DIRT DON'T HURT: HOW RELEVANT SOIL DATA CAN SUPPORT LEARNING AND MOTIVATION AT A HISPANIC SERVING INSTITUTION

Ian Thacker, Ph.D.
UT San Antonio
ian.thacker@utsa.edu

Rebecca Schroeder, Ph.D.
UT San Antonio
rebecca.schroeder@utsa.edu

Sara Shields-Menard, Ph.D.
UT San Antonio
sara.shields-menard@utsa.edu

To create opportunities for meaningful applications of data-science for diverse students, we developed and implemented an online learning module focused on engaging students at a Hispanic Serving Institution (HSI) in an analysis of authentic soil data. Development of the module occurred over three design iterations involving interviews with 10 undergraduate STEM students. We then implemented the finalized module in three undergraduate microbiology classrooms (N=118) using a pretest, posttest, comparison group quasi-experimental study design. Findings revealed that, after adjusting for key variables, the intervention group demonstrated significantly greater microbiology knowledge than the comparison group. Path analyses revealed indirect effects of the intervention through value and interest in STEM.

In today's world, data is ubiquitous and data literacy is essential across industries (Börner et al., 2019). However, traditional data science courses are not meeting students' needs (Baumer, 2015), and racial diversity in STEM is lacking (Cruz et al., 2018; Fry et al., 2021; NSF, 2015). This project aimed to enhance STEM learning among Hispanic students by leveraging their cultural resources (Gonzalez et al., 1995; Wilson-Lopez et al., 2016). We developed an online module on soil microbiology, utilizing data science tools to engage students at a Hispanic Serving Institution (HSI) and assess their impact on STEM learning and motivation.

Theoretical Framework

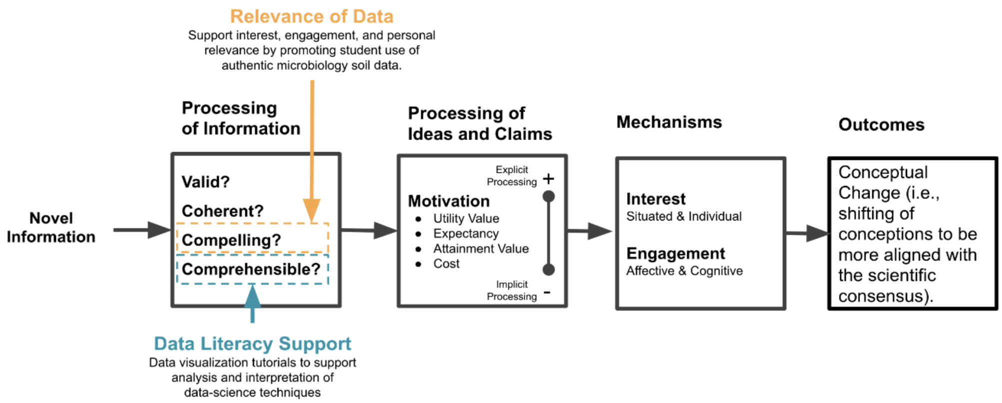
To frame how data visualizations can support science learning for Hispanic students, we integrate theories of Conceptual Change, Data Visualization Literacy, and Expectancy Value. *Conceptual change* theory posits that presenting people with novel information can shift their conceptions about science topics to be more aligned with the scientific consensus (Dole & Sinatra, 1998). For example, the Plausibility Judgments for Conceptual Change model (PJCC; Lombardi et al., 2016) suggests that people process information better if it is *comprehensible*, *coherent* with prior experiences, *compelling*, *relevant*, and stems from *credible* sources. People then judge the plausibility of associated claims and restructure knowledge as a result. Plausibility judgments involve more explicit processing depending on learners' motivation, engagement, and emotion, which then predicts the likelihood that conceptual change will occur.

The *Expectancy Value Theory* (EVT) helps frame motivational factors, and proposes that it is driven by learners' expectancy for success and task value (Eccles et al., 1983; Wigfield et al., 2017). According to EVT, 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. Utility value interventions—especially those that are personal and relevant to students—can significantly enhance value, interest, and learning, particularly for underrepresented students (Harackiewicz et al., 2016).

The *Data Visualization Literacy framework* (DVL-FW; Börner et al., 2019) suggests that data visualizations can help ground abstract concepts. Accordingly, a central process required to interpret data from visualizations is *translating* relevant problems of interest into problems of data. As such, our project aimed to leverage students' knowledge to contextualize data.

We synthesized these theories in a process model (see Figure 1) to inform our intervention development and that we also tested using path model analyses. The model predicts that an intervention intended to expose students to compelling and comprehensible microbiology data would trigger motivational processes (i.e., improve utility value, expectancy, attainment value, and reduced cost), which would activate mechanism variables (student interest and engagement), which would predict increased learning outcomes (i.e., microbiology knowledge). Furthermore, we sought to facilitate change by centering learning around compelling topics that students found relevant, and supporting data literacy to improve data comprehensibility.

Figure 1. *Intervention Logic Model of Conceptual Change, Motivation, and Data Literacy*



To test this theoretical model, we designed and tested an interdisciplinary learning experience for undergraduate students at an HSI. We addressed the following research questions:

- **RQ1.** How can a learning intervention be developed to leverage undergraduate students' motivation for the learning of soil microbiology and data literacy skills?
- **RQ2.** To what extent will such an intervention support students' microbiology knowledge, data literacy skills, engagement, interest, and task value in STEM?
- **RQ3.** Will the hypothesized relationships between task value processes and achievement outcomes be mediated by mechanisms of interest and engagement? (See Figure 1)

This project addressed these research questions through two studies. The first was a formative study focused on creating an online module and the second was a comparative study testing the effectiveness of this intervention using quasi-experimental research.

Study 1: A Design-Based Research Study

To answer the first research question, we used a design-based research (DBR) approach to guide the development and revision of an interactive online intervention (Hoadley & Campos, 2022) on soil microbiology. The re-design, implementation, and revisions occurred over several iterations, resulting in an open-source module for undergraduate microbiology students. The module introduces students to the Tiny Earth Initiative (Hurley et al., 2021), which focuses on identifying new antibiotics in soil. It includes information on the antibiotic resistance crisis, microbial ecology, data visualization tools, and interpretation of soil data visualizations.

We conducted 10 recorded cognitive interviews (Desimone & Le Floch, 2004) via Zoom with a convenience sample of undergraduate students at an HSI in Summer and Fall 2023. Students self-identified as Female (70%), Male (20%), Nonbinary (10%), Hispanic (50%), White (40%), Black (10%), Asian (20%), and English Learners (30%). Interviews focused on student feedback to guide revisions of the module and surveys. Zoom recordings were transcribed and open-coded (Corbin & Strauss, 1990) to examine student engagement and learning.

RQ1 Results: A Module for Soil Microbiology Data Exploration. Revisions focused on improving interactivity, engagement, visual appeal, and ease of data and text interpretation, especially for English learners. The finalized intervention and surveys can be accessed using this link: https://www.softchalkcloud.com/lesson/serve/34E2zGcmQxaWZl/html. Ultimately, this version of the module was used in the second phase of the study.

Study 2: A Quasi-Experimental Study Testing Effects of the Design

To answer the second research question, we recruited 118 undergraduate students from an HSI in a southern U.S. state. Students reported their year of study (1% first year, 13% second year, 38% third year, 38% fourth year, 10% other), gender (76% Female, 21% Male, 1.7% Nonbinary, 1.7% prefer not to say), ethnicity (56% Hispanic), race (1% American Indian/Alaska Native, 13% Asian, 6% Black/African-American, 10% Two or more races, 58% White/Caucasian, 11% Other race), and whether they were enrolled in a STEM major (78% STEM major, 15% not STEM, 3% plan to enroll in a STEM major, 4% Other).

The intervention group consisted of 101 students from two undergraduate microbiology courses and the comparison group consisted of 17 students from a separate course. All participants first completed a 12-item pretest questionnaire on microbiology knowledge and an item that captured students' perceptions of the relevance of data science.

After the pretest, learners either completed the module (treatment group) or continued with "business as usual" (comparison group). All participants then completed an identical post-test of microbiology knowledge and data literacy. Participants also completed a microbiology-specific interest scale (adapted from Hulleman et al., 2010), a Data-Science-specific interest scale (adapted from Hulleman et al., 2010), the Cognitive Engagement scale (Greene, 2015), and a Task Value scale (Kossovich et al., 2015). Internal reliability for all scales at pretest and posttest were judged using Cronbach's alpha, and are reported in Table 1.

Table 1. Descriptive Statistics By Condition and Intercorrelations Between Key Variables

	Total								Control			Treatment		
	items	α	n	Mean	SD	Min	Med	Max	n	Mean	SD	n	Mean	SD
Knowledge (pre)	12	.62	117	0.7	0.2	0.2	0.7	1.0	16	0.5	0.2	101	0.7	0.2
Knowledge (post)	12	.72	117	0.7	0.2	0.2	0.8	1.0	16	0.4	0.2	101	0.8	0.2
Data Relevance (pre)	1	NA	116	4.0	1.0	1.0	4.0	5.0	16	3.6	1.0	100	4.1	0.9
Data Relevance (post)	1	NA	116	4.2	1.0	1.0	5.0	5.0	16	3.4	1.2	100	4.4	0.9
Initial Interest	8	.96	117	5.5	0.8	1.8	5.5	6.8	16	5.1	0.8	101	5.5	0.8
Situated Interest	5	.92	117	5.9	1.1	1.0	6.2	7.0	16	5.2	1.3	101	6.0	1.0
Utility Value	4	.91	117	5.8	1.1	1.0	6.0	7.0	16	5.3	1.3	101	5.9	1.1
Data Science Situated Interest	3	.77	117	4.5	1.3	1.0	4.7	7.0	16	4.7	1.1	101	4.5	1.3
Data Science Utility Value	3	.73	117	5.0	1.0	1.0	5.0	7.0	16	4.9	1.1	101	5.0	1.0
Cognitive Engagement	16	.88	117	3.9	0.6	2.4	4.0	5.0	16	3.7	0.7	101	4.0	0.6
Expectancy	3	.95	117	5.9	1.1	1.0	6.0	7.0	16	5.5	1.3	101	5.9	1.1
Attainment Value	3	.95	117	6.0	1.0	1.7	6.0	7.0	16	5.5	1.2	101	6.1	0.9
Cost	4	.86	117	3.4	1.4	1.0	3.5	7.0	16	4.2	1.5	101	3.3	1.3
Affective Engagement	9	.83	117	3.4	0.6	1.8	3.4	4.9	16	3.1	0.5	101	3.4	0.6

	Intercorrelations												
	k.pre	k.post	dr.pre	dr.post	in.int	sit.int	uv	ds.sit	ds.uv	cog	exp	value	cost
Knowledge (pre)													
Knowledge (post)	.62***												
Data Relevance (pre)	.37***	.28**											
Data Relevance (post)	.30**	.44***	.55***										
Initial Interest	.28**	.38***	.31***	.33***									
Situated Interest	.30***	.33***	.39***	.45***	.88***								
Utility Value	.34***	.33***	.36***	.35***	.67***	.71***							
Data Science Situated Interest	.07	.11	.21*	.12	.41***	.44***	.42***						
Data Science Utility Value	.18*	.24**	.26**	.28**	.47***	.44***	.51***	.62***					
Cognitive Engagement	.14	.23*	.30***	.34***	.63***	.61***	.59***	.27**	.29**				
Expectancy	.17	.25**	.28**	.36***	.80***	.80***	.59***	.34***	.40***	.68***			
Attainment Value	.31***	.38***	.37***	.41***	.82***	.83***	.76***	.35***	.49***	.67***	.82***		
Cost	-.29**	-.27**	-.27**	-.17	-.39***	-.49***	-.25**	-.29**	-.15	-.30**	-.48***	-.38***	
Affective Engagement	.06	.12	.26**	.35***	.40***	.48***	.30***	.45***	.36***	.47***	.43***	.36***	-.22*

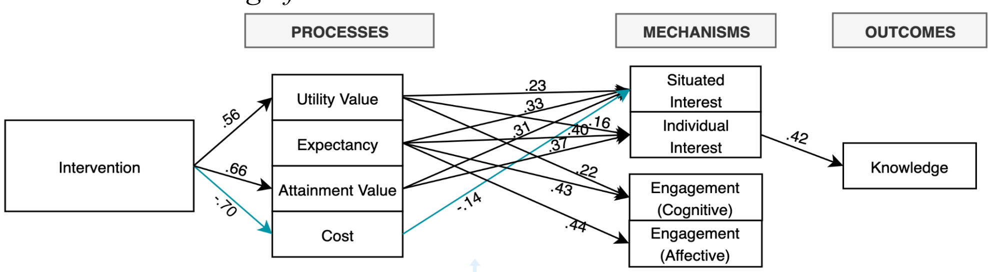
Note: *$p < .05$; **$p < .01$; ***$p < .001$

We assessed if baseline measures differed by condition. Chi-Squared analyses showed gender, ethnicity, race, English speaking status, and STEM status were independent of condition (all $p > .117$). However, year of study ($p < .001$) and pretest knowledge ($p < .001$) were significantly lower in the control group. Thus, we included these as covariates in all analyses. Raw means, standard deviations by condition, and intercorrelations are in Table 1.

RQ2 Results: Module Effects on Learning & Motivation. To assess the module's effects on knowledge, data literacy, engagement, interest, and value in STEM, we used multiple regression with robust standard errors. Predictors included the treatment condition, pre-test scores, and covariates. We predicted the module would improve outcomes due to greater comprehensibility, compellingness, and engaging information (Dole & Sinatra, 1998; Lombardi et al., 2016). Findings showed significant effects on posttest knowledge (β = 1.67, $p < .001$) and data science relevance before adjusting for pretest knowledge and year (β = .66, $p = .012$). The intervention promoted situated interest (β = 0.73, $p = .014$), value of science (β = 0.66, $p = .029$), and reduced perceptions of cost (β = -0.70, $p = .013$), but effects on initial interest and utility value were marginal. No significant moderation effects of gender or ethnicity interactions were found. A summary of the standardized regression coefficients, standard errors, and p-values for analyses can be found in the Supplemental Materials, Table S1 and Table S2.

RQ3 Results: Path Analysis. We tested a model predicting process variables (utility value, expectancy, attainment value, cost), followed by mechanism variables (STEM interest), engagement (cognitive, affective), and academic outcomes (microbiology knowledge). Pretest knowledge and year of study were covariates. The model had satisfactory fit (RMSEA=.081, SRMR=.073, CFI=.982, TLI=.943, AIC=2209, Chi-Square=35, df=20; Hu & Bentler, 1999).

Figure 3. *Path Model with Significant Paths Shown*



Note. Only paths that are significant at the .05 level are shown, blue paths are used when coefficients are negative. All variables shown represent values at posttest. All coefficients represent standardized βs.

Figure 3 shows the full path model with all coefficients. The intervention influenced motivation processes, significantly affecting reported utility value, attainment value, and cost. These motivational process variables were associated with mechanism variables: utility value was positively associated with situated interest, individual interest, and cognitive engagement; expectancy positively predicted situated interest, initial interest, cognitive and affective engagement; attainment value positively predicted situated and initial interest; cost negatively predicted situated interest. Of the mechanism variables only individual interest significantly predicted microbiology knowledge.

Significance

We aimed to develop and test an online learning module for microbiology students at an HSI to enhance STEM learning and motivation. Students using the module had greater posttest knowledge, perceived data-science relevance, science interest, situated interest, utility value, decreased perceptions of cost compared with a comparison group. Findings support prior research showing that interventions supporting perceptions of utility can enhance motivation and achievement in STEM. We also tested predictions by Hulleman and Harackiewicz (2021), finding that value perceptions significantly predicted psychological mechanisms, which predicted achievement outcomes. The intervention had marginally significant indirect effects on achievement through attainment value and individual interest, indicating that interest and engagement may be important mechanisms for expectancy and value.

In summary, this study supports a long-term agenda focused on interdisciplinary STEM applications for Hispanic students, contributing to theory and practice by testing conceptual change models, exploring mathematical reasoning in science learning, and producing a shareable intervention for science instructors and the public.

References

Baumer, B. (2015). A data science course for undergraduates: Thinking with data. *The American Statistician*, 69(4), 334-342.

Börner, K., Bueckle, A., & Ginda, M. (2019). Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments. *Proceedings of the National Academy of Sciences*, 116(6), 1857-1864.

Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology, 13*(1), 3-21.

Cruz, A., Selby, S., & Durham, W. (2018). Place-based education for environmental behavior: a 'funds of knowledge' and social capital approach. *Env. Ed. Res., 24*(5), 627-647.

Desimone, L., & Le Floch, K.(2004). Are we asking the right questions? Using cognitive interviews to improve surveys in education research. *Ed. Eval. Policy Analysis, 26* (1), 1-22.

Dole, J. A., & Sinatra, G. M. (1998). Reconceptualizing change in the cognitive construction of knowledge. *Educational Psychologist, 33*(2-3), 109-128.

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.

Fry, R., Kennedy, B., & Funk, C. (2021). STEM jobs see uneven progress in increasing gender, racial and ethnic diversity. *Pew Research Center Science & Society.*

Greene, B. A., & Miller, R. B. (1996). Influences on achievement: Goals, perceived ability, and cognitive engagement. *Contemporary Educational Psychology, 21*(2), 181-192.

Gonzalez, N., Moll, L. C., Tenery, M. F., Rivera, A., Rendon, P., Gonzales, R., & Amanti, C. (1995). Funds of knowledge for teaching in Latino households. *Urb. Ed., 29*(4), 443-470.

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.

Hoadley, C., & Campos, F. C. (2022). Design-based research: What it is and why it matters to studying online learning. *Educational Psychologist, 57*(3), 207-220.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55.

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.

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.

Hurley, A., Chevrette, M. G., Acharya, D. D., Lozano, G. L., Garavito, M., Heinritz, J., ... & Handelsman, J. (2021). Tiny earth: a big idea for STEM education and antibiotic discovery. *MBio, 12*(1), 10-1128.

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.

Lombardi, D., Nussbaum, E. M., & Sinatra, G. M. (2016). Plausibility judgments in conceptual change and epistemic cognition. *Educational Psychologist, 51*(1), 35-56.

National Science Foundation (2015). *Science and engineering degrees, by race/ethnicity of recipients: 2002-12.* Arlington, VA.

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

Wilson‐Lopez, A., Mejia, J. A., Hasbún, I. M., & Kasun, G. S. (2016). Latina/o adolescents' funds of knowledge related to engineering. *J. of Engineering Education, 105*(2), 278-311.