

Confidence disparities: Pre-course coding confidence predicts greater statistics intentions and perceived achievement in a project-based introductory statistics course

Janet Rosenbaum, Ph.D. (janet.rosenbaum@downstate.edu), Lisa Dierker (ldierker@wesleyan.edu)

Department of Epidemiology and Biostatistics, SUNY Downstate Health Sciences University, Brooklyn
Department of Psychology, Wesleyan University, Middletown, CT

Project-based statistics and self-efficacy

Minoritized young adults are under-represented in STEM careers.

Past research suggests that the project-based course attracts under-represented students who may not otherwise take a statistics courses and improves their interest in further statistics courses.

An original research project is a performance accomplishment, so a project-based course can potentially improve students' self-efficacy.

This study evaluates whether pre-course self-efficacy predicts success of a project-based course: motivation to take further quantitative courses and perceived achievement.

Course description

Students are enrolled in project-based statistics at 29 universities, but project-based statistics has common elements.

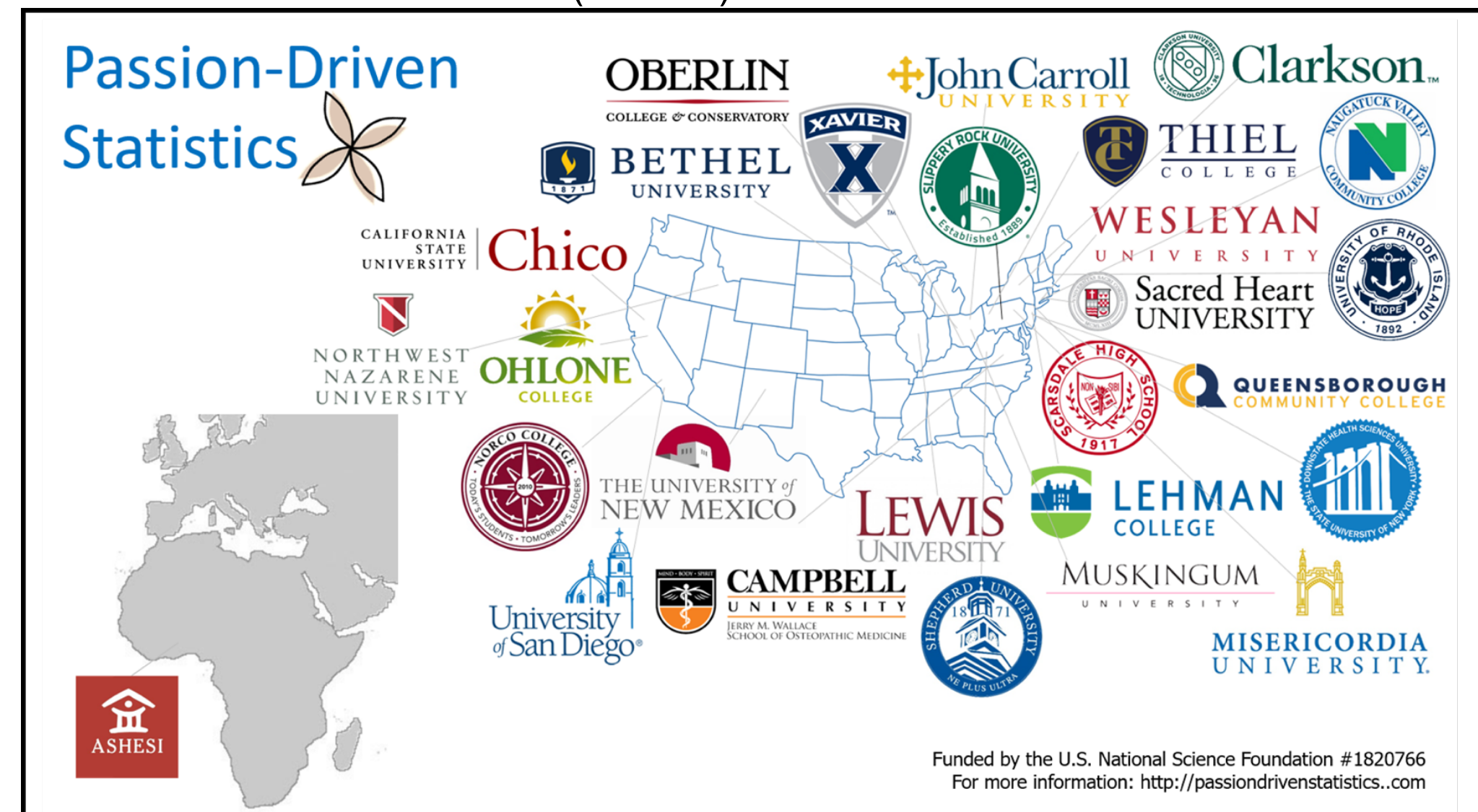
All students complete an individual research project using the analysis of real data using a statistical package.

Text driven: Stata, R, SAS
Menu driven: SPSS, StatCrunch

Materials available at <http://passiondrivenstatistics.com/>

Settings

We describe a project-based introductory statistics course taught in 29 courses at 29 universities (n=917) between Fall 2018 and Winter 2020



11 private liberal arts colleges, 3 flagship state universities, 12 regional city or state universities, and 2 community colleges, located in the United States, and 1 non-profit private university located in Ghana.

Data

Students took self-administered web surveys at the start and end of each semester.

The full data comprised 917 students from 27 classes.

Missing data were due to item non-response, so we used multiple imputation with 20 imputations in a multivariate normal model.

Measures

Perceived achievement = 28 items answering question "How much did you gain in the following areas as a result of your experiences in this course?", Cronbach's alpha = 0.97

Statistical intentions = 13 items, Cronbach's alpha = 0.93
Example item "Are you interested in pursuing advanced coursework in statistics or data analysis?"

Abstract

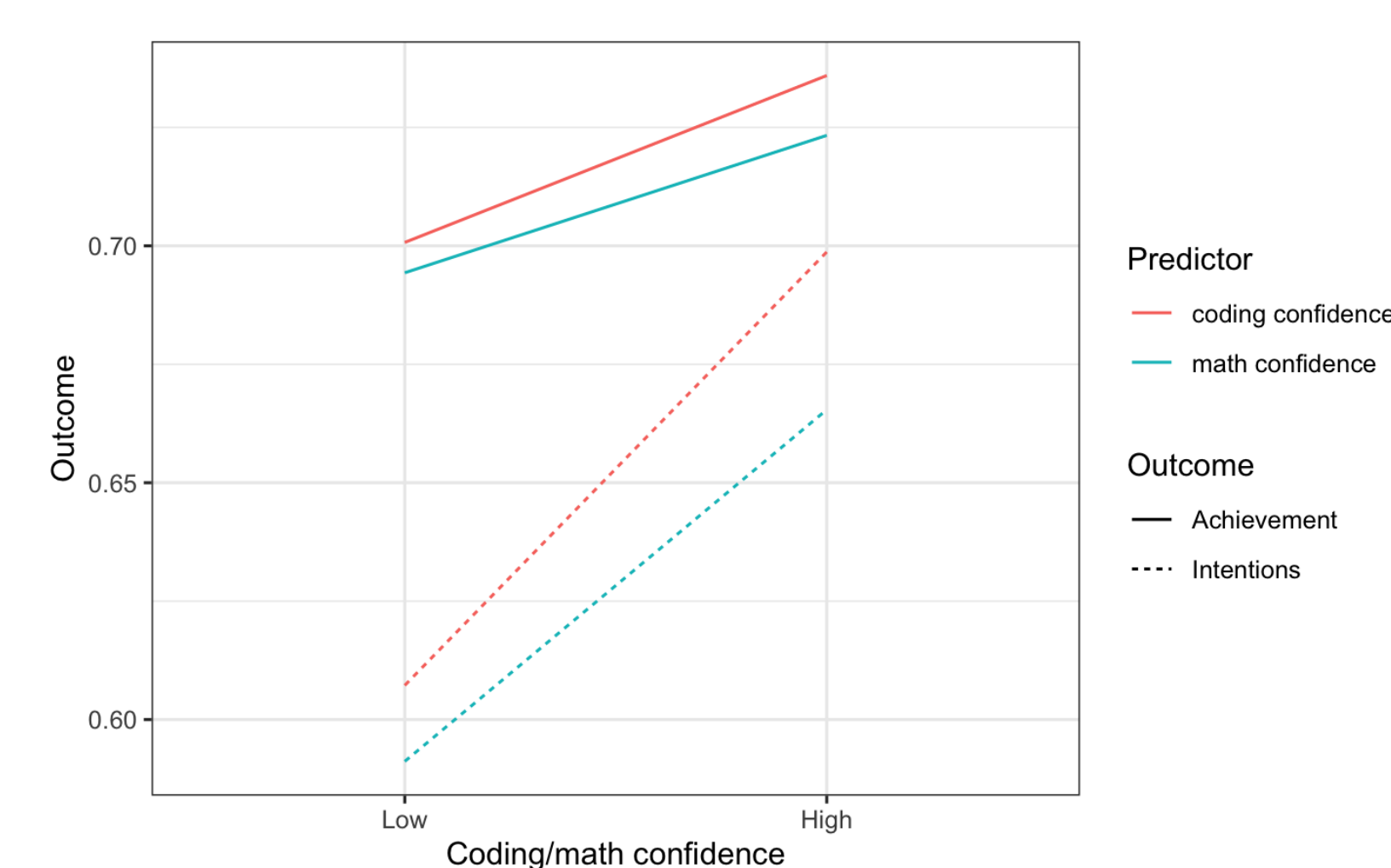
Self-efficacy is associated with many educational outcomes. This study explored associations between pre-course math confidence and coding confidence with post-course statistical intentions and perceived achievement among students in a project-based statistics course at 29 private and public colleges and universities in Fall 2018 to Winter 2020 (n=917) using multilevel mixed-effects multivariate linear regression within multiply imputed data. Pre-course coding confidence and math confidence were associated with greater post-course statistical intentions (coding confidence coefficient: 0.13, 95% confidence interval (0.08, 0.19)); math confidence coefficient: 0.13, 95% CI (0.07, 0.20)). We found greater pre-course coding confidence is associated with higher post-course perceived achievement (95% CI (0.05, 0.17)). Exploratory analyses suggest the association between pre-course math confidence and post-course perceived achievement may be modified by under-represented minority status. Interventions to increase math and coding confidence and reduce stereotype threat may increase students' post-course motivations and perceived achievement and reduce racial and ethnic disparities.

Pre-course math and coding confidence

Pre-course: 45% reported high math confidence and 19% reported high coding confidence

	Confidence Coding	Math
Male	+	+
Age > 22 y	+	
Hispanic	+	
Coding experience	+	+
High school/AP statistics	+	+
Past college statistics	+	
Course uses R	+	+
Course uses SPSS	-	-
Flagship university	+	
Community college		-
Ghana university		+

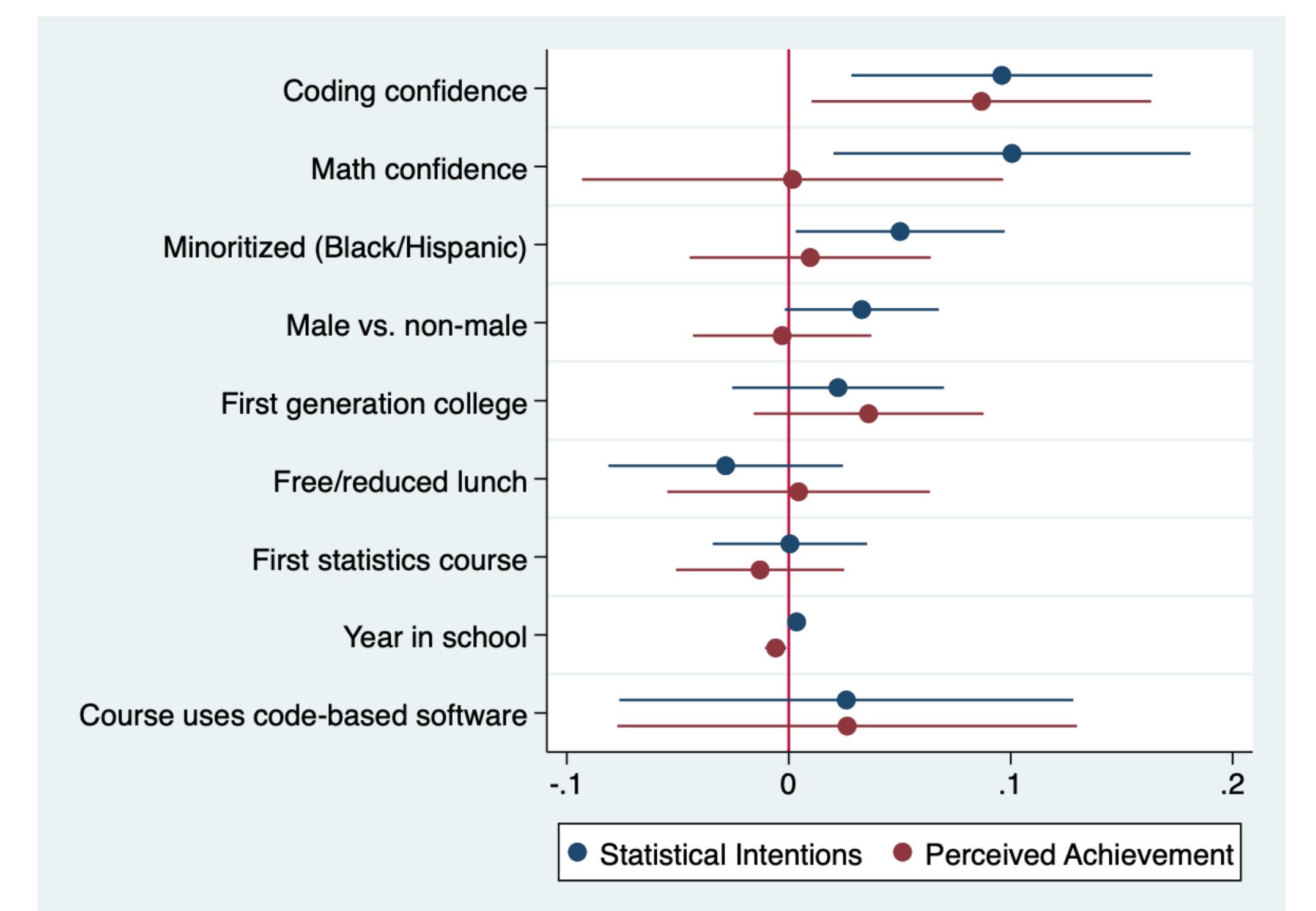
Self-efficacy associated with gains



Statistical analysis

We formulated our model in the data from Fall 2018-Summer 2019 (n=291) and tested the model in Fall 2019-Winter 2020 data (n=624, 21 groups).

We used a mixed effects OLS regression with varying intercepts by course.



Code-based software = Stata, SAS, or R versus SPSS or StatCrunch
The model was formulated in the prior data and applied to these data. These data were multiply imputed with 20 imputations.
Residual standard deviation = 0.19 (0.18, 0.20) and 0.20 (0.18, 0.21)

Varying slopes and varying intercepts

We performed mixed-effects varying intercepts and slopes analysis.

We found little heterogeneity across universities/courses in the slope of coding confidence and math confidence.

Few universities/courses differed in the association between self-efficacy and statistical intentions or perceived achievement.

Conclusions

This course attracts under-represented students who may not otherwise take statistics courses and improves their interest in further statistics courses, according to past research.

Project-based course participants had intentions to take more statistics courses, but students with greater pre-course math and coding self-efficacy experienced greater gains.

Project-based courses could benefit from using brief, evidence-based interventions at the start of the term to improve coding and math self-efficacy.

Pre-college (e.g., pre-frosh) coding experiences may improve coding self-efficacy.