# HOW MUCH STUDENTS VALUE AN INTRODUCTORY STATISTICS COURSE, HOW VALUE LEVELS CHANGE ACROSS THE TERM, AND HOW THEY PREDICT LEARNING

Claudia C. Sutter[1,2], Mary Tucker[2]; Karen Givvin[2], Chris S. Hulleman[1]
[1]University of Virginia, USA; [2]University of California, Los Angeles, USA
ccs3z@virginia.edu

*Motivational processes are a promising avenue for addressing concerns related to interest, persistence, engagement, and learning in STEM courses. We examined changes in utility value in an introductory statistics course overall, by sex and underrepresented racial minority (URM) and tested the relationship between utility value and learning. Data were collected from 810 students, including their perceived utility value of the course, quiz scores, and course grades. Utility value declined from the beginning to the middle of the course. Significant differences were found by URM status, but not by sex. While URM students continued to experience a decline towards the end of the course, White and Asian students did not. Students' utility value ($t_3$) and their formative assessments ($t_2$ and $t_3$) predicted their final grade. The findings highlight the need for a deeper understanding of short-term relationships between motivation, learning, and performance as well as the ongoing concern for underrepresented groups in STEM.*

## INTRODUCTION

The decline in student motivation in science, technology, engineering, and math (STEM) courses continues to pose a major challenge and concern in the United States (Young et al., 2018). Whether students perceive what they are learning as useful and valuable plays a critical role during college and can support students in learning STEM concepts, elicit positive attitudes towards computer programming, and ultimately increase their persistence to pursue STEM courses (Rosenzweig et al., 2019). Often, a single STEM course in college shapes motivational beliefs and future choices about whether to continue to enroll in STEM courses or not. In other words, short-term events such as a single class in STEM, can have long-term consequences in education (Kosovich et al., 2017). However, most research on motivation in the educational context has considered change that spans over a long period of time, neglecting the "short-term dynamic processes" (Kosovich et al., 217, p. 130) within a single course. A better understanding of students' motivational trajectories over the duration of a single course and how it relates to learning process and outcomes is critical to improving students' persistence and will enable us to better design learning contexts and opportunity structures to support students from traditionally marginalized and minoritized backgrounds (e.g., Gray et al., 2018).

The present study investigated changes in utility value within a single introductory statistics course and tested the associations between utility value, formative, and summative learning outcomes. Specifically, we sought to address the following three research questions: (1) What are students' incoming levels of perceived utility value and how do they change over the course of the class? (2) Are there differential trajectories of utility value among subgroups of students (e.g., differentiated by sex and URM)? (3) What are the associations between utility value, and formative and summative learning outcomes?

(1) Based on prior findings (Jacobs et al., 2002; Kosovich et al., 2017), decreasing levels of utility value were hypothesized. (2) Due to the dearth of prior research on motivational trajectories across demographic subgroups, no specific hypotheses about differential trajectories are made. (3) With prior research indicating that utility value relates to effort, persistence, and performance (Hulleman et al., 2008), we expected that utility value relates to formative and summative learning.

## METHOD

### *Participants and Procedure*

The sample included 810 non-computer science majors from one university in California, who used an interactive online textbook as part of an introductory course in statistics. Participants were primarily female (71.1%) and 37.3% identified as Asian, 27.2% White, 18% Latinx, 3% African American, and 14.6% either did not disclose their race or reported another race. We created a variable inclusive of underrepresented racial minorities (URM) to use in our analyses (0 = White or Asian; 1 = African American, Black, Hispanic, Indian Subcontinent, Native American, Mixed race, or Greater Middle Eastern). Students of mixed race were included in the URM group, unless their race was White and Asian. Students from six courses completed a survey at the beginning of the course ($t_1$: prior to chapter 1), mid-course ($t_2$: end of chapter 8) and at the end of the course ($t_3$: after chapter 12, $t_3$). Given the nested nature of the data, intraclass correlations (ICCs) were calculated for utility value ($ICC_{t1}$ = .019; $ICC_{t2}$ = .012; $ICC_{t3}$ = .034) and formative learning ($ICC_{t1}$ = .028; $ICC_{t2}$ = .012; $ICC_{t3}$ = .040). Because these ICC values were below the level of triviality defined by Lee (2000; ICC < .10), and our research questions were focused on student-level indicators, multi-level models were not required.

*Measures*

Perceived *utility value* was assessed using the average of two items ('The content of this course is important for me' ; 'What I learn in this course will be useful in the future'; Kosovich et al., 2015) at three time points ($\alpha\_t_1$ = .84; $\alpha\_t_2$ = .85; $\alpha\_t_3$ = .85), rated on a five-point Likert scale from 1 (strongly disagree) to 5 (strongly agree). *Formative learning outcomes* included students' scores on the practice quiz at the end of each of the 11 chapters; scores on chapter 1 reflected time point 1, scores on chapters 2-8 reflected time point 2, and scores on chapters 9-11 reflected time point 3. The *summative learning outcome* measure, final course grade, was provided by one instructor at the end of the semester and available for 241 students.

*Missing Data*

Missing data was low for utility value (between 3.1% at $t_1$ to 6.5% at $t_3$) and the formative quizzes (between 1.4% and 4%). Full Information Likelihood Estimation (FIML) (Muthén & Muthén, 1998-2012) was estimated for the latent change and SEM models.

*Analysis*

*Preliminary Analysis: Measurement Invariance.* To ensure that the latent construct (i.e., utility value) is being measured in the same way over time (Widaman & Reise, 1997), four models were computed: Model 1 (configural invariance) included the same factor structure over time without constraints on factor loadings or intercepts. Model 2 (weak invariance) constrained the factor loadings to be equal across time points. Model 3 (strong invariance) required the factor loadings and the item intercepts to be invariant over time. Model 4 (strict invariance) constrained the item residual variances to be equal over time. Measurement invariance analyses suggest strong invariance, which allows comparing latent means over time (Cheung & Rensvold, 2002).

*Descriptive Analysis.* To identify incoming and changing levels of utility value overall as well as by sex and URM, paired sample *t*-tests and Cohen's *d* were computed in SPSS.

*Latent Change Model.* To examine the change from the beginning of the course to the middle of the course (Change 1) and from the middle to the end of the course (Change 2) as well as the interactions with sex and URM, a neighbor change model was specified while controlling for prior GPA (see Figure 1 for conceptual model; Geiser, 2012).

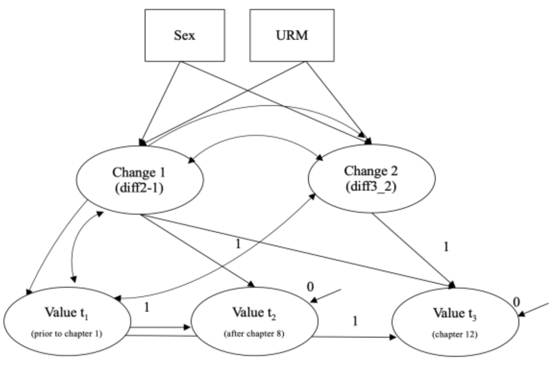


Figure 1. Neighbor Change Model with the Proposed Effects on Students' Utility Value

*Structural Equation Model.* To explore the relationships between utility value, formative assessment, and summative assessment (course grades) among a subset of students, a model relating students' utility value to formative assessment and vice versa as well as their summative assessment was specified (see Figure 2), while controlling for sex and URM.

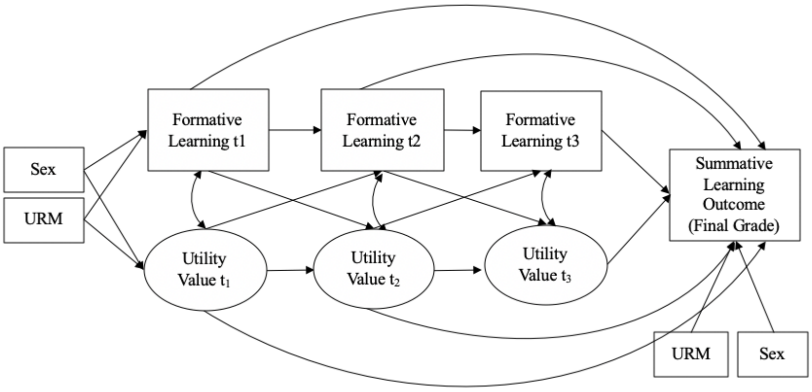


Figure 2. Model relating utility value and formative and summative learning outcomes

RESULTS

*Descriptive Statistics and Correlations*

Means, standard deviations, and correlations between utility value and formative and summative learning outcomes are reported in Table 1. Students' perceived utility value at the beginning of the course was relatively high and not related to formative learning outcomes during the course ($t_1$, $t_2$, or $t_3$) or the final course grade. However, utility value in the middle of the course ($t_2$) was significantly related to their formative learning ($t_2$) and towards the end of the course ($t_3$). Utility value at the end of the course ($t_3$) was significantly related to formative learning outcomes throughout the course as well as their summative learning outcome. Finally, formative learning outcomes were significantly related to the summative learning outcome.

Table 1. Descriptive Statistics and Correlations

| | *n* | Mean | SD | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|---|---|---|
| (1) Utility Value $t_1$ | 785 | 4.22 | .77 | 1 | | | | | | |
| (2) Utility Value $t_2$ | 777 | 3.86 | .98 | .43** | 1 | | | | | |
| (3) Utility Value $t_3$ | 757 | 3.83 | .97 | .32** | .54** | 1 | | | | |
| (4) Formative Learning $t_1$ | 799 | .86 | .10 | -.04 | .06 | .09** | 1 | | | |
| (5) Formative Learning $t_2$ | 810 | .69 | .16 | .03 | .21** | .26** | .38** | 1 | | |
| (6) Formative Learning $t_3$ | 800 | .55 | .20 | .01 | .19** | .21** | .24** | .72* | 1 | |
| (7) Final Grade | 240 | 3.86 | .56 | -.02 | .12 | .21** | .29** | .48** | .40** | 1 |

*Note. **p < .01 level, *p < .05 level*

Descriptive statistics overall as well as by sex and URM, including paired sample t-tests and Cohen's *d* are reported in Table 2. The overall mean change for Change 1 was significant, whereas it was not significant for Change 2, indicating that, on average, students experienced a decline in their perceived utility value from the beginning to the middle of the course but not from the middle to the end of the course. Although mean levels (Change 1) declined for all demographic groups, there were differential rates of decline. More strikingly, whereas the mean levels of utility value continued to decline for URM students, they did not continue to decline for non-URM students.

Table 2. Statistical values (t, df, p) of t-test for paired samples (t1-t2 and t2-t3) of utility value and Cohen's d (d) for the overall sample as well as by sex and URM

| | Change 1 ($t_2$-$t_1$) | | | | | | Change 2 ($t_3$-$t_2$) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *n* | t1 *M* (*SD*) | t2 *M* (*SD*) | *t(df)* | *p* | *d* | *n* | t2 *M* (*SD*) | t3 *M* (*SD*) | *t(df)* | *p* | *d* |
| **Overall** | 754 | 4.22 (.76) | 3.86 (.98) | 10.392 (753) | <.001 | .41 | 734 | 3.86 (.97) | 3.83 (.96) | 1.021 (733) | .308 | .03 |
| **By Sex** | | | | | | | | | | | | |
| Female | 555 | 4.23 (.77) | 3.88 (.98) | 8.820 (554) | <.001 | .40 | 539 | 3.88 (.97) | 3.81 (.96) | 1.706 (538) | .089 | .07 |
| Male | 175 | 4.16 (.74) | 3.80 (1.00) | 4.823 (174) | <.001 | .32 | 168 | 3.81 (.98) | 3.90 (.97) | -1.224 (167) | .223 | -.09 |
| **By URM** | | | | | | | | | | | | |
| URM | 234 | 4.27 (.75) | 3.84 (1.01) | 7.329 (498) | <.001 | .48 | 225 | 3.85 (.99) | 3.70 (.98) | 2.270 (224) | .024 | .15 |
| Non-URM | 499 | 4.20 (.76) | 3.88 (.97) | 7.312 (233) | <.001 | .37 | 484 | 3.88 (.96) | 3.90 (.94) | -.569 (483) | .570 | .02 |

*Latent change models - change in utility value and interactions with sex and URM*

The model fit for the specified model was good: ($\chi^2$ (22) = 37.040, *p* = .0234; CFI = 0.993, TLI = 0.989, RMSEA = 0.030, SRMR = 0.026). Individual differences in latent change were not significantly explained by sex (for both Change 1 and Change 2), indicating that the change in utility value for male and female students are not significantly different. The story was different by URM status. Whereas the regression coefficients for the regression of Change 1 on URM was not statistically significant ($\beta$ = -0.050, *p* = .197), it was for the regression of Change 2 ($\beta$ =-0.112, *p* = .011), implying that there is a significant interaction between time and URM aligning with the descriptive statistics showing whereas both URM and non-URM students experience a decline in Change 1, only URM students continue to experience a decline in Change 2.

*Associations between utility value, formative and summative learning outcomes - Exploratory analysis using a subsample of students*

The fit of the model (see Fig. 2) relating utility value, formative and summative learning to one another was acceptable ($\chi^2$ (32) = 65.949, *p* = 0.0004; CFI = 0.960, TLI = 0.918, RMSEA = 0.067, SRMR = 0.047). Overall, students' utility value at a given time predicted their subsequent utility value and scores on a formative assessment at a given time predicted their scores on the subsequent formative assessments. Utility value and formative learning outcomes in the middle of the course were significantly related. Utility value and formative learning outcomes at the end of the course ($t_3$) positively predicted students' summative learning outcome. Moreover, formative learning outcomes in the middle of the course directly predicted their summative learning outcome.

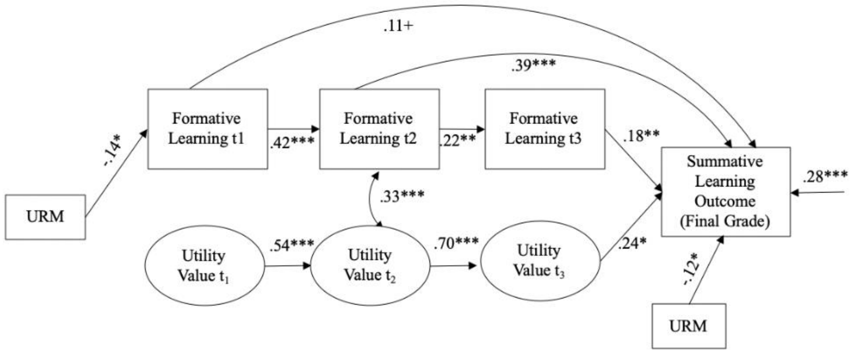


Figure 3. Path model of the significant paths (standardized coefficients) between utility value, formative and summative learning outcome; ***p < .001 level; **p < .01 level, *p < .05 level

DISCUSSION

The present study contributes to prior research by exploring utility value trajectories within the context of a single introductory college statistics course while accounting for sex and underrepresented racial minority status. Students reported relatively high levels of utility value at the beginning of the course. As expected, utility value significantly changed over the term with a significant decline from the beginning to the middle of the course. We did not find evidence of sex differences in the trajectories of utility value, however, there were significant differences by URM status: While URM students continued to experience a decline, Asian and White students did not. The present study thus adds to prior research by highlighting the ongoing concern for underrepresented groups in the sciences. The differential motivational trajectories further highlight the "need to understand students' specific motivational needs" (Robinson et al, 2019), especially among demographic subgroups and in the context of online statistics education. Future research should simultaneously examine the trajectories of different motivational variables, e.g., utility value *and* expectancy levels, to determine how "one construct fluctuates during a semester may be related or unrelated to how the other fluctuates" (Kosovich et al., 2017, p. 132), while accounting for sex, race, and URM status. Further, because different types of values (i.e., intrinsic, utility, attainment value) have been suggested to differentially predict learning outcomes, examining change in said different facets of values would inform education and intervention practice (Kosovich et al., 2017).

Students' perceived utility value of the course ($t_3$) and their formative learning outcomes ($t_2$ and $t_3$) predicted their summative learning outcome (i.e., course grade), and utility value and formative learning outcomes at $t_2$ were significantly associated. Neither incoming perceived utility value of the course nor utility value mid-course predicted their final course grade. Although utility value perceptions have been found to be directly linked to performance (e.g., Wigfield & Eccles, 2002), several studies have shown that they are stronger predictors of future intentions and choice related behaviors (e.g., enrolling in STEM courses) and continued interest, whereas success expectancies within the EVT framework are stronger predictors of performance (Acee & Weinstein, 2010). Thus, future research should simultaneously examine students' course related expectations and utility value as well as include choice related behaviors such as students' intentions and interests to enroll in a statistics course in the future. Overall, however, this pattern of results is compelling with regard to the malleability of utility value as well as implications for motivational interventions since they suggest that whereas students' incoming utility value perceptions about statistics may not serve as a predictor of course outcomes, utility value can change as a result of students' experiences in a course (Rosenzweig et al., 2019). In other words, instructors and instructional material can impact how students' motivation changes throughout the term (Young et al., 2018). Over the past decade, there has been a growing body of research focused on improving student learning and learning outcomes, especially in STEM subjects, by implementing EVT-based interventions (Hulleman et al., 2010; Kosovich et al., 2019). Such interventions are based on perceptions of utility value, facilitated through the emphasis of the relevance of the course topic or coursework for students' future careers. When students believe what they are learning in a course is useful, relevant, and applicable to their lives, they tend to be more interested in the course topic, become more engaged in the material, and more successful in class (Hulleman et al., 2008). Because such interventions have been suggested to be most effective for students who have low competence beliefs or have a history of poor performance (Hulleman et al., 2008), URM students could potentially benefit in particular from utility value interventions.

There are several limitations to our research. First, data were included from one institution and one introductory statistics course only, potentially limiting the generalizability of the findings to students at selective institutions. Second, utility value was only measured using two items, which can undermine reliability and validity. Third, only three time points were collected in this study. Using more than three time points would allow us to run more sophisticated models. Given the ongoing nature of the implementation of the online statistics book, we have since added survey time points into the textbook and can expect up to 14 measurement points - follow up studies will allow us to better understand the reciprocal interplay between the variables.

REFERENCES

Acee, T. W., & Weinstein, C. E. (2010). Effects of a value-reappraisal intervention on statistics students' motivation and performance. *The Journal of Experimental Education*, *78*(4), 487-512. 78(4), 487-512. https://doi.org/10.1080/00220970903352753

Gray, D. L., Hope, E. C., & Matthews, J. S. (2018). Black and belonging at school: A case for interpersonal, instructional, and institutional opportunity structures. *Educational Psychologist, 53*(2), 97-113. https://doi.org/10.1080/00461520.2017.1421466

Hulleman, C. S., Durik, A. M., Schweigert, S. B., & Harackiewicz, J. M. (2008). Task values, achievement goals, and interest: An integrative analysis. *Journal of Educational Psychology, 100*(2), 398-416. https://doi.org/10.1037/0022-0663.100.2.398

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), 880–895. https://doi.org/10.1037/a0019506

Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child development*, *73*(2), 509-527. https://doi.org/10.1111/1467-8624.00421

Kosovich, J. J., Flake, J. K., & Hulleman, C. S. (2017). Understanding short-term motivation trajectories: A parallel process model of expectancy-value motivation. *Contemporary Educational Psychology,* 49, 130-139. https://doi.org/10.1016/j.cedpsych.2017.01.004

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. https://doi.org/10.1177/0272431614556890

Robinson, K. A., Lee, Y.-k., Bovee, E. A., Perez, T., Walton, S. P., Briedis, D., & Linnenbrink-Garcia, L. (2019). Motivation in transition: Development and roles of expectancy, task values, and costs in early college engineering. *Journal of Educational Psychology*, *111*(6), 1081–1102. https://doi.org/10.1037/edu0000331

Rosenzweig, E. Q., Hulleman, C. S., Barron, K. E., Kosovich, J. J., Priniski, S. J., & Wigfield, A. (2019). Promises and pitfalls of adapting utility value interventions for online math courses. *The Journal of Experimental Education*, *87*(2), 332-352. https://doi.org/10.1080/00220973.2018.1496059

Widaman, K. F., & Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance abuse domain. In K. J. Bryant, M. Windle, & S. G. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 281-324). https://doi.org/10.1037/10222-009

Wigfield, A., & Eccles, J. S. (2002). The development of competence beliefs, expectancies for success, and achievement values from childhood through adolescence. In A. Wigfield & J. S. Eccles (Eds.), *A Vol. in the educational psychology series. Development of achievement motivation* (p. 91–120). Academic Press. https://doi.org/10.1006/ceps.1999.1015

Young, A. M., Wendel, P. J., Esson, J. M., & Plank, K. M. (2018). Motivational decline and recovery in higher education STEM courses. *International Journal of Science Education*, *40*(9), 1016-1033. https://doi.org/10.1080/09500693.2018.146077