

Evaluating an Adaptive Equity-Oriented Pedagogy on Student Collaboration Outcomes Through Randomized Controlled Trials

Andrew Estrada Phuong, University of California Berkeley, andrew.e.phuong@gmail.com
Judy Nguyen, Stanford University, judynguyen@stanford.edu

Abstract: This study evaluates an adaptive equity-oriented pedagogy (AEP) through a randomized controlled trial. AEP employs evidence-based teaching practices (e.g., formative assessment, universal design for learning) to address college students' diverse learning needs (i.e., their strengths, interests, and areas for growth). AEP provides collaborative computer-supported project-based opportunities that enable students to extend the application of course concepts to embedded contexts that engage them as learners. To compare AEP and active-learning control conditions, this study utilizes identical measures: validated surveys, interviews, observation notes, anonymous course-feedback forms, and formative and summative assessments. Multivariate regression analyses suggest that students learning through AEP outperformed control conditions by a full letter grade and scored on average 14.20 percentage-points higher on final assessments, when controlling for disability status, gender, and pre-test achievement. AEP has a large standardized effect size on the final ($d=2.40$) and provides a framework to improve student success through embedded and extended learning opportunities.

Adaptive equity-oriented pedagogy (AEP), research questions, and rationale

This study evaluates an adaptive equity-oriented pedagogy (AEP) that seeks to improve diverse students' success through computer-based formative assessment and computer-supported collaborative project-based learning. Applying AEP, instructors administer weekly computer-based student assessment and surveys to diagnose students' learning needs (i.e., their strengths, interests, and areas for growth). Using student data and analytics, instructors iteratively adjust how they provide feedback, model key skills needed for summative assessments, and implement deliberate practice activities. Students pursue their interests by applying course concepts to a novel context through collaborative project-based learning.

AEP seeks to engage learners in an embedded space by making the content and context relevant so that students can extend their learning into the real-world as a scientist. These embedded learning opportunities are beneficial since instructors address skills that are relevant to students' lived experiences and contexts outside of the classroom. To achieve this goal, students' learning is embedded in the real world since they can apply concepts to a range of contexts, which include researching longitudinal data on students' mental health, business management practices, and cancer treatments. Through project-based learning, students apply design-based research to collaboratively examine and address shared problems. Students work in teams to develop competing hypotheses based on existing literature, their experiences, and the concepts presented in class. They then gather data to test these hypotheses through statistical software as they collaborate online to collect data, build databases, generate code, produce visualizations, analyze data, and interpret findings. Students work together to extend course concepts in embedded contexts to make sense of real-world data. Collaboration in these contexts is critical, since students offer diverse perspectives on ways to refine research design and analyses. AEP facilitates this process because it identifies barriers to student collaboration. For example, instructors use formative assessment and surveys to diagnose students' learning needs with respect to collaboration.

Through these strategies, AEP enables instructors to overcome barriers that otherwise inhibit collaborative learning in university classrooms. One such barrier is the conflict that arises in group projects when members do not complete tasks on time. Reasons for incomplete tasks include lack of proper planning, miscommunication, students not knowing how to apply concepts to large tasks, and students working on tasks that do not engage their interests and strengths. The AEP model helps students develop a structure for collaborative work. For instance, students form teams based on common interests and are asked to collaboratively create a task chart via Google sheets or other mediums to assign roles during group projects. To create a system of accountability, students use this chart to designate individuals who work collaboratively on tasks that have deadlines. To foster trust and motivation, group members assign tasks and roles that leverage their members' strengths and interests. Groups also co-construct norms to navigate conflict and ensure accountability. In addition, instructors continuously model collaborative strategies and leverage student data to provide additional support that groups need. Although these approaches may be common in K-12, they are fairly uncommon in U.S. higher education (Phuong et al., 2017). By equipping students with the skills to collaboratively address problems using technology, AEP aims to increase students' sense of belonging in STEM, since they are developing skills that

society values and needs to thrive in a digital age. These experiences provide opportunities for students to practice academic discourse and to connect it to their interests and sociocultural identity (Gee, 1996).

To formally study whether AEP can have strong impacts in a university statistics course, we conducted a mixed-methods randomized controlled trial (RCT). We examine underrepresented minority (URM) student groups that are often overlooked in research on higher education pedagogy. Our research questions below examine an AEP treatment course's impact relative to an active-learning control course and on subgroups (i.e., disability status and gender), since these URM groups have experienced low achievement, persistence, and retention in STEM (McGregor et al., 2016; Stout & Wright, 2016). This study controls for students' pre-test achievement on the final assessment in treatment and control conditions, since prior achievement significantly influences subsequent achievement (Hattie, 2009). Based on these interests, we explore the following research questions:

1. Is there a mean difference on final collaborative assessments between students in the AEP treatment condition and students in the control condition, after controlling for disability status (i.e., students with and without disabilities), pre-test achievement, and gender (i.e., gender non-conforming, female, male)?
2. Is the impact of AEP on student assessment scores consistent across disability and gender groups?

Our research questions evaluating AEP are significant since universities experience difficulties recruiting and retaining STEM students, especially URM (U.S. Department of Education, 2017). Research shows that 90% of students leaving STEM cited poor teaching as a primary concern (Seymour & Hewitt, 1997). Specifically, university statistics courses, a foundational requirement for STEM fields, are typically computation-heavy and lecture-based (Allen et al., 2012). With limited opportunities for real-world application, students can struggle seeing the relevance of learning statistics (Allen et al., 2012). Furthermore, URM students (e.g., students with disabilities, gender non-conforming students) with limited mathematical preparation often experience math-phobia and poor academic achievement (Tishkovskaya & Lancaster, 2012) which can affect their confidence, scientific identity, persistence, and retention (Peters, 2014; Stout & Wright, 2016). Since teaching can directly impact student success (Condon et al., 2016), studying pedagogy to improve students' academic achievement, especially for URM, in STEM is critical.

Potential significance

Increasing recruitment and retention in STEM, especially for URM, is important because this population of students can offer diverse perspectives on problems and solutions. Increasing diversity can enable future leaders to be more responsive to their constituents and environment. Additionally, many industries in today's workforce require competency with computer-based technologies and collaboration in team settings. However, many higher education STEM courses are not adequately equipping students, especially URM, with technology-based competencies and interpersonal, collaborative skills (Gasiewski et al., 2012). This is important since students in a scientific community need to learn to engage with others, recognizing community guidelines and norms. Many of these norms are often not explicit to students, especially for URM and first-generation students who did not have access to academic-oriented scientific communities. AEP seeks to make these guidelines and norms explicit by modeling these norms and providing students with opportunities to collaboratively practice and build on these norms with a community of learners. To draw students into a scientific community, the AEP model assumes that like scientists, students often have an interest and purpose to pursue a research topic. Based on this premise, the AEP model seeks to understand learners' sociocultural experiences and socio-academic interests. Through surveys and formative assessments, AEP enables instructors to understand their students' sociocultural environment and context—their lived experiences and interests that inspire them to pursue STEM. For example, some students are drawn to medicine because their parents suffered from medical conditions and could not address their illnesses due to financial constraints. Hence, AEP empowers instructors to identify what brings students to academic STEM spaces by providing frameworks to understand learners and their sociocultural histories.

As students bring themselves into their projects, we study how students offer diverse strengths and perspectives to practice “thinking scientifically” as they design research questions, make and test predictions, solve problems, and critique their own and others' reasonings (Deslauriers et al., 2011). Students also collaborate synchronously and asynchronously on their computers through Google docs where they work together to share visualizations and to analyze and interpret data. By cultivating environments that mirror a professional scientific community, students have opportunities to collaboratively practice and engage with scientific discourse, where they can embody communicating, behaving, thinking critically, and solving problems like a scientist (Gee, 1996).

Our research represents a significant contribution. First, it is useful for practitioners since it provides a research-based pedagogical framework and tools that instructors can use to increase equity in student outcomes, especially for URM. Second, this study's methodological approach advances higher education pedagogy scholarship practices; this is the first study to our knowledge that uses a randomized controlled trial (RCT) with

the same university instructors where the control condition employs active learning in a university STEM course. In many higher education STEM studies, the control condition differs in multiple respects than the treatment making the source of different outcomes difficult to isolate. For instance, the control is often lecture-based and employs instructors with different characteristics than those in the treatment. In this study, the treatment and control conditions are more precisely contrasted allowing us to focus more precisely on AEP effects of interest.

Theoretical approaches

Conceptual framework of adaptive equity-oriented pedagogy (AEP)

In this section, we describe AEP's key elements and the learning theories that underpin them. Drawing on McCallum's (2013) Assessment-Instruction (A-I) conceptual framework, AEP focuses on collecting ongoing diagnostic assessment data to guide classroom instruction, decisions, and lesson planning. These assessments gather information on students' backgrounds, interests, and experiences to understand the cognitive and noncognitive factors that impact student engagement in light of their sociocultural history. Applying AEP, instructors use weekly data on students' learning needs to continuously adjust how they

1. foreground how course concepts are relevant to students' goals and shared contexts
2. provide brief warm-up active learning exercises where students can engage in productive struggle with concepts and make meaning individually or collaboratively
3. model software skills, expert thinking on concepts, and strategies that students need to excel on summative assessments; during this time, instructors build on and respond to students' critical thinking strategies from previous assessments and/or productive struggle experiences
4. provide written steps and strategies (e.g., task analysis) of how to approach problems that align with rubrics and the rigor of summative assessments
5. include class time for students to practice these skills and strategies to collaboratively analyze and interpret data on computers; during these deliberate practice activities, students articulate to each other which concept-driven step-by-step strategies are useful for their team final project and why
6. provide low-stakes feedback during class to address misconceptions and close gaps in understanding
7. offer in-class and project-based opportunities for students to incorporate feedback and reinforce learning

AEP seeks to support URM students' success by building on McCallum's (2013) A-I model and Vygotsky's (1978) sociocultural theory of learning, in which learning is situated as a cultural and social process. To avoid the lack of social, collaborative interaction associated with more traditional lecture-based pedagogies, AEP draws on formative assessment and universal design for learning (UDL) to address a wide range of learning needs. Throughout this process, AEP provides an adaptive active-learning framework that helps instructors use formative assessment to iteratively adjust their UDL and collaborative learning practices.

Formative assessment

AEP uses formative assessment (Black & Wiliam, 1998) by engaging Vygotsky's (1978) sociocultural theory and the zone of proximal development (ZPD), where an instructor assesses a student's actual developmental level (i.e., a student's level without assistance) to help them reach their potential developmental level (i.e., a student's level with assistance). Using formative assessment data, the instructor modifies evidence-based practices (e.g., modeling key skills, collaborative deliberate practice opportunities, feedback) to help students reach that potential. Instructors assess and provide feedback to all students via formative assessment instead of assuming that students enter the class with the appropriate resources and skills to meet course goals. Furthermore, formative assessment aligned with the final exam's rigor clarifies expectations for students unfamiliar with the dominant cultural capital and academic discourse assumed in many college classrooms (Bourdieu, 1973; Phuong et al., 2017). Formative assessment promotes equity in student outcomes since struggling students are more likely to persist in a field when they develop stronger conceptual foundations, reflect on their growth, have a collaborative community, and are confident in their work (Ambrose et al., 2010; Phuong et al., 2017).

AEP's in-class formative assessment, modeling, collaborative practice, and feedback loops promote equity in student outcomes since they support novice learners and underrepresented students who may struggle with challenging concepts. According to Schwartz et al. (2015), novices "find it harder to engage in deliberate practice on their own, because novices often do not have the experience to know what skills they should be working on or how to go about it" (p. 295). Incorrect types of practice can reinforce misconceptions. AEP's in-class formative assessment, modeling, practice, and feedback loops allow instructors to review challenging concepts. AEP also leverages collaborative learning via peer-to-peer support since peers who more recently

learned concepts often empathize with barriers to understanding concepts. Often, these peers can break down course material in digestible ways for struggling and advanced students. This form of reciprocal teaching enables students to work with their peers and draw on each other's expertise to move farther together in their ZPD.

In AEP, modeling key skills and strategies provide opportunities for instructors to offer various worked examples that show students "what to do and why" when they approach problems (Schwartz et al., 2015, p. 295). Instructors applying AEP provide step-by-step explanations for solving problems in writing so that students can understand different ways to read prompts, identify underlying concepts in the problem, explain why they are doing each step, articulate how each step connects to key concepts, and understand when these steps can and cannot be generalized to other situations. These step-by-step strategies further reduce cognitive load because the steps help chunk students' mental processes and deliberate practice opportunities (Sweller, 1994). This approach can make learning new concepts less daunting for novice learners. These strategies and collaborative practice opportunities can help improve students' success and self-efficacy with mastering challenging concepts, especially for URM students who had less access to college-level coursework. AEP seeks to improve students' conceptual understandings and self-efficacy so they can actively contribute to their project teams and their peer's learning during computer-supported, collaborative activities.

These formative assessment, collaborative practice, and feedback loops can also clarify expectations for success and foster students' metacognition and self-regulation processes (Nicol & Macfarlane-Dick, 2007). For example, students reflect on and monitor whether they understand concepts when they take formative assessments and practice scientific thinking. This process can help students identify their strengths, interests, and areas for growth, which can help them set academic goals and develop strategies to advocate for their learning. For URM students who feel underprepared for college STEM courses, providing weekly ungraded assessments and supportive collaborative activities creates a psychologically safe space. Creating a psychologically safe space is crucial because many URM students have experienced high incidences of microaggressions and low levels of STEM achievement, sense of belonging, persistence, and retention rates (Stout & Wright, 2016). Offering greater opportunities for URM students is necessary because many URM families often do not have equitable access to extracurriculars, resources, and rigorous curricula compared to more privileged families (Martin et al., 2016).

Universal design for learning

To support diverse learners, AEP also incorporates UDL to optimize learning by engaging affective, recognition, and strategic brain networks (Rose et al., 2002). UDL enables students to learn, demonstrate, and reinforce knowledge in various ways by providing multiple means of engagement, representation, and action and expression (Rose et al., 2002). For example, AEP instructors synthesize lectures, simulations, technology, computer-based activities, and project-based learning. Building on Lee's (2005) culturally-responsive-teaching framework, instructors apply UDL strategies to draw on students' funds of knowledge—their background, interests, aspirations (Moll et al., 1992)—to increase their sense of belonging and engagement with developing academic skills. To increase student engagement with academic content, instructors: 1) use student survey data to explicitly articulate how students' interests and aspirations are relevant to course material during interactive lectures and active learning activities, 2) visualize challenging concepts with written annotations, and 3) ask students to apply course content to their own contexts. Consequently, students can see how their backgrounds, interests, and identities are relevant to and are strengths in STEM. In addition, students can exercise agency as they apply concepts to collaboratively address social challenges that relate to their lives and communities using technology.

In sum, AEP's formative assessment and survey strategies provide instructors with data to understand students' interests, barriers to learning, and any microaggressions they face in STEM. AEP addresses cultural and social misalignments by equipping instructors with a framework to clarify expectations for success, create a space that validates students' backgrounds, and provide opportunities for students to code switch and connect their interests with academic discourse. AEP supports struggling students since it empowers learners to make meaning through productive struggle and assists students who are still acquiring the skills to complete conceptual tasks. Seeking to avoid assimilationist approaches to equity, AEP's end goal is not to ask students to think like the instructor and to solely learn a form of routine expertise via didactic modes of teaching (Schwartz et al., 2015). Instead, AEP encourages students to develop a greater sense of adaptive expertise, where students examine problems and apply relevant concepts with and without direct instruction (as described above). Throughout this process, students practice and synthesize different ways of scientific thinking gained from their community, instructors, peer-to-peer collaboration, and their own insights.

AEP is responsive to ongoing learning and is culturally responsive because it does not perceive students' pre-test scores and backgrounds as deficits dictating an endpoint. Instead, AEP sees students' backgrounds as sources for innovation, and AEP therefore empowers students to leverage their existing social and cultural competencies to apply scientific concepts to both novel problems and contexts of their choice. For example, AEP's

instructional and project-based approach enables students to co-construct and make meaning of STEM concepts to utilize innovative strategies that build on their peers' and instructors' critical thought process. Drawing on multiple frameworks, AEP focuses on equipping students with developing the mindset, skills, sense of community, and confidence to tackle meaningful problems where there exists no single solution.

Methods

127 undergraduate student participants from a R-1 US university took the same course, but were randomly assigned into treatment or control sections. The course required students to complete a pre-test on foundational course content and a final assessment. These assessments and the following identical measures were collected in both conditions: validated surveys, demographic information, interviews, observation notes, anonymous course-feedback forms, and formative and summative assessments. Four observers took notes during every class in both conditions to determine if 1) the treatment course applied AEP and 2) the control course did not adjust instruction based on data. The treatment and the control conditions had the same instructors who employed exemplary active-learning strategies (e.g., modeling key skills, deliberate practice, dynamic lecturing, dialogue, case studies). However, instructors adjusted their teaching practices for students in the treatment group based on weekly student data. For the control condition, these same instructors did not adjust their teaching based on weekly student data, since they did not see these data. To ensure that the treatment did not impact pedagogical practice in the control condition, the control condition was taught before the treatment condition in each week of instruction. Instructors were provided with data from the treatment condition only after they had taught the control condition. In the control condition, students received scores on assessments but instructors did not.

We conducted multiple regression at the 5% significance level. We fitted one regression model for research question 1: the final assessment scores (scale 0-100) were regressed on pre-test (scale 0-100), dichotomous variables for treatment group and disability status, and a categorical variable for gender (female and male with gender non-conforming as the reference category). For question 1, the coefficient of treatment is an estimate of the difference in mean final achievement between treatment and control students after controlling for disability status, gender, and pre-test. For question 2, we looked for an interaction between treatment group and disability status within the regression model. The coefficient for this interaction term represents the mean difference in the treatment effect between disabled and non-disabled students, after controlling for pre-test and gender. We also looked for an interaction between treatment group and gender within the regression model. This produced two interaction terms. The coefficient for one interaction term represents the mean difference in the treatment effect between gender non-conforming and female gender groups, controlling for pre-test and disability status. The coefficient for the other interaction term represents the mean difference in the treatment effect between gender non-conforming and male gender groups, controlling for pre-test and disability status.

We also used Saldaña's (2009) methods of qualitative coding, categorizing, and identifying patterns. We inductively coded data by identifying themes and assigning thematic codes. We then created overarching thematic codes based on patterns, which became our analytical focus. To establish inter-rater reliability, we created a codebook with definitions of engagement and achievement with examples to guide our analyses. We triangulated data sources and multiple viewpoints (e.g., instructors, students, and researchers) to further analyze data.

Findings

Descriptive statistics

The treatment and control groups have similar frequencies on disability status and gender explanatory variables (See table 1). The control's mean pre-test score ($M=29.42$; $SD=12.84$) is comparable to the treatment's ($M=29.70$; $SD=12.24$). Additionally, the treatment's mean final score ($M= 97.82$; $SD=1.75$) is higher than the control's ($M=83.66$; $SD=8.10$).

Table 1: Contingency table of disability status and gender categories of control and treatment groups

	Control (n=65)	Treatment (n=62)
Demographic Variable	N (%)	N (%)
Disability Status		
Non-Disabled Students	42 (64.62%)	41 (66.13%)
Disabled Students	23 (35.38 %)	21 (33.87%)

Gender

Gender Non-Conforming	16 (24.62%)	15 (24.19%)
Female	26 (40.00%)	25 (40.32%)
Male	23 (35.38%)	22 (35.48%)

Regression analyses

We conducted a regression analysis with robust standard errors to correct for heteroskedasticity, in order to examine the treatment effect on collaborative final project outcomes, controlling for disability status, pre-test achievement, and gender. See table 2 for regression results. We performed appropriate regression diagnostics to validate the use of this approach. There was no evidence of collinearity between predictors (mean VIF =1.24).

Table 2: Multiple regression with robust standard errors for the effect of treatment group, pre-test, disability, and gender

Variable	Est. Coeff. (Robust Std. Err)	95% Confidence Int.l		p-value
Treatment	14.20 (.99)	12.23	16.17	<0.001
Pre-test	0.03 (0.04)	-0.06	0.12	0.50
Disability	2.65 (1.16)	0.35	4.94	0.02
Gender (Reference: Non-Conforming)				
Female	-2.73 (1.40)	-5.50	0.03	0.05
Male	-2.28 (1.34)	-4.93	0.37	0.09
Intercept	86.38 (2.58)	81.28	91.48	<0.001

Research question 1. The treatment group on average is estimated to perform 14.20 percentage points higher (SE = 0.99, 95% confidence interval from 12.23 to 16.17) on the final than the control, after controlling for disability status, gender, and pre-test achievement. This difference was statistically significant ($t = 14.27$, $df = 121$, $p < 0.001$). This model's R^2 is 0.63, suggesting that 63% of the variation in the final assessment scores can be explained by treatment group, disability, gender, and pre-test.

Research question 2. The effect of AEP persists across disability status and gender groups. No significant interaction exists between treatment group and disability status at the 5% level. Additionally, no significant interaction exists among treatment group and gender at the 5% level. Therefore, students in each of these groups appear to benefit equally from AEP.

Qualitative findings

At the beginning of the semester, survey data showed that treatment and control students had anxiety about collaborative learning. Many of them were unconfident in their skills and they often felt uncomfortable interacting with peers because they did not know each other or their peers' strengths. Students also felt that project requirements were overwhelming since they had multiple parts and asked for deep conceptual application. One student stated, "Working with others on a project that has no clear answer feels much more difficult than a multiple-choice exam." Students also added that using statistical software felt daunting and they had fears about their team not completing project tasks. To alleviate student anxiety, instructors applied AEP strategies to address interpersonal and conceptual shortcomings that students voiced on formative assessments and surveys. In surveys, treatment students indicated that it was helpful when instructors provided team activities and peer-to-peer validation exercises to address these shortcomings and foster a sense of community.

Coding of survey data suggests that AEP provided a structure that helped improve students' collaborative learning skills and experiences. Treatment students indicated that they learned effectively from applying statistical competencies via peer-to-peer activities to succeed in a collaborative scientific community. These students noted that group members contributed effectively to the team since peers and instructors ensured that everyone understood how to apply course concepts and use software. These treatment students highlighted how instructors

supported groups by providing additional examples of how other organizations dealt with interpersonal challenges. Moreover, treatment students appreciated completing point-people, task, and deadline sheets, because they could plan and break down projects into manageable tasks that could be completed each week within an hour. These tasks had deadlines with enough time for the team to provide feedback to each other and to allow extensions that did not significantly slow down the team. A majority of students also found it less stressful when they could submit project pieces in small chunks to a folder or other medium (e.g., Canvas page, Google drive, project Dropbox folder), because they could earn points for these smaller and intermediate submissions. Treatment students also highlighted how instructors used assessment data to provide resources and page numbers from readings to address group project needs. Moreover, the teaching assistants monitored who was completing tasks on time. They provided suggestions to group members on how they could ask questions to understand and address the barriers that team members had with completing tasks on time. The teaching assistants regularly encouraged team members to play to their strengths and to offer each other collaborative support with completing tasks. Students also mentioned that ungraded individual reflections mitigated conflict and contributed to more productive collaboration, since students self-acknowledged and addressed their areas for growth. This study shows that formative assessment data in higher education is not only useful for addressing students' conceptual difficulties, but is also beneficial for addressing barriers to collaborative learning.

Based on surveys and interviews, treatment students reported excelling on the collaborative final project because 1) they received weekly feedback from peers and instructors and 2) instructors adjusted teaching in response to student performance. When asked to indicate effectiveness of various teaching practices, treatment and control students appreciated how steps for challenging problems were broken down, linked to definitions of concepts, provided in writing, and modeled for them with reference to a rubric. Treatment students reported being more prepared for collaborative in-class activities when instructors provided class-discussion questions and lesson-plan outlines before class. Treatment students also consistently emphasized how instructors addressed misconceptions from weekly assessments, reviewed challenging concepts, and validated and built on the different ways students arrived to accurate solutions. Additionally, these students reported improved learning when instructors offered written cues and strategies to help students analyze question prompts. These students said this approach helped them apply different skills and thought processes when completing assessments. Moreover, treatment students stated that the collaborative learning process was helpful because they were held accountable to a peer and had to explain and answer questions about the rationale behind their thought processes. Treatment students appreciated using various concept-driven step-by-step strategies to identify their problem and goal, select the strategies to address their goal, and then develop a rubric for assessing their work and their peers' work. These students mentioned that this process helped them engage in purposeful application and avoid mindless application of code and strategies when analyzing data.

By contrast, data from the control condition show consequences when instructors were not aware of barriers to student collaboration. Many control students mentioned experiencing difficulties with collaborative assessments since they needed more review of dense concepts and software skills to contribute effectively to their teams. Moreover, several control students indicated that they were not satisfied with group projects because collaboration was strained. For example, some control students felt that they could not voice their concerns to their instructors, fearing repercussions for their teammates. Many control students indicated that they decided to work on tasks individually to avoid conflict. Fortunately, a majority of students in both conditions found synchronous and asynchronous computer collaboration to be helpful for sharing code, editing visualizations, and analyzing data together, especially when they used Google doc functions. For example, these students used collaborative editing features and clicked a user's icon in the Google doc, which would bring them to the location within the document where their counterpart was working. Students in both conditions reported learning effectively when they collaboratively worked on clicker-esque exercises because they could explain concepts to each other and question each other's thinking.

Implications and conclusion

AEP's large standardized effect size on the final ($d=2.40$) is consistent with previous studies that examined the impacts of formative assessment and active learning (Deslauriers et al., 2011; Froyd, 2008; Phuong et al., 2017). AEP provides a framework on how to leverage the benefits of extended and embedded collaborative learning. Students in the treatment reported how learning to diagnose and break down problems from instructors and peers helped them better understand concepts since they could see the application of these concepts in collaborative tasks and via computer-based visualizations. Highlighting the benefits of embedded learning, treatment students emphasized that the material felt more relevant when they were asked to collaboratively discuss different step-by-step modeled strategies that were and were not useful for the context of their team project and why. Alluding to the advantages of extended learning, treatment students also said the concepts and statistical tests made more

sense and were worth learning when they were applicable to contexts and data that were meaningful. These findings suggest that AEP can help instructors address their students' learning needs (i.e., their interests, strengths, and areas for growth) to improve computer-supported collaborative learning experiences in real-world contexts.

This research is novel since no study to our knowledge has used an RCT with the same university instructors to compare an equity-oriented framework with a control condition that employed exemplary active-learning strategies. This control condition avoided the pedagogical limitations cited in multiple studies across different higher education contexts for introductory STEM courses. Pervasive across several colleges, including the university in this study, these limitations include curved grading, competitive and un-collaborative peer environments, overpacked lecture-based curricula, limited faculty interaction, and artificially difficult exams that are disconnected from both classroom instruction and the real world (Bettinger 2010; Barr, Gonzalez, & Wanat 2008; Crisp, Nora, & Taggart 2009; Eagan et al. 2011; Seymour & Hewitt 1997). At the R-1 institution we studied, statistics instructors in introductory courses and electives primarily teach through lecture and typically do not adjust teaching based on weekly formative assessments aligned with the final's rigor. They often have problem sets and classroom exercises that are lower in rigor than the final and midterms. For the active learning control condition in this study, instructors modeled key skills and provided deliberate practice and active learning activities to students. However, like many higher education STEM faculty who have not been formally trained in teaching, the control instructors did not adjust instruction based on weekly assessments. Nevertheless, control students, on average, performed about a third to half a letter grade higher than department norms for introductory statistics courses. By comparison, the treatment group, on average, scored over a full letter grade higher than these department norms. Therefore, there are significant implications for AEP in introductory college statistics and data science courses that are foundational to STEM disciplines. These courses can be gateway courses, which impact students' attitudes, confidence, sense of belonging, persistence, retention, and identity in STEM. Applying AEP can benefit universities and STEM-related departments that want to create collaborative spaces that retain URM.

Limitations to this study include institutional context and that the disability variable combines physical, invisible, and mental health disabilities. Another possible limitation is that the control was taught before the treatment for each week of instruction. However, in previous studies, the standardized effect size ($d=2.36$) was similar when the AEP treatment condition was taught before the active learning control condition (Phuong et al. 2017). Future research directions include controlling for additional variables (e.g., race, class) and finding ways to support faculty with implementing AEP to improve student learning. It would be particularly interesting to examine whether students in the treatment group would perform better in subsequent courses. We would also like to examine if the treatment would impact other academic (e.g., persistence, retention) or psychosocial outcomes (e.g., sense of self-efficacy, sense of belonging, stereotype threat) in STEM courses.

Faculty development centers and equity units can provide programs, guides, and incentives that support instructors with applying elements of AEP. AEP's elements (e.g., formative assessment) can be useful across disciplines, academic departments, universities, and schools committed to narrowing academic achievement gaps, especially for URM. By increasing URM students' success in higher education, the AEP model strives to diversify perspectives on STEM and respond to systemic inequities, such as retention issues. These students' success, which can be enhanced through AEP, can inspire others in their communities to pursue STEM-related majors and careers.

References

- Allen, R. A., Folkhard, A., Abram, B., & Lancaster, G. A. (2010). Statistics for the biological and environmental sciences: improving service teaching for postgraduates. *Journal of Statistical Education*.
- Deslauriers, L., Schelew, E., & Wieman, C. (2011). Improved learning in a large-enrollment physics class. *Science*, 332(6031), 862-864.
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 111(23), 8410-8415.
- Hattie, J. (2009). *Visible learning: A synthesis of meta-analyses in education*. Routledge.
- McGregor, K. K., Langenfeld, N., Van Horne, S., Oleson, J., Anson, M., & Jacobson, W. (2016). The University Experiences of Students with Learning Disabilities. *Learning Disabilities Research and Practice*, 31(2), 90-102.
- Phuong, A. E., Nguyen, J. & Marie, D. (2017). Evaluating an adaptive equity-oriented pedagogy: A study of its impacts in higher education. *The Journal of Effective Teaching*. 17(2), 5-44.
- Schwartz, D. L., Tsang, J. M., & Blair, K. P. (2016). *The ABCs of how we learn: 26 scientifically proven approaches, how they work, and when to use them*. WW Norton & Company.
- Seymour, E., & Hewitt, N. (1997). *Talking About Leaving: Why Undergraduates Leave the Sciences*. Westview Press.