their data literacy are also an important metric of their confidence and motivation in data-specific fields. Specifically, college students’ perceptions of their own data-specific competencies have been linked to heightened persistence and achievement (Chen, 2025). In this study, we aimed to develop activities to enhance students’ data literacy efficacy.

Current Study

Despite current research focusing on experiential learning opportunities in data-science related fields, none focus specifically on applications to promote skills in emerging technology fields, nor do they investigate the important role that expectancies, values, and the need to belong play in such pursuits. As such, we intend to address the following research questions:

- **RQ1:** How does participation in a 15-week experiential learning course enhance students’ motivation and confidence in applying data analytics to emerging technologies?
- **RQ2:** What are students’ perceptions of the benefits and challenges of pursuing data-intensive careers after completing the course?

Methods

To address these questions, we designed and evaluated a 15 week course focused on experiential learning intended to promote data literacy, academic motivation, and interdisciplinary applications in emerging technologies. The course included training in foundational content, specialized content, and an opportunity to apply for a paid internship. The *foundational content* consisted of nine modules: six technical modules on statistics, database management with SQL, and AI fundamentals using Google Colaboratory (a free, cloud-based Jupyter notebook environment for Python) and three modules on project management, oral and written presentation, and leadership to strengthen students’ soft skills (see Appendix A for an overview). Students then completed six *specialized application* modules wherein students choose one of three tracks: smart manufacturing, smart construction, or smart retail. These tracks were co-designed with industry partners and led by faculty with expertise in applying AI, digital twins, and analytics tools to solve real-world problems. This specialization phase prepared students for a final phase (not required from students or measured in this study) where students submit applications to participate in ten-week paid *internships with local industry partners*.

Participants & Procedure

We recruited undergraduate students to participate in the course in the Spring and Summer of 2025. The study was conducted at a university in the Southwestern United States, where undergraduate students were recruited though a dedicated online recruiting network via emails and flyers. Students qualified to participate if they were in good academic standing and received a $4000 tuition stipend (but no course credit) for participating upon completion of the course. Participation in the program was competitive, our budget allowed for only 15 available spaces, thus we required students to submit an application in advance to be considered.

Before and after the 15 week course (before the internship), students completed a pretest and a posttest survey. Eight students initially enrolled in the experience but opted out before finishing the course, and thus we used the remaining funds to recruit additional students to take their place.

In all, N = 23 students completed the pretest, and N = 13 completed the posttest. At pretest, participants reported their gender (57% female, 43% male), ethnicity (52% Hispanic), and race (52% White, 30% Asian, 9% American Indian/Alaska Native, 4% Black, 4% multiracial). Most were STEM majors (74%), primarily third- and fourth-year students (86%), with 13% international students. The mean age was 21 years.

Materials & Measures

Both the pretest and posttest included identical measures of data literacy efficacy (Kim et al., 2024, 31 items), interest in emerging technology (adapted from Hulleman et al., 2010), expectancy, value, cost (Kosovich et al., 2015), and two open ended items reporting their perceptions of challenges and benefits of pursuing a career in emerging technology (researcher created). For all items and materials, see Appendix B. For internal reliability metrics across all scales using Cronbach's alpha, see Table 1.

Data literacy efficacy items were adapted from Kim et al., 2024 and consisted of 31 items that capture student's confidence in their data skills. Items asked students to rate the extent to which they felt knowledgeable about specific data skills, such as "Defining what data is" on a scale from 1 (no knowledge) to 5 (expert). The measure has three subscales, one measuring students' efficacy in their *data identification skills* (e.g., "Defining what data is" ; 6 items), *data processing skills* (e.g., "Displaying and summarizing data using descriptive statistics", 12 items) and *data management and sharing skills* (e.g., "Differentiating between different data storage devices" and "Determining data ownership and use rights…", 8 items).

Measures of expectancy, value, and cost were adapted from Kossovich et al., (2015) consisting of 10 items measuring students' expectations for success (e.g., "I know I can learn the material in this class" 3 items), value (e.g., "I think this class is useful", 3 items), and cost (e.g., "I have to give up too much to do well in this class", 4 items). Participants responded on a seven point agreement scale.

Students' personal interest in emerging technology was measured using three items (adapted from Huelleman et al., 2010, 2009) and asked students to respond to statements (e.g., "I think the field of emerging technology is interesting.") on a seven point agreement scale.

Lastly, we created two open-ended questions prompting students to "...write a paragraph to explain your perceptions of the *challenges* you might encounter if you were to pursue a career in fields of emerging technology" and another asking students to describe the *benefits*.

Results

Changes in Motivation and Confidence (RQ1)

To examine change in data efficacy and motivational outcomes from pretest to posttest (RQ1), we fit a linear mixed-effects model using the "nlme" package in R (Pinheiro & Bates, 2000) because it models individual differences in change, handles missing data, and offers more statistical power than traditional ANCOVA or T-test methods (Jost & Jansen, 2022). Time (pre vs. post) was treated as a within-subjects fixed effect and participant included as a random intercept to account for individual differences in baseline levels. This approach uses full-information maximum likelihood estimation, allowing all available data to contribute to the analysis and providing unbiased estimates. The fixed effect of time tests whether mean scores significantly changed from pretest to posttest, while the random structure models correlation among repeated measures.

Preliminary Analysis

Prior to fitting our models, we conducted a series of diagnostic checks to evaluate whether the data met the assumptions of the analytic approach. Visual inspection of Q–Q plots indicated that residuals and random effects were approximately normally distributed. Plots of residuals against fitted values and by time point suggested that distributions were relatively homoscedastic. Autocorrelation functions revealed no meaningful temporal dependence in the residuals. T-tests revealed no significant differences across participant characteristics when comparing students with and without missing data (all $p > .133$), supporting the plausibility of the missing at random (MAR) assumption for the linear mixed-effects model. Overall, the data satisfied the assumptions of the mixed-effects modeling framework, and no violations were severe enough to warrant model modification. See Table 1 for descriptive data for all analytic variables.

Table 1

Descriptive Statistics

	n	mean	sd	min	median	max	skew	kurtosis	#items	α
dat.efficacy.pre	9	2.59	0.76	1.32	2.77	3.58	-0.13	-1.45	31	.98
dat.efficacy.post	13	4.01	0.50	3.19	4.00	4.90	0.19	-1.11	31	.96
data.identification.pre	9	2.85	0.66	2.00	3.00	3.83	0.08	-1.54	6	.88
data.identification.post	13	4.04	0.30	3.50	4.00	4.83	0.98	1.77	6	.64
data.processing.pre	9	3.67	1.06	1.75	4.08	5.08	-0.35	-1.29	12	.97
data.processing.post	13	5.73	0.83	4.08	5.67	6.92	-0.09	-0.94	12	.96
data.management.pre	9	2.39	0.95	1.00	2.25	4.00	0.20	-1.28	9	.98
data.management.post	13	3.90	0.61	3.00	4.00	5.00	0.23	-1.02	9	.89
pers.interest.pre	23	6.07	0.64	5.00	6.00	7.00	-0.22	-1.05	3	.51
pers.interest.post	13	6.67	0.53	5.67	7.00	7.00	-1.05	-0.71	3	.92
expectancy.pre	23	4.54	0.48	3.67	4.67	5.00	-0.30	-1.67	3	.87
expectancy.post	13	4.59	0.49	4.00	5.00	5.00	-0.35	-1.96	3	.96
value.pre	23	4.33	0.63	3.00	4.00	5.00	-0.27	-1.19	3	.94
value.post	13	4.59	0.49	4.00	5.00	5.00	-0.35	-1.96	3	.96
cost.pre	23	2.41	0.85	1.00	2.50	5.00	0.73	1.56	4	.95
cost.post	13	2.52	0.70	1.25	2.50	4.00	0.18	-0.35	4	.63

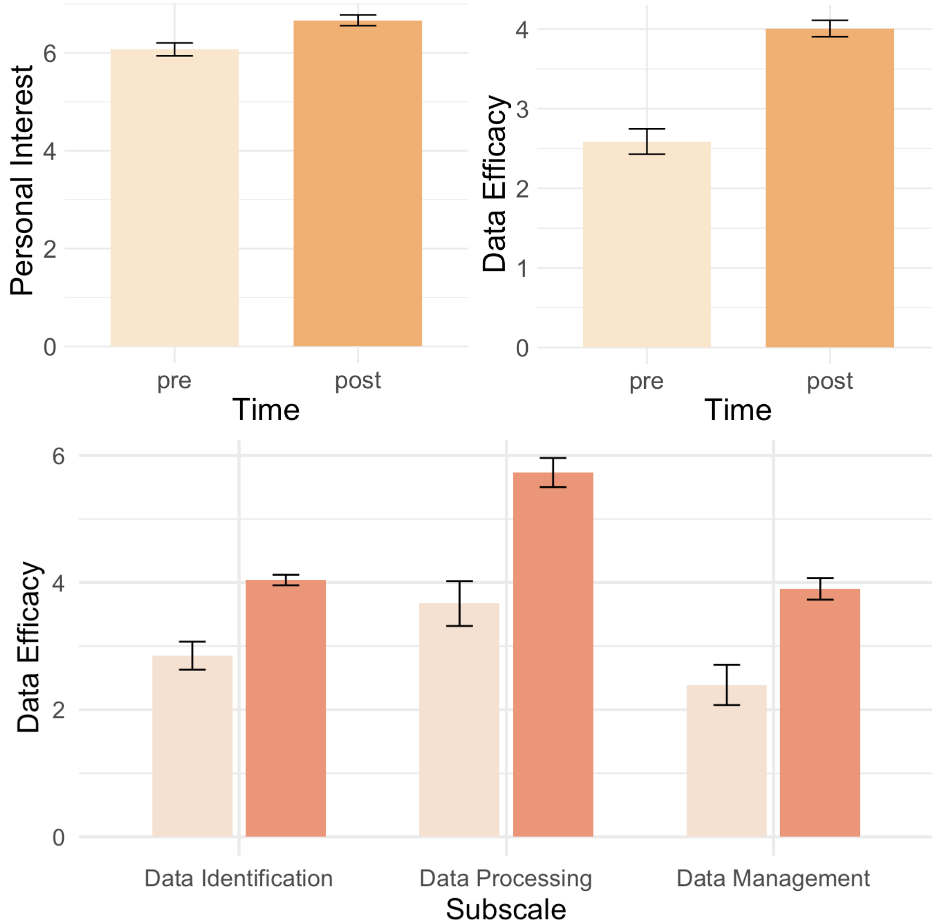
Attrition Analysis

We conducted an attrition analysis to assess whether characteristics of participants who dropped out of the study differed from those who did not. Wilcoxon signed rank tests revealed no significant mean differences across all continuous pretest variables (efficacy, data-efficacy scales, interest, expectancy, value, cost, age; all ps $\geq .268$). Pearson's Chi-Square tests revealed no significant relationships between attrition and categorical variables (gender, race, ethnicity, year of study, STEM major, English as a first language status, and international student status; all *ps* ≥ .27).

Main Analysis

We found that students' general data-efficacy dramatically improved from pretest (M = 2.5, SD = 0.76) to posttest (M = 4.01, SD = 0.50)—an increase by more than two standard deviations (d = 2.18). The fixed-effect estimate from the linear mixed effect model was b = 1.42, indicating that participants showed an average increase of 1.42 points from pretest to posttest on the five point scale. This change was statistically significant, F = 30, p = .012. The variance of the random intercept was τ^2 = .047, demonstrating meaningful individual differences in baseline efficacy. The residual variance was σ^2 = .333. Together, the model explained approximately 62.4% of the total variance in efficacy (conditional R^2), with the fixed effect of time accounting for 57.2% (marginal R^2).

Figure 1. *Mean Interest, Data Efficacy, and Data Efficacy Subscales Over Time*



Note. Intervals represent ±1 standard error above and below the mean.

We also ran parallel analyses with the three subscales of data efficacy (data identification, data processing, and data management) to investigate student growth across these specific dimensions. We found similar growth from pretest to posttest across all three subdimensions of data identification (d = 2.15, b = 1.18, F = 35.0, p = .010, τ^2 = .034, σ^2 = .192), data processing (d = 2.06, b = 2.05, F = 26.8, p = .014, τ^2 = .095, σ^2 = .774), and data management (d = 3.15, b = 1.52, F = 21.1, p = .019, τ^2 = .023, σ^2 = .566). For a visualization of unadjusted mean differences in data efficacy from pretest to posttest, including disaggregated subscales, see Figure 1.

Regarding personal interest in the field of emerging technology, we also observed significant improvement from pretest (M = 6.07, SD = .64) to posttest (M = 6.67, SD = .53), about two thirds of a SD increase (d = .77). The mixed effect analysis revealed significant gains in interest over time (b = 0.58, F = 8.36, p = .014, τ^2 = .036; σ^2 = .324) with time accounting for about 18% of the variance in personal interest ($R^2_{\text{conditional}}$ = .263, R^2_{marginal} = .181).

We also tested whether expectancy, value, or cost changed from pretest to posttest. While we did observe increased levels of expectations for success and task value and decreased levels of cost perceptions from pretest to posttest (see Table 1), these differences were not significant when tested in mixed effects models (all p > .05 in predicting).

Perceived benefits and challenges of pursuing data-intensive careers (RQ2)

To explore students' perceptions of the challenges and benefits of pursuing careers in emerging technologies, we used inductive qualitative coding processes (Chamraz; 2015). Namely, the first author openly-coded all posttest responses (N = 11) to two open-ended prompts asking students to write a paragraph to explain their perceptions of the (a) benefits and (b) challenges they might encounter if they were to pursue a career in fields of emerging technology. After open coding all responses, the first author identified common codes across student responses which we used to systematically recode all responses (Saldaña, 2021).

Perceived Challenges

We found that students perceived several challenges in pursuing careers in emerging technologies. Four of 11 students (36%) shared that a challenge was the *competitive job market* in the field of emerging technology. E.g., one student shared that "Entry-level jobs might be very competitive, so finding a good job might be a difficult task." Another concern among 36% of students—all of whom either identified as a woman or student of color—was that of *gender and racial bias in emerging technology* fields. For example one student shared "I will keep it simple - it is difficult to be a women in STEM. I have been underestimated many times or given the more 'feminine' tasks in project…" Another challenge identified among 36% of students was that, in the field of emerging technology where advances happen rapidly, there is a constant need to improve one's knowledge and learn new skills. For example, a student wrote "...the rapidly evolving nature of emerging technologies demands constant upskilling, which can be both time-consuming and mentally taxing…" Lastly, two students (18%) shared that the content was difficult, e.g., "My background isn't in technology or STEM, so I feel that would be a challenge. I don't have the education or technical training required for those types of careers."

Perceived Benefits

Students also identified several benefits in pursuing careers in emerging technologies. Eight of eleven students (74%) reported that they viewed emerging technology fields to offer *opportunities to develop skills*. For example, a student wrote "The benefits of pursuing a career in emerging technology is that it is emerging which means that there are alot of growth opportunities in this field." Students (36%) also noted that careers in emerging technology offer opportunities to *improve the lives of others*, e.g., "I believe these careers will continue to grow and will allow to make a meaningful impact in people's lives." Lastly, many students (27%) say potential for *financial rewards* in emerging technology fields, e.g., "Apart that are the roles with more pay I got interested in emerging technologies and not falling behind."

Overall, students appeared to offer thoughtful and balanced consideration of potentially pursuing careers in emerging technologies.

Discussion

This study sought to design and evaluate an experiential, interdisciplinary learning intervention that prepares undergraduates—many of whom are first-generation and from historically underrepresented backgrounds—to participate meaningfully in an emerging, AI-driven workforce. This program specifically contributes to educational research on supporting first-generation college students in STEM fields. Namely, our findings offer insight into (a) the impact of experiential learning practices for promoting student interest and confidence, and (b) students perceptions of the benefits and challenges of pursuing careers in emerging technology.

Regarding our first research question pertaining to change from pretest to posttest, results showed that students significantly improved their data-literacy self-efficacy by more than two SDs across all subdimensions, as well as their interest in emerging technologies by two thirds of an SD. These findings align with research demonstrating that experiential learning environments can cultivate students' motivation in technical domains by situating complex concepts within authentic, real-world applications (Hassan et al., 2021; Wu, 2022), and extends this research by offering additional specificity to what students gain: data-literacy efficacy and interest. Models of data-literacy development and motivation similarly highlight the importance of building learners' self-efficacy and perceived value of content as core mechanisms enabling persistence and learning in STEM (Thacker & Herrick, 2025; Kim et al., 2024). Our findings suggest that a structured yet flexible module sequence—balancing technical skills on statistics, SQL, and AI with explicit instruction in leadership, project management, and communication—may be a promising approach for broadening participation in data-intensive careers.

Students' open-ended reflections provided additional insight into how experiential learning shaped their thinking. At the end of the course, many students articulated an awareness of relevant opportunities and challenges associated with entering emerging-technology fields. On the one hand, students frequently emphasized the value of developing concrete technical and professional skills and highlighted the potential for data-driven technologies to improve lives. This aligns with prior work showing that when students perceive STEM learning as socially relevant and potentially impactful, their interest and sense of purpose are strengthened (Hulleman & Harackiewicz, 2021; Harackiewicz et al., 2016; Seyranian et al., 2023; Thacker et al., 2025). On the other hand, students also expressed concerns about the competitive nature of the field, the rapid pace of technological change, and persistent gender and racial inequities in technology sectors. These dual appraisals—optimism about impact and apprehension about barriers—suggest that while experiential instruction can enhance students' readiness, broader systemic and structural challenges remain salient in shaping students' long-term career decisions.

Limitations and Future Directions

Several limitations should be considered when interpreting these findings. First, the current study was a proof of concept, and involved a relatively small sample, which limits statistical power and the precision of estimates. While we selected analyses that optimized our power, future research should aim for recruiting larger numbers of students. Relatedly, the sample was not fully representative of the broader undergraduate population, as participants self-selected into a paid, intensive experiential course and were primarily upper-division STEM majors who identified as Female and Hispanic. While some may view the diversity of our sample as a strength, we acknowledge this also constrains the generalizability of our conclusions. In addition, the absence of a control or comparison group prevents us from attributing observed gains solely to the intervention, as other unmeasured factors may have contributed to changes over time. Future research should consider using a comparison group to enhance internal validity and enable causal inference. Finally, although most measures demonstrated strong internal consistency, some scales warrant further refinement and validation in future work to ensure that they accurately capture the constructs of interest across diverse student populations. Despite these limitations, the study provides promising initial evidence for the potential of experiential, interdisciplinary instruction to enhance students' data-literacy and emerging-technology readiness.

Implications and Conclusion

Overall, our results demonstrate that an interdisciplinary, experiential approach can simultaneously strengthen students' confidence in their data skills, academic motivation, and emerging-technology career readiness. For institutions seeking to diversify the AI and data-analytics workforce, these findings underscore the value of embedding hands-on, context-rich data experiences in undergraduate curricula—particularly for students who may have limited prior exposure. Future research may explore how to scaffold these experiences longitudinally, how mentorship and community-building influence career persistence, and how programs can explicitly address students' concerns regarding equity and competition in technology fields.

References

Acemoglu, D., Autor, D., Hazell, J., & Restrepo, P. (2022). Artificial intelligence and jobs: Evidence from online vacancies. Journal of Labor Economics, 40(S1), S293–S340. https://doi.org/10.1086/718327

Acemoglu, D., & Autor, D. (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings. In Handbook of Labor Economics (Vol. 4, pp. 1043–1171). Elsevier. https://doi.org/10.1016/S0169-7218(11)02410-5

Acemoglu, D., & Restrepo, P. (2020b). The wrong kind of AI? Artificial intelligence and the future of labour demand. Cambridge Journal of Regions, Economy and Society, 13(1), 25–35. https://doi.org/10.1093/cjres/rsz022

Allen, G. I. (2021, March). Experiential learning in data science: Developing an interdisciplinary, client-sponsored capstone program. In Proceedings of the 52nd ACM Technical Symposium on Computer Science Education (pp. 516-522).

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The Skill Content of Recent Technological Change: An Empirical Exploration*. The Quarterly Journal of Economics, 118(4), 1279–1333. https://doi.org/10.1162/003355303322552801

Börner, K., Bueckle, A., & Ginda, M. (2019). Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(6), 1857–1864. https://doi.org/10.1073/pnas.1807180116

Bessen, J. (2019). Automation and jobs: When technology boosts employment. Economic Policy, 34(100), 589–626.

Bile Hassan, I., Ghanem, T., Jacobson, D., Jin, S., Johnson, K., Sulieman, D., & Wei, W. (2021, March). Data science curriculum design: A case study. In Proceedings of the 52nd ACM Technical Symposium on Computer Science Education (pp. 529-534).

Bowler, L., Acker, A., Jeng, W., & Chi, Y. (2017). “It lives all around us”: Aspects of data literacy in teen's lives. Proceedings of the Association for Information Science and Technology, 54(1), 27–35.

Brynjolfsson, E., & McAfee, A. (2014). The second machine age: Work, progress, and prosperity in a time of brilliant technologies (p. 306). W W Norton & Co.

Carlson, J., & Johnston, L. (2015). Data information literacy: Librarians, data, and the education of a new generation of researchers. Purdue University Press.

Ceci, S. J., Williams, W. M., & Barnett, S. M. (2009). Women's underrepresentation in science: sociocultural and biological considerations. Psychological Bulletin, 135(2), 218-261.

Campbell, D. T., Stanley, J. C., & Gage, N. L. (1963). Experimental and quasi-experimental designs for research. Houghton, Mifflin and Company.

Charmaz, K. (2015). Grounded theory. Qualitative psychology: A practical guide to research methods, 3, 53-84.

Chen, F. (2025). The relationship between digital literacy and college students’ academic achievement: the chain mediating role of learning adaptation and online self-regulated learning. *Frontiers in Psychology, 16,* 1590649.

Cheryan, S., Ziegler, S. A., Montoya, A. K., & Jiang, L. (2017). Why are some STEM fields more gender balanced than others?. Psychological Bulletin, 143(1), 1-35.

Dewey, J. (1938). *Experience and education.* New York: Macmillan.

Eccles, J., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), Achievement and achievement motives: Psychological and sociological approaches (pp. 75–146). San Francisco: W. H. Freeman.

Frank, M. R., Autor, D., Bessen, J. E., Brynjolfsson, E., Cebrian, M., Deming, D. J., Feldman, M., Groh, M., Lobo, J., Moro, E., Wang, D., Youn, H., & Rahwan, I. (2019). Toward understanding the impact of artificial intelligence on labor. Proceedings of the National Academy of Sciences, 116(14), 6531–6539. https://doi.org/10.1073/pnas.1900949116

Gottfried, A. E., Marcoulides, G. A., Gottfried, A. W., & Oliver, P. H. (2013). Longitudinal pathways from math intrinsic motivation and achievement to math course accomplishments and educational attainment. *Journal of Research on Educational Effectiveness, 6*(1), 68–92. https://doi.org/10.1080/19345747.2012.698376.

Gould, R. (2017). Data literacy is statistical literacy. *Statistics Education Research Journal*, *16*(1), 22–25. http://iase-web.org/Publications.php?p=SERJ

Hulleman, C. S., Godes, O., Hendricks, B. L., & Harackiewicz, J. M. (2010). Enhancing interest and performance with a utility value intervention. J. of Ed. Psychology, 102(4), 880.

Harackiewicz, J. M., Canning, E. A., Tibbetts, Y., Priniski, S. J., & Hyde, J. S. (2016). Closing achievement gaps with a utility-value intervention: Disentangling race and social class. Journal of Personality and Social Psychology, 111(5), 745.

Hulleman, C. S., & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. science, 326(5958), 1410-1412.

Hulleman, C. S., & Harackiewicz, J. M. (2021). The utility-value intervention. In G. M. Walton & A. J. Crum (Eds.), Handbook of wise interventions: How social psychology can help people change (pp. 100–125). The Guilford Press.

Hulleman, C.S., Wormington, S. V., Tibbetts, C. Y., & Philipoom, M. (2018, August). A meta-analytic synthesis of utility-value interventions in education. Paper presented at the bi-annual meeting of the International Conference on Motivation. Aarhus, Denmark.

Jost, L., & Jansen, P. (2022). Using linear mixed models to analyze learning processes within sessions improves detection of treatment effects: An exemplary study of chronometric mental rotation. Methods in Psychology, 6, 100092.

Kim, J., Hong, L., Evans, S., Oyler‐Rice, E., & Ali, I. (2023). Development and validation of a data literacy assessment scale. Proceedings of the Association for Information Science and Technology, 60(1), 620-624.

Kim, J., Hong, L., & Evans, S. (2024). Toward measuring data literacy for higher education: Developing and validating a data literacy self‐efficacy scale. Journal of the Association for Information Science and Technology, 75(8), 916-931.

Kolb, D. A. (2014). *Experiential learning: Experience as the source of learning and development*. Pearson FT press.

Kolb DA (1984) Experiential learning: Experience as the source of learning and development, Prentice Hall, Englewood Cliffs NJ.

Kosovich, J. J., Hulleman, C. S., Barron, K. E., & Getty, S. (2015). A practical measure of student motivation: Establishing validity evidence for the expectancy-value-cost scale in middle school. The Journal of Early Adolescence, 35(5-6), 790-816.

Lewin, K. (1951). Field theory in social science: selected theoretical papers (D. Cartwright, Ed.). Harpers.

Lohfink, M. M., & Paulsen, M. B. (2005). Comparing the determinants of persistence for first generation and continuing-generation students. Journal of College Student Development, 46(4), 409–428. https://doi.org/10.1353/csd.2005.0040.

Mäkelä, E., & Stephany, F. (2024). Complement or substitute? How AI increases the demand for human skills. arXiv preprint arXiv:2412.19754.

Martin, J. P., Stefl, S. K., Cain, L. W., & Pfirman, A. L. (2020). Understanding first-generation undergraduate engineering students' entry and persistence through social capital theory. International Journal of STEM Education, 7(1), 37.

McGee, E. O., & Martin, D. B. (2011). "You would not believe what I have to go through to prove my intellectual value!" Stereotype management among academically successful Black mathematics and engineering students. American Educational Research Journal, 48(6), 1347-1389.

National Center for Science and Engineering Statistics. (2023). Women, minorities, and persons with disabilities in science and engineering 2023. Arlington, VA. Retrieved from ncses.nsf.gov/pubs/nsf23315

National Science Foundation. (2020). Science and engineering degrees, by race/ ethnicity of recipients: 2008–18. Arlington, VA. Retrieved from https://ncsesdata.nsf.gov/sere/2018/

Pascarella, E. T., Pierson, C. T., Wolniak, G. C., & Terenzini, P. T. (2004). First generation college students: Additional evidence on college experiences and outcomes. Journal of Higher Education, 75(3), 249–284. https://doi.org/10.1353/jhe.2004.0016.

Piaget, J. (1971). Inconscient affectif et inconscient cognitif. Raison Présente, 19(1), 11–20. https://doi.org/10.3406/raipr.1971.1490

Pinheiro, J. C., & Bates, D. M. (2000). Mixed-effects models in S and S-PLUS. New York, NY: Springer New York.

Portes, A. (1998). Social capital: Its origins and applications in modern sociology. Review of Sociology, 24(1), 1–24. https://doi.org/10.1146/annurev.soc.24.1.1
Prado, J., & Marzal, M. A.´ (2013). Incorporating data literacy into information literacy programs: Core competencies and contents. *Libri, 63*(2), 123–134.
Pugh, K. J. (2011). Transformative experience: An integrative construct in the spirit of Deweyan pragmatism. Educational Psychologist, 46(2), 107-121.
Qiao, C., Chen, Y., Guo, Q., & Yu, Y. (2024). Understanding science data literacy: a conceptual framework and assessment tool for college students majoring in STEM. *International Journal of STEM Education*, *11*(1). https://doi.org/10.1186/s40594-024-00484-5
Ridsdale, C., Rothwell, J., Smith, M., Ali-Hassan, H., Bliemel, M., Irvine, D., Kelley, D., Matwin, S., & Wuetherick, B. (2015). *Strategies and best practices for data literacy education: Knowl- edge synthesis report.* Dalhousie University. https://dalspace.library.dal.ca/handle/10222/64578
Rosenzweig, R., Nillegoda, N. B., Mayer, M. P., & Bukau, B. (2019). The Hsp70 chaperone network. Nature reviews molecular cell biology, 20(11), 665-680.
Saldaña, J. (2021). *The coding manual for qualitative researchers.* Sage Publications
Seyranian, V., Thacker, I., Madva, A., Abramzon, N., & Beardsley, P. (2023). A Utility Value Intervention to support undergraduate student interest, engagement, and achievement in calculus and calculus-based physics. Proceedings of the forty-fifth annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education (Vol 2, pp. 464-473). Reno, NV. https://doi.org/10.51272/pmena.45.2023
Thacker, I. (2023). Climate change by the numbers: Leveraging mathematical skills for science learning online. Learning & Instruction, 86, 101782. https://doi.org/10.1016/j.learninstruc.2023.101782
Thacker, I. (2024). Supporting secondary students' climate change learning and motivation using novel data and data visualizations. Contemporary Educational Psychology, 78, 102285. https://doi.org/10.1016/j.cedpsych.2024.102285
Thacker, I., French, H.*, & Feder, S. (2024). Estimating climate change numbers: Mental computation strategies that can support science learning. International Journal of Science Education, 47(1) 1–22. https://doi.org/10.1080/09500693.2024.2307473
Thacker, I., Shroeder, R., Shields-Menard, S., & Goforth, N.* (2025). Dirt Don't Hurt: How relevant soil data can support learning and motivation at a Hispanic Serving Institution. International Journal of Science and Mathematics Education, 23, 803–826.
Thacker, I., Herrick, I. (2025). Data Literacy for Conceptual Change: An Integrated Model With Special Applications in Climate Change Education. In R. M. Zbiek, X. Yao, A. McCloskey, & F. Arbaugh (Eds.), Proceedings of the forty-seventh annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education (pp. 1722–1733). State College, PA. https://doi.org/10.51272/pmena.47.2025
Vu, T., et al. (2022). Motivation-Achievement Cycles in Learning: a Literature Review and Research Agenda. Educational Psychology Review 34 (1), 39–71. https://doi.org/10.1007/s10648-021-09616-7
Wigfield, A., Rosenzweig, E.Q., & Eccles, J.S. (2017). Achievement values: Interactions, interventions, and future directions. In A.J. Elliot, C.S. Dweck, D.S. Yeager (Eds.), Handbook of competence and motivation (2nd ed.), Guilford, New York, NY (2017), pp. 116-134

Wu, W. (2022). Investigating Internship Experiences of Data Science Students for Curriculum Enhancement. Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE, 1, 505–511. https://doi.org/10.1145/3502718.3524741

Appendix A: Summary of Course Materials and Student Data Collected
(Researcher Created)

Table 1. Annual Schedule. *WS: workshop; S: seminar; * indicates an optional training, all else is core training.*			
Date	**Activity**		**Task Leader**
Sep-Dec	Design and implement project activities Confirm Partnerships		All mentors
Oct	Identify eligible participants		All mentors
Nov	Recruit eligible participants and invite applicants		All mentors
Dec	Select participants		All mentors
Jan	Notify participants, Evaluate recruitment efforts		UC
Jan-May	**Externship**		
	Topic	**Objectives and Activities**	
Week 1	Work ethic and Professional development	Foster a respective and collaborative working environment; Working email and professional behaviors; What are the expectation from your managers	Career Center Staff
Week 2	Basic Statistics specific for data preprocessing	Descriptive statistics, visualizations, introduction to inferential statistics. **Deliverable**: Report and interpret descriptive and inferential statistics for authentic data.	Ian Thacker
Week 3	Presentation Skills	Tips and tools for developing and delivering formal and informal presentations. **Deliverable**: Presentation of industry-specific topic.	Ian Thacker
Week 4	Data Analytics and Interpretation	Students will learn how to access, analyze, and interpret large retail datasets using tools like Excel. **Deliverable**: Data analysis where they clean, analyze and visualize an authentic dataset, providing insights into retail trends and suggesting actionable business decisions.	Isil Koyuncu
Week 5	Communication Skills	Sharing complex ideas to various audiences across different modes of communication. General tips for professional listening, speaking, writing, and reading. **Deliverable**: Read and reflect on industry-specific articles and presentations fit for different audiences.	Ian Thacker
Week 6	Data Processing and Management	Students will learn to write basic SQL queries to manage retail data, such as querying sales trends and updating inventory. **Deliverable**: Queries to analyze a sample retail dataset, extract key insights and submit a report summarizing their findings and recommendations.	Isil Koyuncu
Week 7	Collaboration Skills	How and why diverse perspectives can enhance creativity and productivity on collaborative projects. **Deliverable**: Interdisciplinary, industry-specific, collaborative project.	Career Center Staff
Week 8	Programming with Patterns	Sharing python codes to read data in different formats, commonly used packaged for data processing, output data with various format such bar chart, pie chart, box plot, and save projects. **Deliverable:** Analyze authentic data with python and share code.	Yufang Jin

<table>
<tr><td>Week 9</td><td>Writing Reports/Technical Writing Skills</td><td>Strategies for clear and coherent scientific writing. Clear and transparent communication of purpose, audience, research methods, results, and implications. Deliverable: Write an industry-specific technical report.</td><td>Yufang Jin</td></tr>
<tr><td>Week 10</td><td>AI Algorithms and Digital Twins</td><td>Introduce basics for Artificial Neural Network (ANN), run the ANN algorithm for data processing with Python. Deliverable: Analyze authentic data using ANN.</td><td>Yufang Jin & Anthony Rios</td></tr>
<tr><td>Week 11</td><td>Leadership skills</td><td>Strategies for promoting collective motivation, creativity, and constructive feedback in collaborative settings. Deliverable: Lead an activity or discussion.</td><td>Career Center Staff</td></tr>
<tr><td>Week 12</td><td>Digital Twins</td><td>Introduce digital twin applications</td><td>Anthony Rios</td></tr>
<tr><td>*Track 1: Week 13-15</td><td>AI applications</td><td>Hands-on experiences with various financial data, read, write, basic process, plot. Deliverable: Analyze authentic financial data.</td><td>Yufang Jin</td></tr>
<tr><td rowspan="4">*Track 2: Week 13-15</td><td>Smart Retail Engineering</td><td>Students will explore how data analytics drives marketing strategies in retail through readings, case studies, and a hands-on data exercise. Deliverable: Apply tools to analyze customer behavior, measure the effectiveness of retail marketing campaigns, customer segmentation, and personalized marketing strategies.</td><td rowspan="4">Isil Koyuncu</td></tr>
<tr><td>Omnichannel Strategy Implementation</td><td>Students will explore how to integrate physical and digital retail channels to create seamless customer experiences. Deliverable: Propose their own solutions for a fictional retailer and present their omnichannel strategies.</td></tr>
<tr><td>Supply Chain Engineering</td><td>Students will explore the importance of sustainable practices in retail, including but not limited to sustainable sourcing, logistics and operations. Deliverable: Analyze supplier scorecards and critically assess how retailers can implement sustainable strategies in areas like supply chain management and waste reduction.</td></tr>
<tr><td>Supply chain resilience / mitigating disruptions</td><td>Students will explore how transit and supplier disruptions impact lead times and overall supply chain performance. Deliverable: Analyze the provided dataset on disruptions and propose resilience strategies, balancing cost-efficiency and long-term sustainability in supply chain operations.</td></tr>
<tr><td>Jun-Aug</td><td colspan="3">Course Wrap-up, Post-Survey.
Participants engage in summer internships with partners
Evaluation;</td></tr>
<tr><td>Post-program</td><td colspan="3">Participants survey analysis, Networking, Cohort meeting, and summarize the preliminary for funding application and results dissemination;</td></tr>
<tr><td>Field trips</td><td colspan="3">Tower Semiconductor, SWRI, H-E-B headquarter</td></tr>
</table>

Appendix B: Data Literacy Self-Efficacy

(Adapted from Kim, Hong, Evans, 2024)

To what extent do you feel knowledgeable about the following data skills?
1 = *“No knowledge” and 5 = “Expert.”*

1. Defining what data is
2. Describing what different types and formats of data are
3. Explaining the real and potential value of data
4. Collecting primary data, such as surveys, observations, experiments, questionnaires, or interviews
5. Locating secondary data from print/electronic sources, such as journal articles and government publications
6. Accessing and locating secondary data from electronically stored information, such as websites, APIs, and databases
7. Identifying data errors, including corrupted, incorrectly formatted, duplicate, or incomplete data
8. Displaying and summarizing data using descriptive statistics
9. Using inferential statistics to draw meaningful inferences from data
10. Employing predictive analytics to make predictions and direct decisions
11. Applying data mining to explore patterns and trends of data
12. Communicating in writing and verbally the results of data analysis/analytics
13. Selecting appropriate tools to create data visualizations
14. Creating useful and meaningful data visualizations using charts, tables, and graphics using a chosen dataset
15. Interpreting data visualizations to identify points of interest
16. Presenting data insights tailored to a specific audience
17. Using a storyboarding technique to plan the content and structure of a data presentation
18. Using a narrative structure to create compelling data presentations
19. Assessing data sources to ensure they meet the defined need
20. Determining whether data meets the quality required for a given purpose
21. Evaluating data critically for the right type and quantity to support its intended use
22. Categorizing and classifying data to make it more accessible for use
23. Organizing and distributing data among files, including file naming conventions, directory structure, version control
24. Documenting data to improve data findability and accessibility (e.g., codebook, data dictionary, README file, metadata)
25. Differentiating between different data storage devices
26. Identifying various data backup methods and strategies
27. Determining the best storage options for short- and long-term data storage
28. Locating laws, regulations, and guidelines, such as FERPA, HIPAA, and GDPR, for handling data security and privacy
29. Utilizing de-identification techniques to protect personal data
30. Determining data ownership and use rights – i.e., intellectual property rights and licensing
31. Identifying current best practices in data citation, including the use of permanent identifiers, notably Digital Object Identifiers

Subscales:
Items 1–6 = Data Identification
Items 7–23 = Data Processing
Items 24–31 = Data Management and Sharing

Appendix G: Expectancy Value Cost Scale
(Adapted from Kosovich et al., 2015)

1 = Strongly disagree, 2 = Disagree, 3 = Slightly disagree, 4 = Neutral, 5 =Slightly agree, 6 = Agree, 7 = Strongly agree

Please rate the following items.

E1 I know I can learn the material in my [math or science] class.
E2 I believe that I can be successful in my [math or science] class.
E3 I am confident that I can understand the material in my [math or science] class.
V1 I think my [math or science] class is important.
V2 I value my [math or science] class.
V3 I think my [math or science] class is useful.
C1 My [math or science] classwork requires too much time.
C2 Because of other things that I do, I don't have time to put into my [math or science] class.
C3 I'm unable to put in the time needed to do well in my [math or science] class.
C4 I have to give up too much to do well in my [math or science] class.

Kosovich, Jeff J., Hulleman, Chris S., Barron, Kenneth E., & Getty, Steve. (2015). A practical measure of student motivation: Establishing validity evidence for the Expectancy-Value-Cost Scale in middle school. The Journal of Early Adolescence, Vol 35(5-6), 790-816. doi: https://dx.doi.org/10.1177/0272431614556890

Appendix H: STEM Interest

(Adapted from Hulleman et al., 2010; 2009).

Participants will respond to these items on a 7-point Likert-type agreement scale from 1 (strongly disagree) to 7 (strongly agree).

Please respond by stating your agreement with the following statements. Note that the term "emerging technologies" refers to topics such as: digital twins, smart manufacturing, smart engineering, artificial intelligence (AI), data science, machine learning, and robotics.

Initial/Personal Science Interest

1. I think the field of emerging technology is interesting.
2. To be honest, I just don't find the field of emerging technology interesting. (REVERSED)
3. I've always wanted to learn more about the field of emerging technology.

Hulleman, C. S., Godes, O., Hendricks, B. L., & Harackiewicz, J. M. (2010). Enhancing interest and performance with a utility value intervention. *Journal of Educational Psychology, 102*(4), 880.

Hulleman, C. S., & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science, 326*(5958), 1410-1412.

Appendix J: Open Ended Questions

(Researcher Created)

The following questions ask about your thoughts on careers in "**emerging technology.**" In this case, we use the term to refer to STEM fields relating to study of topics such as: *digital twins*, *smart manufacturing*, *smart engineering*, *artificial intelligence (AI)*, *data science,* machine learning, and *robotics*.

Please write a paragraph to explain your perceptions of the ***challenges*** you might encounter if you were to pursue a career in fields of emerging technology?

Please write a paragraph to explain your perceptions of the ***benefits*** of pursuing a career in fields of emerging technology?

Appendix K: Demographics Survey

Posttest Demographics:

- Screener Questions:
  - What is your age? _____ (must be 18 or older to continue)
  - Do you currently live in the USA? [Y/N] (must select Yes to continue)
  - Are you a pre-engineering student? [Y/N] (must select Yes to continue)
  - Do you agree to share your email and complete a follow-up survey at the end of the semester? [Y/N] (must select Yes to continue)
- [Consent materials here] Do you agree to participate? [Y/N] (must select Yes to continue)
- Name: ___________________ (needed to link pretest to posttest)
- abc123: ___________________ (needed to access student records)
- Email: ___________________ (needed to link pretest to posttest)
- Do you currently describe yourself as male, female, transgender, or nonbinary? (Male, Female, Transgender, Nonbinary, Prefer not to say)
- Are you Hispanic, Latino, or of Spanish origin? (Yes, No)
- I identify my race as: (American Indian/Alaska Native, Asian, Black/African-American, Native Hawaiian/Pacific Islander, White/Caucasian, Two or more races, Other)
- Is English your first (native) language? (Yes, No)
- What year of study are you in? (First year, second year, third year, fourth year, other)
- Are you a STEM Major? (No, Yes, Not currently but I plan to eventually enroll in a STEM major, Other)
- If you have any general feedback, please write it in the space below. (Optional)