Increasing College STEM Instructors' Equity-Oriented Teaching Competencies and Students' Success

By

Andrew Estrada Phuong

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Education

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Bruce Fuller, Co-chair
Associate Professor Michelle Hoda Wilkerson, Co-chair
Professor Sophia Rabe-Hesketh
Teaching Professor Daniel D. Garcia

Fall 2022

ABSTRACT

Increasing College STEM Instructors' Equity-Oriented Teaching Competencies and Students' Success

by

Andrew Estrada Phuong

Doctor of Philosophy in Education

University of California, Berkeley

Professor Bruce Fuller and Associate Professor Michelle Hoda Wilkerson, Co-chairs

My dissertation research explores the factors that drive college STEM instructors' adoption of equitable teaching practices, such as Adaptive Equity-Oriented Pedagogy (AEP). AEP is a framework for adjusting teaching to address equity barriers to learning based on student data collected through formative assessment, observations, and surveys (Phuong et al., 2017a). Through randomized controlled trials, I had previously found that compared to an active learning control course, instructors applying AEP improved average student achievement by over a letter grade and narrowed achievement gaps for all students (Phuong & Nguyen, 2019; Phuong et al., 2022).

In my dissertation research, 129 student-instructors (henceforth "instructors") were randomly assigned to treatment and control pedagogy courses. While the control course *taught* instructors *about* AEP, the treatment course *modeled* AEP explicitly by using weekly instructor reflection data to continuously adjust course discussion and activities. Using multilevel regression modeling, I found that the treatment, on average, significantly improved instructors' AEP competencies over time compared to the control (2.63 standard deviations, $p<0.001$), when controlling for instructors' gender, URM status, years of teaching experience, years of tutoring experience, and number of students. Compared to the control condition, more instructors in the treatment condition used multiple data sources to adjust teaching and address equity barriers to advance student learning. This research demonstrates that modeling AEP principles explicitly (e.g., reviewing and addressing instructors' beliefs and lenses based on written reflections) in educational development programs can improve instructors' equitable teaching competencies.

Through the novel application of mixed-methods research and validated AEP measures, this study also offers a comparative analysis of a subset of instructors who exhibited low- versus high-growth on AEP competency. My research identifies the rationales instructors provided in their written statements for enacting particular teaching practices. These rationales were analyzed for evidence of a variety of factors driving instructors' learning, such as their use of student data, conceptions of equity, emotions, and sharing the responsibility for learning. Compared to low-growth instructors, I found that high-growth instructors more frequently shared the responsibility for learning with their students, and used multiple data sources (e.g., formative assessment,

observations, surveys) to understand equity barriers (e.g., imposter phenomenon, stereotype threat) and to adopt teaching practices that address these barriers.

My dissertation validates an Adaptive Equity-Oriented Pedagogical Competency (AEPC) assessment, a measure of college instructors' effectiveness with inclusive teaching, that was used to track instructor growth in equitable teaching competencies in my study. The AEPC assessment includes reflection questions that ask instructors to provide evidence of demonstrating AEP competencies, such as clarifying learning outcomes; aligning formative assessments and activities with outcomes; understanding student equity barriers; and continuously adapting teaching to address these barriers. More specifically, the AEPC assessment applies a rubric to score self-reported data on instructors' teaching philosophies, learning outcomes, written reflections on teaching and assessment practices, and instructor-provided evidence of inclusive pedagogies and student learning (e.g., student gain score visualizations, peer/student feedback on teaching). The AEPC assessment is thus built on a portfolio framework commonly used to evaluate college teaching (Seldin et al., 2010) but extended to incorporate well-defined criteria to measure equitable practices.

Based on partial credit Rasch modeling (Masters, 2010), the AEPC assessment has high reliability (0.94), has strong validity based on Wilson's (2005) strands of validity framework, shows no gender assessment bias, has high inter-rater reliability, and correlates with researchers' independent review of instructors' teaching methods and pedagogical materials (Phuong et al., 2022). My research demonstrates that instructors can use the AEPC assessment to demonstrate competency and progress in applying equitable teaching practices. This assessment approach enables instructors to document evidence of inclusive teaching practices, which provides insight on how they might plan, teach, reflect, and alter their pedagogical practice to address students' equity barriers. In addition, the AEPC assessment does not exhibit the known gender and racial bias found in course evaluations (Boring et al., 2016; MacNell et al., 2015; Phuong et al., 2022; Stark & Freishtat, 2014). My work contributes to conversations on how studies of college instructor professional development programs focused on equity can leverage randomized controlled trials, partial credit Rasch modeling, and multilevel modeling to reduce bias, mitigate the impact of confounding variables, create comparable groups, and support causal inference. This study provides recommendations for creating adaptive professional development programs that can advance equity in student outcomes.

Keywords: *Higher Education STEM Pedagogy, Equitable Teaching Competencies, Instructor Professional Development, Adaptive Equity-Oriented Pedagogy, Equitable Teaching, Assessment of Equitable Teaching, Reducing Assessment Bias*

ACKNOWLEDGMENTS

Image credit: mykidsdream.com; used by permission.

Growing up, I loved dinosaurs, and my favorite was the Tyrannosaurus Rex, also called the T-Rex. Some might think that T-Rexes were formidable, solitary hunters in the Late Cretaceous period. Recent evidence, however, suggests that T-Rexes hunted socially in a pack (Titus et al., 2021). Indeed, even millions of years ago, community and collaboration were key ingredients in the recipe for survival.

As a T-rex committed to peace and equity, I was able to conquer the rocky terrain of a doctoral dissertation through the support of my extraordinary and awe-inspiring pack of dissertation committee members: Professors Michelle Wilkerson, Bruce Fuller, Sophia Rabe-Hesketh, and Dan Garcia. Michelle, thank you for your dedication in ensuring that my dissertation work would not only be robust enough to withstand an asteroid, but also earth-shattering enough to be the comet that ushers in a new era of equity research in higher education. There is no one like you and I thank you for emboldening my voice in the learning sciences. You truly mentor the whole person and helped me feel like I can belong in the learning sciences. You're a true equity champion and you represent the best wave of mentorship that will transform academia into a more inclusive space for learning. Bruce, thank you for embracing my authentic self when I literally dressed up as a dinosaur to give presentations in your classes. You have given me valuable insights into organizational change and educational policy that influenced much of my work. Thank you for being an amazing primary advisor and for supporting me as a whole human being. You are a force of nature who has truly cared about me. You were a perfect advisor for me during my doctoral career, you gave me the agency and support I needed to thrive. Sophia, thank you for your guidance on quantitative methods throughout my projects. Being equipped with prodigious research skills, you helped me become a T-rex that could take on multilevel modeling and advanced quantitative research. You're a true multilevel modeling queen and learning from you has been invaluable. Your quantitative wisdom will continue to live on in me and was perfect for my dissertation. Dan, thank you for helping to ground my work and bringing in a practical focus for my dissertation. Your practical perspectives energize me, and I am inspired by your passion for equity! You are a true trend setter and equity champion whose work will continue to influence the millennia to come. Your incredible insight into Computer Science and Data Science education was perfect for me as I navigated my dissertation research. Your support will help my work impact the field – for T-rexes, velociprators, mosasauruses, pterodactylls, bracchiosauruses, plesiosauruses, and so forth.

Beyond the pack of my dissertation committee, I would also like to acknowledge the generous support and time of my mentors. Echoing the words of Kris Valloton, my committee and mentors have poured water on my fears and have thrown gasoline on the passions that have ignited my dreams! In this spirit, each and every one of you is the best!

I will start by thanking Richard J. Patz. Rich, thank you for your generous time supporting me professionally throughout my career. You are the epitome of what it means to be the ultimate and true equity-oriented mentor, patiently answering my questions and working in partnership with me to advance my learning. You truly create a home away from home and have supported me in taking on some of the most exciting challenges and opportunities. Thank you for supporting my success both inside and outside of the classroom. No statistical model can account for the profound impact that you have had on this world and on my life. You have truly exponentialized

my passion, potential, and impact on advancing equity, inclusion, justice, and belonging. Thank you for being the best, especially at what you do– whether it is measurement or building a Christmas Tree farm. I'll see you soon at your Christmas Tree & Pumpkin Farm, Rancho Siempre Verde (RSV Trees), where the most epic swing exists. You have the wisdom and mind of an entire galaxy and a heart of gold! Your brilliant, energetic, patient, and kind spirit will always live in my heart and in all the work that I do. Meeting you was truly a miracle and was life changing! You have made a difference in my life that will last for eternity! Through your words and actions, you have taught me that one person can transform the world for so many people around them! Our conversations truly bring me to life and empower me to reimagine and advance the greatest future possible! You're a god-tier superhero!

Doug Holton, thank you for the gracious hours chatting with me about our shared passion for faculty development and higher education literature. I am constantly amazed by the depths of your knowledge of work in the education field, which has helped my work be in conversation with the latest and greatest in educational equity research. Talking with you is like talking to Google Scholar; you are so well-versed in the field of educational development and the learning sciences.

I would like to thank the following mentors who have supported my learning during my graduate and undergraduate studies.

Fabrizio Mejia, thank you for your partnership in working to improve student success both in theory and practice. Fabrizio, you're one of the main reasons why I went and made it through graduate school. You have helped me see everything is possible and you have focused on helping me find a path that will support me effectively in the future. You're a true brother to me and someone who helped me bring my authentic self to the settings I've been in. You're also the best at what you do and my true superhero!

Richard Freishtat, thank you for your mentorship and guidance which connected me to a network of professionals in faculty development. You have helped bring the best out of me and have shaped so many of my career and life goals. Your impact on my life is truly invaluable and your spirit continues to shape me.

Marcos Ramos, thank you for being an exemplar of a collaborative, equity-oriented advisor who creates a brave space for growth. Working with and learning from you has been transformative. You bring the feeling of being at home to our interactions and you're one of the main reasons why I decided to pursue graduate school. Your unwavering belief in me and our conversations have inspired me to advance equity work that extends beyond the classroom walls. You are the EOP college counselor who taught me what it means to create a home away from home. Most importantly, you embodied that, and I always felt that you truly understood me since we first talked. As Fab said, you are the model for CE3 and equity-oriented practices!

Richard McCallum, thank you for grounding me in the assessment-to-instruction model. Your work has been foundational in my thinking and has been essential to the development of my AEP framework. You saw the potential in me and nurtured me into the scholar I am today. The AEP framework, the ways I do research, and how I think about teaching and learning bear your

imprint. Your words of "Don't guess, assess!" have been life changing as I think about formative assessment and using multiple forms of data rather than my assumptions to guide my practice. And by talking with you, I learned that "We need to reject the pernicious belief that the only good students are the ones who teach themselves and don't need help." You're right that it truly takes a village to build many scholars. More importantly, you reinforced who I wanted to be—an educator who places the learner at the center of a complex learning process. When I was an undergraduate, you helped me normalize that it was okay for me to grade for equity, provide flexibility with deadlines, prioritize students' wellbeing, and focus on helping students bring their hearts, minds, passion, and authentic selves to the classroom. Your mentorship shifted my trajectory as an undergraduate and has inspired me to build a rigorous research agenda in higher education that advances practical strategies for increasing equity in student outcomes. Your impact has truly been life changing!

Glynda Hull, thank you for your continuous feedback and support on my work since my undergraduate career. Improving the AEP framework with your mentorship has been invaluable. Thank you for introducing me to the world of faculty and educational development and for bringing Richard Freishtat to campus. You are truly the quintessential critical cosmopolitan and thank you for helping me learn how not only to read the word, but the world (as Paolo Freire would say!).

Linda von Hoene, thank you for connecting me with the Professional and Organizational Development Network in Higher Education. Your time and support have been invaluable in my development as a scholar. There could be no better Director of the GSI Teaching and Resource Center than you! Ever since I met you as an undergraduate, I felt that you have understood me. You have supported me, nurtured me, and helped me become the educational developer that I have dreamed of becoming. You have designed the greatest graduate professional development resources and programs on campus that I have seen. I thank you for allowing me to attend your workshops when I was an undergraduate. You trained me in cutting-edge pedagogies and helped me see how the AEP framework truly builds on the field of educational development. You have truly been my Jedi master and your spirit and force will live through my work and everything I do. You're the best at exemplifying the highest quality mentorship for all on campus. I love your energy and you truly apply the principles from Claire Kramsch's *Language and Power* seminar to advance the most exciting and groundbreaking workshops, scholarship, and practical strategies. There is no mentor like you!

Stephanie LaBarge, thank you for being a true champion for students. Your time, support, and deep listening skills have transformed my educational trajectory. I appreciate you always supporting me and helping me navigate campus resources. You are the greatest DSP specialist I have ever seen, and you are one of the main reasons why I made it to the finish line for graduation. You have truly changed and transformed my life trajectory!

Martha Velasquez, you are my accessibility tech superhero! I have learned so much from you, including how to use accessible technologies and how to apply UDL principles to education, my research, and life! You have been truly life changing for me as well!

Raymond Budelman, thank you for being an amazing friend and collaborator. Your life advice and our conversations about critical theory and disability studies are so important to me. You're my superhero!

Paolo Martin, thank you for being my Cal Reads, critical theory, and Let's-go-Learn friend. I enjoy thinking with you on how to infuse critical theory perspectives into learning sciences and the work that I do. I try to emulate the patient way you work with learners and the passion you bring to our world. You have an amazing heart, and I am truly inspired by you!

Isela Pena-Rager, thank you for being a star advisor, coach, and so much more! You are like family to me, and I am so happy that I had the opportunity to learn from and collaborate with you to launch radical and transformative work! You are a true advocate and superhero! The love, passion, and care that you bring to your work has been a force of nature that energizes me.

Colin Gerker-Junsay, you are the god-tier coach. Thank you for cheering me on and always supporting me. Your energy, enthusiasm, and interpersonal skills light up my life!

Jimmy Vo and Jenny Vo, thank you for being super siblings! Jimmy, I have thoroughly enjoyed collaborating with you on research and having fascinating conversations about deep academic topics! Jenny, thank you for helping Judy and me think about equity scholarship from the perspectives of Gen-Z students.

Susan Payrovi and Cynthia Tsai, thank you for being supportive of my health and well-being. Susan, you have taught me so many life strategies that have transformed my wellness. Cynthia, you're an amazing advocate who has truly supported my wellness!

Calvin Nguyen and Toai Nguyen, thank you for the smile you bring to my life! Calvin, our conversations about academia, spirituality, and living a healthy life has been a pillar of support for my PhD. You're the best dentist!

Christopher Hunn, thank you for supporting me for years! You're my hero and a true equity champion. You're one of the main reasons why I'm graduating, and I thank you so much for the pivotal role you have played in my success as a scholar. Together, we have transformed pedagogy and practice on campus and beyond. You're a brother and family to me, and I hope to continue moving forward with you on our radical vision for transformative change!

Abhi Roy and Ryan Ng, you're my heroes, my brothers, and the rock stars of my life!

Jasmine Cordero, Emily Hoang, Dena Marie, and Brendan Henrique, you're also my superheroes and rock stars!

Juliana Chiarelli, you're my hero! I appreciate the time and mentorship you have provided. You have a true equity-oriented way of working with students and I am so happy to have met you! Thank you for changing my life!

Heather Yaden, I appreciate your time and support. You're amazing and a wonderful mentor who inspires me! You have been instrumental in my growth as a scholar and human being.

Kasra Sotudeh, you're MVP status and a true social justice superhero! I enjoy our conversations. They invigorate me, inspire me, and catalyze my passion and drive to continue doing equity work! You truly go above and beyond in exemplifying the spirit of Mary Slakey, and you bring your own captivating charisma to all the work that you do! You have been life changing for me and have transformed my development, perspectives, and growth as a human being.

Monica Gutierrez, you are my hero and I have learned so much about EOP from you! Your wisdom, compassion, and student-centered nature inspire me.

Liz Keithley, you are another one of my true Jedi masters. Learning from you has been a miraculous experience. I appreciate you and the heart you bring to this world. You have inspired me, and your spirit will forever live through the work that I do. You're my life-changing, superhero!

Todd Zakrajsek and Deb Van Etten, thank you for being my faculty development superheroes and mentors. The POD and Lilly conferences are my favorite conferences. Todd, you have led the best 8 a.m sessions for these conferences and you are a dynamic lecturer and equity champion superstar! You and Deb Van Etten are my heroes!

Isis Artze-Vega, Carl S. Moore, Hoag Holmgren, and Hilary Schuldt, thank you for also being my educational development superheroes! You have truly supported my growth as a scholar, educational developer, researcher, and human being! Isis, I appreciate the time you took to help orient me to the higher education organizational change literature in my earlier years as a scholar. You, Carl, Hoag, and Hilary and the POD Network helped me feel like I belong in the field of educational development as a Latinx and Vietnamese scholar and learner.

I am also grateful for the energy, support, and community of folks in the Equity-Oriented Advising and Coaching Program: Susan Hagstrom, Isela Pena-Rager, Colin Gerker-Junsay, Denzil Streete, Donna Vivar, Heather Yaden, Katrina Pantig, Larry Jackson, Layla Naranjo, Liza Schlosser-Olroyd, Sunny Lee, Syrhonda Calhoun, and Ying Kuah. Thank you for being part of the movement of cultivating a community of equity champions.

Dania Matos, I could not thank you enough for being not only a key collaborator in my work but also an equity champion and Vice Chancellor for Equity and Inclusion. I am still inspired when you said the following at the Equity-Oriented Advising and Coaching program: "I might have wanted something to look a certain way, but my people may not need that." That is a vivid example of perspective taking and it continues to fuel the heart of the adaptive equity-oriented pedagogical framework. Echoing your words, one of my core life goals is to lead from and with love. With you, Fab and so many more, I hope to continue disrupting, decolonizing, and radically transforming institutions to center equity, justice, and belonging.

Albert Phuong, Tony Phuong, Jenny Phuong, Luann Phuong, Lucas, Steve, thank you for creating a wonderful environment and for being amazing people who have transformed my life. Much of my passion for supporting learners is inspired by all of you.

Thank you to my other heroes: Dolores Tolano, Tio Miguel, Lauren Rios Tolano, Sabrina Rios Tolano, Michael Rios Tolano, Kevin Rios Tolano, Marie Miles, Ricky Baja (Chris Miles), Brian Miles, Mami Pera, and Isidro. You all are a major part of my life and have shaped me into who I am today.

I also appreciate my Papi Felix, Bertha, Liz Estrada, Elsa Estrada, Chris Estrada, Hugo Estrada, Bella Estrada, Alex Estrada, Tio Nino, Tia Santos, Edgar Quezada, and the rest of the Estrada family. You have played a pivotal role in my development and in my passion for social justice and Latinx topics.

Esthela, Matthew, and William Phuong, thank you for being the world that nurtured me and helped me become the person I am today. You have instilled a passion for equity in me that will forever live and animate the work that I do!

Fan Huang, thank you for being supportive through health issues and life. Your work, designs, and art are groundbreaking. You're a true superhero and your energy and enthusiasm inspire me. You truly make a difference and can figure out problems that are complex and feel impossible to overcome. I have never met anyone like you, and I thank you for having my back and supporting me. Even though we're not part of the same research team or program, I feel like I have learned so much from you, and you have helped me grow so much. You're everything and more, a true measurement and assessment superstar! The love and passion that you put into your work is empowering. No measurement model can capture the brilliance, intellect, and awe that you bring to this world. You energize, invigorate, and excite me to think beyond my imagination.

I would not have made it through my dissertation without the unconditional support of baby yodas including Duy, Yen, An, Sippy Mo, Maddy Mo, Pandan, Mungbean, Mochiball, Coco, Eggy, Tiato, Dr. Grogu, and many more.

I would like to thank Judy Nguyen for keeping me happy, healthy, and hopeful. Your intellect and powers transcend the infinite walls of the universe. You empower so many with your therapeutic nature, and you bring a god-tier level of empathy to all your interactions. You bring kindness, compassion, and hope to this world, looking for the best in people and bringing out their strengths. Thank you for being there for me health-wise and emotionally. You're a groundbreaking researcher with a brilliant and creative imagination— you have an ability that is beyond this world. Your work and rigor have always inspired me to be better. You have helped me learn critical perspectives that are essential to pursuing the equity and anti-racist work that I care so much about. Your skills with language and understanding people's emotions are truly unparalleled. You have demonstrated and identified the social-emotional competencies that can forever change this world. Thank you for embodying and bringing those principles to life.

I want to thank one of my early mentors, John Hurst, who helped me catalyze my work on equity and inclusion. Thank you for being my first faculty sponsor and advisor in the Graduate School of Education.

From the bottom of my heart, I would also like to give special thanks to faculty, researchers, professionals, family, and friends who have supported me on my academic journey. These individuals include Igor Chirikov, Alexander Blum, Alex Chaparro, Kurt Organista, Laura Perez, Oscar Dubon, Matthew L. Miller, Brendan C. Russell, Terrence Tivnan, John Hansen, Drew Bridewell, Jacqueline Martinez, Chris Dede, Darren Zook, Melanie Hom, Austin Alquisola, Tony Mirabelli, Carol Christ, Eileen Berger, Michael Berger, Hoag Holmgren, Brandon Geller, Amy Scharf, John Michael Scott, José Ramón Lizárraga, Artur Daylidonis, Matt W. Courtney, Judy Alessandri, Tom Alessandri, Karen Brennan, Joe Campos, Tenzin Kyizom, Elsie Carrillo, Amy Cheung, Maleka Donaldson, Americ Azevedo, Sam Bliss, Martha Velasquez, Silvia E. Marquez, Meng So, Liliana Iglesias, Ruben E. Canedo, Nick Creech, Michael Downs, Carol Giraudo, Eugene Arthur Haggerty, Abhishek Roy, Peter Hurtubise, Jeremy Lum, Russ Marcel, Esteban Leiva Parker, Jerry W. Sanders, M. Pera, Jon Scherbart, Julie Shackford-Bradley, Carolyn Swalina, Sonia Torna, Leah Johnston, Liz Grossman, Steve Pinkston, Roben Torosyan, Kitty Boles, Rilda Kissel, James Ryan, Katie Kearns, Elizabeth Luoma, Cecilia Lucas, Olufemi Ogundele, Yuki Burton, Julian Ledesma, Zach Pardos, Cruz Grimaldo, Brigid Barron, Bertrand Schneider, Matt Young, Chris Meyercord, Kristina Luscher, Kevin Barry, Phuong family, Estrada family, Tolano family, S Yeh, Catherine Koshland, Cyrell Roberson, Annie Liang, Patrick Shields, Sandra Zuniga Ruiz, Yuki Watanabe, Roberto Gonzales, Thomas Hehir, Rick Mintrop, Aaron Ninsei Zagory, Eric Shed, Crystal Ward, Kate Perry, Don Davis, Trevor Osaki, Lauryn Holloway, Berrin Boyce, Esthela Estrada, Matthew Phuong, and William Phuong, and all the folks in my communities.

I would also like to thank the following UC Berkeley units: the Division of Equity and Inclusion, Centers for Educational Equity and Excellence, Educational Opportunity Program, Undocumented Students' Program, the Disabled Students' Program, the GSI Teaching and Resource Center, the Graduate Division, the Chicanx and Latinx Standing Committee, the HSI Taskforce, the Latinx Thriving Initiative, Division of the Vice Chancellor for Undergraduate Education, Cal Reads, the Berkeley School of Education, and the Center for Teaching and Learning.

Beyond Berkeley, I would like to also acknowledge the Harvard Graduate School of Education (HGSE), HGSE Access and Disability Services, and the Professional and Organizational Development (POD) Network in Higher Education.

In community and in collaboration with my pack, I have not only survived, but have also thrived as a T-rex who is on a life-long mission to advance equity, inclusion, belonging, and social justice at scale.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: Introduction

1.1 Statement of the Problem and Research Questions

Research shows that 90% of students leaving STEM have cited poor teaching as a primary concern (Seymour & Hewitt, 1997). Reasons enumerated by students include a lack of support, limited real-world applications, and a lack of alignment between instruction and assessment. High attrition rates in STEM continue to persist, especially for underrepresented minority (URM) students (Sax et al., 2018).

Engaging STEM faculty and instructors[1] in professional development presents challenges (Wieman, 2017). In R-1 universities, many STEM faculty and instructors do not find professional development engaging, relevant, worthwhile, or rewarding (Wieman, 2017). Additionally, in R-1 universities, teaching is often not rewarded as much as research (Fairweather, 2008). Therefore, many faculty and instructors often teach the way they have been taught in the past (Davis, 2009), which can be inequitable. Nevertheless, Condon et al. (2016) have found that instructor professional development can directly impact instructors' pedagogy and their students' learning. Improving STEM students' learning is critical, because students' academic achievement is often associated with their subsequent achievement, retention, and pursuit of professional goals (Cole & Espinoza, 2008; Sax et al., 2018).

Unlike K-12 teachers, many college faculty have limited to no training on how to teach because teaching certification is often not required for higher education instructors (Davis, 2009). Consequently, there is less training on how to more equitably teach diverse groups of students who hold intersectional identities (e.g., race, ethnicity, gender, disability). The training that college instructors might receive is often through educational development programs, which refers to pedagogical training for college instructors (e.g., faculty, graduate students). These trainings are offered through pedagogy workshops, pedagogy courses for graduate student instructors, and teaching certificates. I will refer to these trainings collectively as "educational development" which includes both faculty and graduate student professional development on pedagogy.

To create a more equitable learning environment for diverse groups of students, I have co-designed and studied equitable teaching practices that improve success for all students. With collaborators at Berkeley and Stanford, I have co-developed the Adaptive Equity-Oriented Pedagogy (AEP) framework. AEP is a framework for adapting teaching to address equity barriers to learning based on student data (e.g., formative assessment, observations, surveys) (Phuong et al., 2017). For example, AEP entails using such data to adapt instructional practices (e.g., productive struggle, modeling key skills, deliberate practice, providing feedback) (Phuong et al., 2017). Prior studies show that STEM instructors applying AEP have improved student achievement by over a letter grade, on average, when compared to active learning control conditions (Phuong & Nguyen, 2019). I provide access to a tip sheet on AEP for instructors in Appendix A.

[1] In this dissertation, I refer to faculty members as a "faculty". I use the term "instructor" for graduate and undergraduate students who serve in an instructional capacity under faculty direction and are enrolled in the pedagogy courses. The term "facilitators" refers to those who teach the pedagogy courses.

My dissertation research focuses on studying what factors drive college STEM instructors' adoption of equitable teaching practices, such as AEP. I explore the impact of these teaching practices on student success. For my study, STEM instructors participated in AEP pedagogy courses where they completed a teaching portfolio that included written reflections, evidence of student learning, and peer/student feedback on their teaching. Instructors were randomly assigned to participate in either a treatment pedagogy course that adapted weekly to instructors' learning needs, or a control course that was less adaptive and responsive to their learning needs. The treatment adapted the pedagogy course curriculum based on instructors' weekly written teaching reflections. By contrast, the control did not. Thus, the treatment focuses on explicitly modeling AEP on the instructors so that they can apply AEP practices with their students.

In particular, I am interested in examining how treatment and control pedagogy courses impact instructors' growth in AEP competency from pre- to mid-semester and pre- to post-semester. I control for instructors' gender and URM identities because instructors with non-male and URM identities may have a more positive disposition towards learning AEP, since STEM is often a toxic environment for women and students of color (Sax et al., 2018). I control for years of teaching experience and tutoring experience because teaching experience has been shown to impact teaching effectiveness and student success (Podolsky et al., 2019). I also control for the number of students of each instructor, since adapting based on student learning data may be more feasible for smaller classes. Based on these interests, I explore the following research questions:

1. **Quantitative research question 1:** Is there a greater increase in mean AEP competency scores across time in the treatment pedagogy course than in the control pedagogy course, controlling for instructors' gender, URM status, years of teaching experience, years of tutoring experience, and number of students?
2. **Qualitative research question 2:** What mechanisms for instructional change characterize differences between low- and high-AEP growth instructors in the treatment?

My inquiry is inspired in part by my own experiences as a low-income historically underrepresented student who comes from a Latinx and Vietnamese immigrant household whose first language is not English. I have experienced how the K-16 pipeline can feel like a maze without a map, especially with access to limited resources. I am committed to restructuring this maze into a highway towards greater access, equity, and inclusion.

1.2 Summary of Findings

Research question 1 findings. Using multilevel regression modeling, I found that the treatment condition, on average, significantly improved instructors' AEP competencies over time compared to the control condition (2.63 standard deviations, $p<0.001$), when controlling for instructors' gender, URM status, years of teaching experience, years of tutoring experience, and number of students. Compared to the control condition, more instructors in the treatment condition used multiple data sources to adjust teaching and address equity barriers to advance student learning. These findings suggest that educational developers can model principles from AEP (e.g., reviewing and addressing instructors' beliefs and lenses based on written reflections) to improve instructors' equitable teaching competencies in professional development programs.

Research question 2 findings. Qualitative analyses of treatment instructors' teaching reflections revealed differences between low- and high-growth instructors that illuminate mechanisms for adopting adaptive, equity-oriented teaching practices. I found that, compared to low-growth instructors, high-growth instructors more frequently shared the responsibility for learning with their students, and used multiple data sources (e.g., formative assessment, observations, surveys) to understand equity barriers (e.g., imposter phenomenon, stereotype threat) and to adopt teaching practices that address these barriers and improve students' success.

Compared to low-growth instructors, the high-growth instructors more frequently aligned their academic and equity goals for student learning (e.g., promoting students' sense of belonging as they developed academic skills). Similarly, high-growth instructors more often defined equity as adapting to students' needs than low-growth instructors. A few low-growth instructors defined equity as giving everyone the same thing, although they did not invoke that definition consistently when describing pedagogical practices. Furthermore, compared to low-growth instructors, high-growth instructors more frequently mentioned how equity barriers fostered negative emotions and affect (e.g., feeling excluded as a student), which motivated them to adapt their teaching and promote a more inclusive and welcoming environment.

1.3 Contribution: How this research advances the current state of knowledge in the field

This study contributes insights on how to support instructors with learning and adopting equitable practices, thereby addressing a long-standing barrier to pedagogical reform in STEM. The treatment focuses on explicitly modeling equitable teaching practices, such as AEP, on the instructors so that they can apply these practices with their students. The results obtained from analyses of the treatment and control conditions reveal successful strategies for supporting instructors in using multiple forms of data effectively– rather than relying on their assumptions– to improve student success. The adaptive pedagogy course model, which leverages AEP, can be beneficial for the implementation of effective educational development programs and the advancement of equity in student outcomes.

The original contribution of my work includes leveraging learning sciences theories to build and validate equity-oriented competencies that foster more inclusive classrooms. These competencies are grounded in learning theory from higher education and have been shown to significantly improve student success (Phuong et al., 2017; Phuong & Nguyen, 2019; Phuong et al., 2022). Through deeper qualitative analysis of instructors' reflections, I have also identified mechanisms for instructional change to explain differences between low- and high-growth instructors of AEP competency.

CHAPTER 2: Literature Review

The field of higher education has found that educational development programs can increase instructors' adoption of active learning practices, which have been shown to improve students' success. However, less research has been conducted on how educational development programs can support instructors' use of equitable pedagogies. In addition, there is limited research on how such programs address and support the mechanisms that motivate instructors' learning of equitable teaching practices. To build on this literature, I explore theories of pedagogical change to enrich our understanding of how instructors learn and adopt equitable teaching practices. These theories of pedagogical change provide insight on the mechanisms that can help explain instructors' barriers and opportunities for learning, which can be addressed through educational development interventions. These theories inform the design, implementation, and analysis of the educational development interventions for my dissertation.

Broadly, the field of "educational development" in higher education includes faculty development and graduate student development around teaching, learning, and mentoring (POD Network, 2016). This field began to emerge in the 1970s (North & Scholl, 1979) with professional development that focused on developing competencies in teaching. Boyer's (1990) *Scholarship Reconsidered: Priorities of the Professoriate* amplified the need for educational (faculty and graduate student) development initiatives that were already gaining momentum at colleges and universities. Boyer (1990) reimagines teaching as a site of scholarship that exists parallel to the scholarship of discovery (e.g., original research). Boyer (1990) acknowledges that faculty and graduate students are often not trained in teaching and are often not taking a scholarly, or research-based, approach to teaching. Therefore, Boyer (1990) recommends the use of research-based teaching practices in higher education.

The academic subcommunity known as the Scholarship of Teaching and Learning (SOTL) emerged from Boyer's (1990) identification of teaching as a site of scholarship. SOTL asks instructors to conduct classroom research and improve their pedagogy. As interest in this form of pedagogical scholarship grew, some forms of educational development focused on pedagogy across disciplines, while others became more discipline-based. An example of the former can be found in the integration of SOTL principles by scholars and practitioners of Faculty Learning Communities (FLCs). FLCs bring together different disciplines so that faculty and instructors can learn from each other (Cox, 2004).

An example of the latter can be found in Discipline-Based Education Research (DBER), which has been particularly common in Science, Technology, Engineering, and Mathematics (STEM) disciplines. The DBER community examines how student outcomes are impacted by particular instructional approaches such as active learning, think-pair-share or clickers in STEM education (Deslauriers et al., 2011; Freeman et al., 2014; Mazur, 2009). DBER research has influenced educational development programs to incorporate more active learning strategies and evidence-based teaching (Wieman, 2017).

A subset of DBER researchers have also looked more specifically at educational development programs themselves as a focus of research, and how such programs impact the adoption of active learning practices in STEM. Literature in this field has found that sustained educational development programs (focused on evidence-based active learning practices) can improve instructors' active learning practices and student outcomes (Brown & Kurzweil, 2014; Condon et al., 2016; Derting et al., 2016; Wheeler & Bach, 2021). Theobald et al. (2020) found

that, on average, active learning in STEM courses narrowed achievement gaps in passing grades and exam scores for all students, but was particularly beneficial for underrepresented students.

Less research on educational development programs, however, has focused specifically on equitable teaching. By *equitable*, I refer to pedagogies that address students' learning needs (e.g., interests, areas of growth, strengths) so that they have greater opportunities and outcomes for academic achievement and psychosocial development (Phuong et al., 2017). Therefore, in the remainder of my review of empirical work in educational development, I restrict my attention to literature that explores whether and how sustained educational development programs focused on equity can impact instructors' equitable beliefs and practices.

I begin with a brief review of literature from the educational development field focused on defining the practice and assessment of equitable teaching. I then focus more specifically on empirical studies that explore multi-session educational development programs geared toward equitable STEM instruction that include pre- and post-assessments of teacher belief or practices. Motivated by my research questions, I then examine what these studies can tell us about whether and under what circumstances such programs (a) *shift instructors' beliefs and practices related to equitable teaching*, (b) *impact student outcomes*, and (c) how programs *conceptualize and attempt to motivate instructors' learning or pedagogical change*. To better understand the processes by which educational development drives instructor learning, I then review more general theories of pedagogical change and adoption that are focused on data-based and equitable pedagogy from the fields of educational development and the learning sciences.

2.1 Equitable Teaching in the Field of Educational Development

In the field of educational development, equitable teaching beliefs are often defined as honoring and addressing students' backgrounds and identities in ways that do not marginalize students (Parker et al., 2016). The field of educational development has provided multiple examples of equitable teaching practices, which include clarifying learning outcomes and expectations, aligning assessments and activities with outcomes, and responding to equity barriers and student needs in the classroom.

2.1.1 Practices and Policies that Promote Equitable Teaching

Fink (2013), Winkelmes (2014), Howard et al., (2020), Davis (2009), Weinstein (2009), and McCallum (2013) highlight the importance of communicating, clarifying, and supporting students in meeting rigorous learning outcomes. To support students in reaching expectations, Fink (2013) emphasizes the need to align formative assessment and course activities with learning outcomes so that students can have exposure and a clear understanding of the expectations. Phuong et al. (2017) further argue that one of the major challenges in STEM college courses is a lack of alignment between summative assessment, formative assessment, and instruction. This lack of alignment engenders equity issues, since it often privileges students who have more background knowledge or those who have more access to more financial resources to support their success.

To further support all students' success, designing for equity also requires one to build and clarify course policies (e.g., grading, assessment, accommodations) that are informed by the equity literature (Feldman, 2018; Phuong et al., 2017). Examples of equitable course policies include showing flexibility with deadlines, providing opportunities for re-grading, not curving down, and being willing to adjust policies based on students' needs and contextual

circumstances.

Assessment can also play a role in promoting equity, since it can help instructors diagnose student needs and adjust instruction accordingly. McCallum's (2013) assessment to instruction model attends to equity by directly linking assessment to instructors' diagnosis of students' existing competencies and interests. Black and Wiliam (2001) have described how formative assessments, including diagnoses of students' current abilities, is an important part of adjusting instructional practices to address students' learning needs in the classroom.

Equitable learning experiences can also be promoted through universal design for learning, which offers multiple means of engagement, representation, and action and expression in the learning process (Griful-Freixenet et al., 2017; Pace & Schwartz, 2008; Rose et al., 2006). The field of universal design for learning (UDL), a popular model in education, provides students with and without disabilities different ways to learn, demonstrate, and reinforce knowledge so that they have more equitable opportunities and greater access to academic success.

In addition to specific instructional practices such as diagnosing student needs and enabling multiple means of engagement, research suggests that instructors' own explicit and implicit beliefs around teaching and learning (e.g., growth mindset) have been shown to impact student achievement (Canning et al., 2019). In fact, Canning et al. (2019) state that a longitudinal study of 150 STEM professors and over 150,000 students suggested that faculty with a fixed mindset had achievement gaps that were twice as large as those with a growth mindset, when controlling for the faculty members' teaching experience, age, tenure status, race and ethnicity, and gender.

Culturally relevant and culturally responsive pedagogy (CRP, grounded in the scholarship of K-12 educators such as Gloria Ladson-Billings (1995), Gay (2010), Hammond (2014), and García and Ortiz (2006)) is another model of equitable teaching that has been used in higher education. CRP addresses the cultural and identity-based needs and backgrounds of students (García & Ortiz, 2006; Gay, 2010; Ladson-Billings, 1995). For example, through CRP, instructors can engage students in project-based learning where they can expand their critical consciousness by examining and addressing real-world problems that impact social inequities (Hammond, 2014; Ladson-Billings, 1995). In the K-12 and higher education literature, culturally responsive learning experiences can enable instructors to broaden their interpretation of culturally and linguistically diverse students' learning, reflect on the broader systemic and institutional barriers to learning, recognize triggers around race and culture, be aware of cultural stereotypes and archetypes of individualism and collectivism, examine their biases and limitations, and promote a sense of belonging (Cochrane et al., 2020; Hammond, 2014). Furthermore, within a CRP framework, instructors can use formative assessments and feedback to increase learners' engagement and success (Hammond, 2014). Under this approach, students can be prepared for more rigorous tasks since they are offered repeated opportunities to collaboratively practice the habits of mind, academic discourse, and expert thinking within a scientific community of learners; such opportunities are often unavailable for URM (Hammond, 2014). In related literature, pedagogies that promote minoritized students' success (e.g., Latinx, Black, Native American students) have engaged students' cultural backgrounds, funds of knowledge, interests, and identities (Bernal, 2002; Brockenbrough, 2016; Garcia & Okhidoi, 2015; Grande, 2015; Moll et al., 1992; Valenzuela, 2010).

2.1.2 Measures on Equitable Teaching in Higher Education

Course evaluations are often used as one of the primary measures of teaching effectiveness in higher education (Stark & Freishtat, 2014). A potential flaw of course evaluations is that they have been found to be biased against women and people of color (Boring et al., 2016; MacNell et al., 2015; Stark & Freishtat, 2014). Stark and Freishtat (2014) also found that course evaluations appear to have little to no correlation with learning.

There have been efforts in the field of educational development to measure equitable teaching in higher education. These primarily take the form of observation-based equity-based measures of teaching and learning in higher education. For example, there are observation checklist measures of universal design for learning (CAST, 2018) and culturally responsive pedagogy (Hammond, 2014) in higher education. These measures ask observers to identify whether instructors are using formative assessment, are explaining course content in different ways, and if they promote participation for all students. The validity evidence associated with such a checklist is quite limited, including no statistical analysis.

I also found research that applied factor analysis to self-report survey responses focused on equitable teaching in higher education. For example, O'Leary et al. (2020) used exploratory factor analysis to support the validity and internal reliability of surveys measuring equitable teaching. These surveys had Likert scales that asked instructors the extent to which they agreed with statements, such as: whether they could improve student success through their teaching styles, whether all students could succeed, and whether students should be encouraged to leave STEM if they are not cut out for the major. Rooney et al. (2020) also adapted a validated survey to examine equitable teaching in higher education. The survey had been previously validated in K-12 using exploratory and confirmatory factor analyses (Yoon Yoon et al. 2014). Rooney et al. (2020) focused on measuring instructors' confidence in a) promoting positive student-student and faculty-student interactions and b) mitigating harmful psychosocial effects and unconscious biases.

Item response theory (IRT; Lord, 1980) has been used to develop more psychometrically rigorous measures of student learning (Blum et al., 2020; Glas, 2008) and pedagogical effectiveness (Chang et al., 2019; Chang, 2022). A primary advantage of using item response theory (IRT) is that doing so allows for coherent analyses at both the item-by-item and overall instrument levels (Hambleton & Van der Linden, 1982; Lord, 1980; Wilson, 2005). As such, IRT effectively accounts for differences in item level difficulty reflecting, for example, the reality that some equity-oriented practices require more advanced skills than others. Simultaneously, IRT distinguishes overall levels of proficiency on the part of instructors, reflecting, for example, that some instructors are more competent in applying equitable teaching practices than others. Importantly, IRT permits aggregation of information collected from a variety of item formats, so one can, for example, score written responses that contain evidence beyond self-report. In addition, IRT enables the analysis of differential item functioning, which can detect when an assessment contains items that may be biased against subgroups of people taking the assessment (Paek & Wilson, 2011). This approach would allow, for example, the examination of whether an assessment of equity-oriented teaching competency demonstrates bias based on gender or URM status. Taken together, the literature suggests that rigorously validated equity-oriented practice measures that go beyond course evaluations should be achievable.

2.2 Empirical Literature on STEM Educational Development for Equitable Teaching

To understand prior work exploring the impact of STEM educational development on equitable teaching practices such as those described above, I conducted a systematic literature search utilizing Esterhazy et al.'s (2021) framework for reviewing literature in college educational development. Table 2.1 below describes my inclusion and exclusion criteria. Because there is a lot of literature in this area, I elected to restrict my search to studies that leveraged methods that are considered more rigorous in the field and that are similar to the methods used in my own research. Most notably, papers were included if they utilize a pre- and post-assessment (including self-reported surveys), which is considered a "gold standard" for examining impact in the field of educational development (von Hoene et al., 2017).

In Google Scholar, I entered the following search phrases one at a time with no quotation marks or Boolean operators:

- STEM faculty development higher education
- STEM educational development higher education
- STEM graduate student development higher education
- STEM college training equity
- STEM college training PERC
- STEM college training SIGCSE
- STEM college training ASEE
- STEM college training RUME
- STEM Professional development inclusive teaching
- college instructor professional development STEM equity[2]

[2] I selected the discipline specific conferences that I was aware of at the time that have discussed teaching and learning in higher education STEM disciplines. These include Physics Education Research Conference (PERC), Special Interest Group in Computer Science Education (SIG CSE), American Society for Engineering Education (ASSEE), and Research in Undergraduate Mathematics Education (RUME).

Table 2. 1

Inclusion and Exclusion Criteria for Literature Review

Topic	Inclusion	Exclusion
Program Type	● Multiple Sessions over multiple months ● Supports STEM instructors and discusses STEM content ● Focuses on equitable and inclusive teaching ● Focuses on supporting students with disabilities and those from Latinx, Indigenous, LGBTQ+, African, and gender identities among others	● Single Session ● Non-STEM programs only ● No pre- and post-assessment of teaching
Study Design	● Describes data collection and analysis methods ● Participants include STEM instructors (e.g., faculty and graduate/ undergraduate student instructors) ● Pre- and post-assessment of teaching beliefs and practices ○ Assesses impact on teaching practices ○ Contains self-report (e.g., surveys), observations, and/or evidence of growth in teaching ● Assesses impact of educational development programs of student outcomes (e.g., achievement) ● Contains rich, qualitative empirical data (e.g., pre- and post-observation measures, instructor/ student feedback)	● Conceptual only (no data) ● Only a literature review ● No description of data collection and analysis methods ● Focuses only on summative impact of educational development programs
Activity choices for instructors in educational development programs	The activity choices for instructors may include any of the following: ● Lecture ● Discussion and Active Learning ● Q+A ● Classroom observations and feedback ● Modeling of evidence-based practices ● Reflection activities ● 1-1 consultations/ office hours ● Discipline specificity occurs in the context of the program	● Not situated in higher education ● No information on practices of discussing teaching
Publication characteristics	● English language ● Peer reviewed ● Published before February 2022	● No peer review ● No full text available

I found three studies in peer-reviewed journals (see Table 2.2) and two conference proceedings (see Table 2.3) that met the inclusion criteria. All the studies identified suggest that

educational development programs can positively shift instructors' equitable teaching beliefs and practices. The literature summarized in Table 2.2 and 2.3 all used self-report pre- and post-survey responses to gauge the effectiveness of educational development programs on instructors' equitable teaching beliefs and practices. As I describe in further detail below, while these studies provide useful starting points to consider how such programs impact instructor beliefs and practices (RQ1), the studies do not examine the impact of these programs on student achievement. In addition, most of the studies do not explicitly offer mechanisms for explaining pedagogical change and adoption (RQ2).

Table 2. 2

Peer-Reviewed Journal Studies that met Inclusion Criteria

Questions	Parker et al. (2016)	O'Leary et al. (2020)	Harrison-Bernard et al. (2020)
What equitable teaching competencies were focused on?	Addressing micro-inequities	Culturally responsive practices	Anti-racist practices
What did the program do to improve equitable teaching competencies?	Instructors evaluated a teaching intervention	Instructors reflected on their social identity, engaged in dialogue, listened to student perspectives and read pedagogy research about barriers to student learning	Instructors learned through didactic modules, videos, teaching modules, and dialogue
What theory or mechanism(s) is used to explain adoption of equitable teaching?	Instructors' equitable teaching beliefs are developed when their prior beliefs and predispositions are addressed	Equitable pedagogy is fostered when instructors reflect on their social identity, develop self-awareness, mind the privilege gap, and reduce implicit bias	Equitable pedagogy will improve if one is more knowledgeable and attentive to diversity, equity, inclusion, and implicit bias
What was the measure of equitable teaching? How, if at all, was the measure validated?	Pre- and post-survey States survey was previously validated but offers no information on validation	Pre- and post-survey Used exploratory factor analysis to support the validity and internal reliability of surveys	Pre- and post-survey Not validated
What is the type of control group? What is the type of assignment to treatment and control?	No control group, not randomized	No control group, not randomized	No control group, not randomized
What statistical or analytical framework was used to measure growth?	*t*-test / Bonferroni Method	Wilcoxon signed-rank tests	*t*-test
Did the program impact equitable beliefs?	Mixed results via self-report	Yes via self-report	Yes via self-report

Did the program impact equitable teaching practices?	No evidence	No direct pre- and post-assessment measure; however, faculty retrospectively mentioned in post-surveys and discussions how they would promote equitable teaching and culturally responsive teaching	No direct evidence; provides retrospective evidence from faculty who mention that they would use equitable practices

Table 2. 3

Higher Education Conference Proceedings that met Inclusion Criteria

Questions	Rooney et al. (2020) Not archived	Metevier et al. (2010)
What equitable teaching competencies were focused on?	Culturally Responsive Classroom Management, self-confidence in equitable teaching and mentoring practices	How to teach equitably and be responsive to students' backgrounds through engaging their interests and collaborative learning
What did the program do to improve equitable teaching competencies?	Mini-modules instructed faculty on equitable teaching and mentoring methods	Workshops helped instructors weave equitable instructional strategies into the inquiry laboratory activities they design and teach
What theory or mechanism(s) is used to explain adoption of equitable teaching?	Evidence-based mini-modules can increase faculty self-confidence in equitable teaching and mentoring practices.	Equitable teaching can improve when participants learn about equity topics/ growth mindset and design and teach their own inquiry activities
What was the measure of equitable teaching? How, if at all, was the measure validated?	Pre- and post-workshop surveys (3 questionnaires) Survey was previously validated in K-12 using exploratory and confirmatory factor analyses (Yoon Yoon et al. 2014)	Pre-and post-workshop surveys Not validated
What is the type of control group? What is the type of assignment to treatment and control?	No control group, not randomized	No control group, not randomized
What statistical or analytical framework was used to measure growth?	Compared means, *t*-test	Scored survey responses Compared means, *t*-test
Did the program impact equitable beliefs?	Yes, faculty self-report more confidence in using equitable teaching practices, such as Culturally Responsive Classroom Management	Yes, faculty self-report being more responsive to the different ways students learn

Did the program impact equitable teaching practices?	Yes, self-report on survey but no direct pedagogical evidence or observation data are provided	Yes, self-report in written survey responses but no direct pedagogical evidence or observation data are provided

2.2.1 Effects on Faculty Beliefs and Practices

All three of the articles I identified in peer-reviewed journals, Parker et al. (2016), O'Leary et al. (2020), and Harrison-Bernard et al. (2020), found that educational development programs can positively impact instructors' equitable teaching beliefs. All these authors measured growth in equitable teaching beliefs through pre- and post-surveys.

Parker et al. (2016) examined how educational development programs improved 28 STEM community college faculty members' beliefs and understandings of equitable teaching practices that address micromessages and microinequities. The program focused on asking instructors to learn through experience by designing and evaluating a teaching intervention in their own classrooms. The study employed pre- and post-surveys and found that there were statistically significant shifts in faculty members' understanding of seven of the nine constructs related to equitable teaching measured in the surveys. Shifts were determined by examining whether there was a significant gain for each construct from pre- to post-survey on a Likert Scale (1= strongly agree, 2= agree, 3= disagree, 4 = strongly disagree). The seven items for which participating faculty demonstrated significant pre-post gains in beliefs were all related to knowledge and confidence in equitable teaching. However, there were two items where faculty did not demonstrate gains, both related to instructors' beliefs that all students can succeed in STEM. There was no control group, and the study did not observe or directly measure adoption of equitable teaching practices.

Over the course of 3 years, O'Leary et al. (2020) studied how educational development programs shifted 3 different cohorts of STEM faculty members' (n=115) equitable teaching beliefs and attitudes. These programs prompted instructors to reflect on their social identity, engage in dialogue, listen to student perspectives, and read pedagogy research about barriers to student learning. The authors measured shifts in beliefs and attitudes as recorded in self-reported pre- and post-surveys. They found statistically significant positive shifts in responses to three items related to instructors' beliefs that all students can succeed in STEM. In particular, two items focused on whether the instructor agreed that it was their role to provide greater opportunities for success. Another item focused on whether the instructor thought they could improve student success through changing their teaching style. The last item asked whether instructors agreed if some students are not cut out to be science majors and should be encouraged to leave the major. The study also found statistically significant increases in faculty's knowledge of students' learning barriers and social identities in college STEM classrooms, especially for those from URM and lower socioeconomic status backgrounds (O'Leary et al., 2020).

In the O'Leary et al. (2020) study, there was no direct pre- and post-assessment measure of equitable teaching; however, faculty retrospectively mentioned in post-surveys and in discussions that they would promote equitable teaching and culturally responsive teaching upon returning to campus. The authors expressed that these responses provided evidence that the educational development program inspired action to modify teaching practices; however, the authors acknowledged that their measure of impact on teaching practices faces limitations and proposed classroom observations as a way to address this issue.

Harrison-Bernard et al. (2020) researched how educational development programs improved STEM faculty members' (n=47) awareness of equitable teaching strategies and topics. These programs created learning experiences where instructors learned through didactic modules, videos, teaching modules, and dialogue. Based on pre- and post-surveys, the educational development program significantly increased instructors' perceived ability to a) leverage equitable communication strategies with students from diverse backgrounds, and b)

recognize biases and stereotypes in graduate education and in mentorship for URM students (Harrison-Bernard et al., 2020). The educational development program significantly improved faculty members' knowledge of equitable teaching topics, such as “implicit bias”, “color-blind racial attitudes,” “failure to differentiate”, and “tokenism” (Harrison-Bernard et al., 2020, p. 1).

Harrison-Bernard et al. (2020) expressed that their study faces limitations with examining the impact of educational development on equitable teaching. The authors state that, in the future, they will create a follow-up survey to assess faculty perceptions of how they integrate information from workshops into their pedagogical “practice, perceptions, thoughts,” and “decision-making” (Harrison-Bernard et al., 2020, p. 293).

The two conference papers yielded by my search also reported positive impacts of educational development programs focused on equitable teaching. Rooney et al. (2020) share research on an educational development program that used mini modules to instruct faculty on equitable teaching and mentoring methods. Rooney et al. (2020) found that faculty reported gains in confidence in their ability a) to promote positive student-student and faculty-student interactions and b) to mitigate harmful psychosocial effects and unconscious biases. In this study, faculty self-report changes in their confidence of applying equitable teaching practices on surveys but cite no direct pedagogical evidence or observation data that demonstrates such application.

Finally, Metevier et al. (2010) examined how educational development programs shifted 2 different cohorts of STEM faculty members’ (n=98) equitable teaching beliefs and attitudes. These programs contained workshops that helped instructors weave equitable instructional strategies into the inquiry laboratory activities they design and teach. The authors developed a rubric to score and analyze faculty members’ pre- and post-survey workshop responses about equitable STEM teaching and assessment practices and found positive significant gains in the reported use of these practices after the workshops. The rubric was not validated, but it focused on equitable teaching competencies such as drawing on students’ interests, addressing learning differences, and creating a collaborative environment. In this study, faculty self-reported equitable teaching beliefs and the application of equitable teaching in written survey responses; however, no direct pedagogical evidence or observation data was provided. These two conference studies appear in proceedings for which it is difficult to determine the stringency of review, and likely represent preliminary work.

Looking across these studies, Parker et al. (2016), O’Leary et al. (2020), Harrison-Bernard et al. (2020), and Metevier et al. (2010) suggest that instructors can learn equitable teaching practices through experience, designing activities, and reflection. This approach to educational development reflects a more social constructivist perspective, where instructors develop their own understandings of pedagogy through observing student work and examples. Rooney et al. (2020) focus more on information transfer through mini-modules that instructors are expected to reproduce. All the studies I identified describe long-term, cohort-based educational development programs. Four of the studies mention that these programs positively affect instructor’s equitable teaching practices based on self-reported data. The studies suggest that these educational development approaches were effective, but that the field would benefit from more robust measures of instructor change and learning.

2.2.2 Effects on Student Outcomes

The five studies I examined above do not investigate the impact of educational development on student achievement using pre- and post-assessments.

2.2.3 Mechanisms for Adopting Equitable Teaching

The five studies I reviewed identified different mechanisms to explain the adoption of equitable teaching beliefs and/or practices. Parker et al. (2016) found that addressing instructors' prior beliefs and dispositions was key for helping shift instructors' beliefs in addressing microinequities. Parker et al. (2016) also highlighted that sustained educational development programs and discussions of microaggressions in a STEM context were key in shifting beliefs. O'Leary et al. (2020) mentioned that equitable beliefs and pedagogy are fostered when instructors reflect on their social identity, develop self-awareness, mind the privilege gap, and reduce implicit bias. O'Leary et al. (2020) found that this approach led to significant gains in instructors' self-reported equitable teaching beliefs. Similarly, Harrison-Bernard et al. (2020) stated that instructors adopt equitable pedagogy if they became more knowledgeable and attentive to diversity, equity, inclusion, and implicit bias. Rooney et al. (2020) also identified that exposure to equity topics in evidence-based mini-modules can increase faculty self-confidence in equitable teaching and mentoring practices. Metevier et al. (2010) found that equitable teaching can improve when participants learn about equity topics (e.g., growth mindset) and design and teach their own inquiry activities.

2.3 Theories of Pedagogical Change in Higher Education Settings

Given the existing empirical literature's lack of focus on the underlying mechanisms for pedagogical change, I turn to the learning sciences to better understand mechanisms for why instructors might adopt equitable teaching practices. A better understanding of this literature can help address what Amundsen and Wilson (2012) pose as a core question of the field, "What are the features of educational development that make it effective?" (p. 112). After reviewing the literature, my goal is to explore the mechanisms that help me understand the features that make educational development effective in supporting the adoption of equitable pedagogies. Moreover, Mulnix (2016), Henderson et al. (2012), and Henderson and Dancy (2011) discuss how educational developers need to address the prior knowledge and beliefs of instructors in order to promote greater adoption of evidence-based teaching practices. Smith (2020) adds that not addressing instructor beliefs and what motivates their reasoning will lead to failures in adoption. Smith (2020) and Mulnix (2016) state that solely providing faculty with articles and telling them about data-driven practices is not sufficient to shift instructor's practices.

There are different theories of pedagogical change that focus on different levels of analysis. I am interested in examining instructor-level changes within a professional development context. I focus in detail here on six broad fields of theory: Communities of Practice, Teacher-Centered Systemic Reform (TCSR), the Four Frames of Leadership, Neo-institutionalism, and Teacher Noticing. Communities of Practice, TCSR, and the four frames of leadership are particularly prevalent theories for explaining pedagogical change in the field of educational development (Reinholz et al., 2021). Kezar & Bernstein-Serra's (2020) influential work in the field has cited neo-institutional theory as a key organizational behavior framework for explaining institutional and pedagogical change. From the learning sciences literature, I review teacher noticing because it provides content (e.g., academic concepts) and non-content (e.g., equity barriers) specific mechanisms that can motivate pedagogical change (van Es & Hand, 2017).

2.3.1 Communities of Practice

Communities of Practice (CoP) have been used widely to support instructors to work together, which can motivate them to learn and adopt innovative pedagogies within a community context. Wenger (1998) discusses that a community of practice involves the following elements: mutual engagement, joint enterprise, a shared domain of interest, and shared repertoire. Mutual engagement entails the community's shared behaviors and interactions. Joint enterprise refers to "the goals or requirements for the practice, as defined and negotiated, informally, by members of the community" (Kisiel, p. 2009, 98). A shared domain of interest can refer to a common area or topic for community members. In a community, members can pursue the domain of interest, participate in activities together, support each other, and share information. Expertise in this domain can distinguish a member from others in the group. Lastly, shared repertoire includes "the resources that facilitate practice—tools, artifacts, definitions, and common experiences" (Kisiel, p. 2009, 98). These shared tools, definitions, and experiences can inform how members respond to problems. Like many scholars focusing on K-12 teacher education and professional development, Judith Warren Little (2002) has leveraged the CoP theoretical framework to examine "professional development at the levels of individual experience, professional community, and organizational capacity for teacher support and instructional improvement" (p. 918).

CoP has also been applied to instructor learning in the higher education context. Leveraging the CoP framework, Robert and Carlsen (2017) described case studies of four science professors' teaching practices and research at a large university. They took a sociocultural perspective by viewing faculty "not as isolated subjects, but rather as members of complex social settings with rich personal histories and connections" (p. 939). Citing Vygotsky (1978), Robert and Carlsen (2017) state that faculty "learn and develop through interactions embedded in social and cultural activities" (p. 939). These authors emphasize that these "activities take place within 'communities of practice' (Lave & Wenger, 1991) in which new members of communities 'legitimately and peripherally participate' in culturally valued activities with the purpose of preparing them for 'mature participation' (Rogoff, 1995)." (Robert & Carlsen, 2017, p 939). Within a CoP, Robert and Carlsen (2017) suggest that faculty can change their practices through what Rogoff (1995) calls "guided participation" and "participatory appropriation".

Communities of Practice have been widely applied and researched in higher education through Faculty Learning Communities (FLCs) (Baker, 1999; Cox, 2004; Dalrymple et al., 2017; Herman et al., 2015; Pelletreau et al., 2018; Reinholz et al., 2021; Tinnell et al., 2019). In the field of educational development, FLCs are "relatively small groups[s] that may include students, teachers, administrators, and others who have a clear sense of membership, common goals, and opportunity for extensive face-to-face interaction" (Baker, 1999, p. 99). Cox (2004) and the higher education community (Dalrymple et al., 2017; Herman et al., 2015; Pelletreau et al., 2018; Tinnell et al., 2019) have drawn on CoP to professionalize teaching and formalize teaching communities focused on improving pedagogy over extended periods of time. The work on FLCs is important because K-12 requires certification to professionalize teaching, and there are CoPs in teacher education programs.

In Cox's (2004) study, the FLC engages in an active, collaborative, yearlong program with a curriculum about enhancing teaching and learning. Within the FLCs, there are frequent seminars and activities that promote learning, development, the scholarship of teaching, and community building. It has been shown that FLCs "create connections for isolated teachers,

establish networks for those pursuing pedagogical issues, meet early-career faculty expectations for community, foster multidisciplinary curricula, and begin to bring community to higher education" (Cox, 2004, p. 1). In order to be effective, this sense of community needs to be established early in the term (Duffy & Jones, 1995). Cox (2004) provides evidence that FLCs improve faculty learning through feedback from surveys of 50 past FLC participants. In the FLCs, instructors were, in part, motivated to adopt active learning practices because they had opportunities to discuss and address student data, such as their classroom observations. Over 90% of participants reported an increase in each of the following five categories: an increase in students' ability to apply learned concepts to new problems and solutions, to ask intelligent and relevant questions, to become open to new ideas, to work collaboratively with others, to think independently, and to integrate and synthesize information (Cox, 2004). The study reported approaches that increased discussion, writing, technology, cooperative learning, active learning, and student-centered learning.

Gehrke and Kezar (2016) build on the notion of CoP to introduce the concept of Communities of Transformation (CoT) to think about pedagogical change. A CoT includes 1) a philosophy to rethink STEM education, 2) structures that promote social interaction among members, and 3) mentorship that supports leadership in the community. Gerke and Kezar (2017) studied how CoTs promoted institutional and departmental STEM reform, which included sharing pedagogical tools and resources.

Departmental Action Teams (DATs) build on FLCs by offering faculty a space for faculty, staff, students, and administrators to work together in a department to advance pedagogical and institutional change (Ngai et al., 2020). These stakeholders often elect a problem of practice to discuss and address. For example, DATs can bring faculty and stakeholders together to examine departmental curriculum and how the curriculum and its courses could better serve equity and inclusion goals that promote robust student learning and belonging. DATs can contain stakeholders from different departments and have been found to be effective for implementing pedagogical innovations in higher education (Reinholz et al., 2020).

2.3.2 Teacher Centered Systemic Reform (TCSR)

Teacher Centered Systemic Reform (TCSR), developed by Woodbury and Gess-Newsome (2002), has been used to support instructors in adopting new teaching practices by considering the structural and cultural context of teaching and learning (Birt et al., 2019). These approaches to reform have been applied in STEM disciplines (e.g., chemistry, physics, biology, mathematics, computer programming, engineering) at R-1 universities (Birt et al., 2019; Enderle et al., 2013; Ferrare, 2019; Stains et al., 2015). The TCSR is a multifaceted model that simultaneously takes into account the structural and cultural context of teaching, instructors' profile, and instructors' thinking to study the success or failure of classroom reform efforts. The TCSR model views instructors and their contexts as drivers for change (Lund & Stain, 2015; Reinholz et al., 2021). TCSR asks instructors to have students engage in authentic scientific practices such as problem-solving and constructing evidence-based explanations using inquiry and experimentation.

Using TCSR as a framework, Birt et al. (2019) examined three case studies of new undergraduate science instructors after they had completed a graduate-level introductory course on college science teaching. The authors identified different mechanisms that prevented and supported the development of new teaching practices. Birt et al. (2019) discuss how the departmental teaching culture and routines served as an obstacle to pedagogical change:

> The local and departmental teaching environment was the biggest obstacle. The new instructors felt they needed to act according to the norms and expectations of their departments, which impeded their teaching practice in some aspects.

Birt et al., (2019) also emphasize that instructor beliefs and agentive goals play a role in adopting new teaching practices and that these goals. With respect to these goals, they highlight reproductive and transformative agency as mechanisms for pedagogical change. They express how instructors expressed reproductive agency by adhering to curriculum that was given to them, but they also expressed transformative agency by making curriculum more relevant to students' lives and backgrounds (Birt et al., 2019, p. 792).

2.3.3 Four Frames

Higher education scholars have used Bolman and Deal's (2008) Four Frames to highlight how symbolic, structural, political, and human resources frames can influence the adoption of innovative, active learning teaching practices in STEM. The symbolic frame emphasizes that change can occur when a coherent mission and culture is visually valued and celebrated. The structural frame focuses attention on the organizational rules, policies, goals, technology, and environmental factors that could influence pedagogical change. The political frame highlights the need for advocacy for new change, where a power base for such change needs to be developed. An example could be administrators who profess change and find allies as well as opinion leaders to achieve both vertical and horizontal support within their institution. Finally, the human resources frame emphasizes how organizations support and develop people. An example is having communities of practice where collaborative problem-solving can occur. Although the four frames have often been applied in the context of school-level reform, many scholars in the field of educational development have applied it to support the professional development of instructors (Rämö et al., 2019; Reinholz & Apkarian, 2018; Reinholz, Matz, et al., 2019; Reinholz, Ngai, et al., 2019).

Reinholz et al. (2019) moves Bolman and Deal's frames from a meso-level to a more micro-meso level. Reinholz et al. (2019) describes how the symbolic frame can be invoked when universities and departments value and prioritize professional development programs that teach active learning, learning outcomes, and backward design (p. 1140). Within the context of professional development programs, Reinholz et al. (2019) discusses that the structural frame can be leveraged when the training has explicit goals of using new active learning practices that respond to student data (e.g., clickers, formative assessment). Regarding the political frame, Reinholz et al. (2019) talks about how elevating the status of instructors and educational practices can promote the adoption of novel pedagogies in professional development. One can invoke the human resources frame by bringing science teaching fellows, or experts in STEM teaching, into professional development efforts to support the learning and adoption (e.g., human capital development) of novel pedagogies (Reinholz et al., 2019).

2.3.4 Neo-Institutional Theory

Kezar and Bernstein-Serra (2020) have discussed how neo-institutional theory (NIT), an organizational behavior framework, can be used to explain barriers and drivers for adopting active learning practices. According to NIT, historically, embedded social norms can inhibit change but can also be addressed to promote change (DiMaggio & Powell, 1991). Kezar and

Bernstein-Serra (2020) use NIT, specifically the mechanism of institutional logics (Thornton et al. 2012), to explain culture change. Institutional logics are "made up of assumptions (tacit beliefs and values that drive behavior), values (stated priorities of campuses and organizations) and norms (collective assumptions around values)" (p. 358). These authors discuss that competing institutional logics, or beliefs, can undermine change because there are administrators who advocate for change, but that research universities tend to prioritize research over teaching in tenure, review, and promotion processes for ladder-rank faculty. Furthermore, the authors discuss that devaluing instructor professional development from Centers for Teaching and Learning can thwart change. According to NIT, one has to disrupt the intersubjective understandings of how work is done (e.g., current norms and logics) by providing counter logics, or beliefs. This can be accomplished by highlighting active learning approaches that administrators, the field, and the department value. Consistent with Durkheim's beliefs (2004), NIT posits that instructors would adopt pedagogies if the pedagogies are valued and supported by the tribe, or members of the department. Mimicking these valued pedagogies would increase one's success in the tribe, sense of belonging, membership, and social integration. Consequently, instructors may be more motivated to use certain active learning practices and forms of student data if they align with the epistemologies and cultural values of the academic department, community, or field. Thus, Kezar and Bernstein-Serra (2020) describe that practitioners need to address the organizational and campus culture in order to support broader adoption.

2.3.5 Teacher Noticing

I am also inspired by the concept of teacher noticing, which can also be used as a framework for understanding how instructors adopt equitable teaching practices. Teacher noticing is grounded in the idea that instructional change can occur when instructors are reflective practitioners (Schön, 1987). For example, in higher education settings, McCoy and Bocala (2022) found that communities of practice can encourage instructors to reflect on equity and use equitable practices. These activities and extensive dialogue on pedagogy can promote reflection-in-action and reflection-on-action (Schön, 1987).

Sherin and Star (2011) discuss how teacher noticing is an "intuitive model of teacher reasoning" which involves the following process: "first, teachers attend to some classroom events; then, interpret, make sense of, or otherwise reason with the attended event; and finally, take some actions based on their reasoning (for a discussion; see Sherin and Star 2011, pp. 69–70)" (Scheiner, 2020, p. 87). In doing so, teacher noticing activates teachers' perception, cognition, and action. As such, there is a dynamic, interrelated, and cyclical interplay between "framing, perceptual activity and exploration, and interaction with the world of classroom events (modified from Neisser 1976, p. 21)" (Scheiner, 2020, p. 85-86). Thus, perception, interpretation, and decision-making are processes that mediate teachers' "transformation of dispositions into practice" (Scheiner, 2020, p. 87), which is critical for learning and adopting equitable pedagogies.

Focusing on K-12 mathematics classrooms, Sherin and Star (2011) discuss that "Teachers select and ignore on the basis of their sensemaking; the way they respond shapes subsequent instructional events, resulting in a new and varied set of experiences from which teachers attend and make sense" (p. 5). This suggests that "reality is not so much constructed, as it is filtered through an individual's understanding of the world, which is affected by her knowledge, values, and intentions. Such cognitive resources cause some aspects of the world to be more relevant and meaningful than others, thereby creating an understanding of the world with a particular bias.

Perception, as such, is always biased" (Scheiner, 2000, p. 87). Teacher noticing can therefore serve as a framework for understanding how teachers adopt equitable teaching practices, because the framework highlights how instructors' perspectives and values can serve as lenses that motivate pedagogical dispositions and decision-making.

van Es, Hand, and Mercado (2017) and van Es, Hand, Agarwal, and Sandoval (2022) extend the teacher noticing literature to focus on equity issues and equitable teaching practices in math for K-12 teachers. In the teacher noticing literature, expert teachers learn to notice and be more sensitive to differences in student learning.

2.3.6 Main Themes Across the Theories

I examine common themes across these frameworks for pedagogical change that I have reviewed. Below, I have identified common mechanisms for pedagogical change, such as valuing and acting on data, departmental and disciplinary cultures, instructor perspectives and definitions of equity, and instructor goals focused on equity and/or content.

Valuing and acting on data. Scholars who have leveraged TCSR, Communities of Practice, the four frames, neo-institutional theory, and teacher noticing would agree that data could motivate teacher learning. All these theories would agree that instructors should look at data to make pedagogical changes. Teacher noticing focuses on observational data. TCSR, Communities of Practice, the four frames, and neo-institutionalists would agree with focusing on observational, assessment data, and survey data if such data is valued by a group or organization.

Departmental and disciplinary cultures. TCSR, Communities of Practice, the four frames, and neo-institutionalists would agree that departmental and disciplinary cultures could motivate pedagogical change and the use of evidence-based teaching in higher education. Reinholz et al. (2019) discuss how departmental initiatives, such as department action teams, are important for initiating pedagogical change since they focus on the context and needs of the department. Due to the scope of my study, I do not focus on departmental initiatives. Instead, I aim to contribute a greater understanding of how departmental and disciplinary cultures can impact instructors' micro-interactions with students and their use of equitable teaching practices.

Instructor perspectives and definitions of equity. Scholars who have leveraged TCSR, Communities of Practice, the four frames, neo-institutional theory, and teacher noticing would agree that instructors' perspectives on equity could motivate teacher learning. TCSR asks instructors to reflect on equity issues that have systematically marginalized students. McCoy and Bocala (2022) have highlighted how faculty learning communities have focused on how faculty can reflect on equity issues to increase the use of equitable teaching practices.

Instructor goals focused on equity and/or content. Scholars who have leveraged TCSR, Communities of Practice, the four frames, neo-institutional theory, and teacher noticing would agree that instructors' curricular content and equity goals could motivate teacher learning. TCSR and Communities of Practice scholars have examined how instructors reflect on their content and equity goals to promote the use of more equitable teaching strategies. Additionally, a neo-institutionalist and a scholar using the four frames would argue that the content and equity goals of an instructor could motivate their pedagogical change if these goals are aligned with the institutional logics, values, and beliefs of the instructors' organizational culture. van Es and Hand (2017) would agree that content-specific and non-content specific noticing focused on equity is important for promoting a more equitable classroom.

2.4 Limitations in the Existing Research

There are limitations in terms of evaluation and underlying theories that motivate pedagogical change. Understanding these limitations is important for advancing the field of educational development and pedagogical change in higher education.

In terms of evaluation, beyond self-report, none of the five empirical studies identified in this systematic review report observation data or other direct evidence for how each instructor *applied* equitable teaching practices. Ebert-May (2011) highlights how there can be a major difference between self-report and what the instructor practices. These studies also do not focus on learning progressions or other measures to better understand *processes* of instructors' development or adoption. Instead, the existing studies focus on knowledge, awareness, beliefs, and/or perceptions of equitable teaching practices.

Additionally, the 5 empirical articles had a limited discussion of theory for why instructors adopted equitable teaching beliefs and/or practices. In fact, these articles do not focus very much on the theoretical mechanisms of how and why educational development interventions can support adoption of specific equitable teaching beliefs and practices. Therefore, I would characterize these studies as "implementation studies" (Fowler, 2021). As such, these articles take a practical perspective since they are tailored towards practitioner audiences. Moreover, Levinson-Rose and Menges (1981), authors of an early empirical review of college pedagogy, provided the critique that research and theory on improving college teaching is not well-developed. Similarly, the field of learning sciences is still building theories of how college instructors learn.

To address what Fowler et al. (2021) call the "theory crisis" and Levin-Rose and Menges's (1981) concerns, I am interested in going beyond implementation studies. In my research, I am interested in understanding the more micro-level processes (e.g., values and beliefs around data, the organization, the responsibility for learning) of how instructor professional development interventions can influence adoption of equitable teaching. I plan to examine instructors' deficit thinking, values, beliefs, and resistance to equitable teaching as key mechanisms that can impact adoption.

In terms of the specific interventions used during educational development programs, the studies mentioned above do not examine how faculty developers explicitly adapt the pedagogical training based on instructors' learning needs, perspectives, and values. I argue that addressing instructors' needs, perspectives, and values is critical for shifting their equitable teaching and practices. As described in Chapter 1, my study examines how facilitators adapt professional development learning experiences based on instructors' reflections that capture their learning needs, perspectives, and values. One of my main goals is to design educational development that increases equitable teaching competencies through adaptive professional development. I evaluate whether this specific approach can change teaching practices in ways that directly impact student outcomes such as achievement.

In addition to the literature on educational development interventions, I also identified limitations with the theories of pedagogical change that I reviewed above. I noticed that these theories do not explicitly consider how pedagogical change is motivated by emotion and where instructors place the responsibility for learning. Therefore, I plan to explore how the adoption of equitable pedagogies can be motivated by emotion and where instructors place the responsibility for learning. In particular, I am interested in the role of positive and negative emotions that impact how instructors learn and adapt equitable pedagogies. I also hope to examine whether or

not instructors who share the responsibility for learning with their students are more likely to learn and adapt their use of equitable practices, such as AEP.

Moreover, the teacher noticing literature often focuses more on observations and student behaviors. One drawback is that this literature focuses on unaided observations, which can overlook student equity barriers and instructors' perspectives that may not be immediately visible. To augment instructors' unaided observations, I hope to examine how instructors notice equity barriers through other forms of student learning data (e.g., surveys, formative assessment responses).

Moving beyond K-12, I seek to add an equity lens to the teacher noticing literature that examines higher education STEM courses. My study focuses more on the specific equity barriers that STEM students face at the college level. I hope to explore how college STEM instructors can use data to inform their decisions and how data can be leveraged as a mechanism for supporting the adoption of equitable teaching practices. In my study, I am interested in how teachers perceive their classrooms, learn about their students, and act based on student data (e.g., formative assessment, student survey responses) (Scheiner, 2020).

2.5 Conclusion

Equitable teaching for instructors entails clarifying expectations, aligning instruction with these expectations, grading for equity, expressing a growth mindset, and being responsive to students' backgrounds, identities, and existing competencies. A small sample of papers in the literature including five empirical studies show that educational development programs can positively impact instructors' equitable beliefs and practices, as measured by self-reported survey responses (RQ1). However, none of these studies examined the impact of these programs on student achievement, and most of the studies did not explicitly offer mechanisms for explaining pedagogical change and adoption (RQ2).

To identify and explore such mechanisms in the literature, I reviewed theories of pedagogical change which included Communities of Practice, Teacher-Centered Systemic Reform, the Four Frames of Leadership, Neo-institutionalism, and Teacher Noticing. These theories emerge from the fields of educational development, learning sciences, and organizational behavior. Among these theories, common mechanisms for motivating pedagogical change include valuing and acting on data, departmental and disciplinary cultures, instructor perspectives and definitions of equity, and instructor goals focused on equity and/or content. In the next chapter of my dissertation, I discuss how theories of change inform my analytical approach to analyzing why instructors adopt or do not adopt equity-oriented pedagogies in STEM.

CHAPTER 3: Definition of AEP Competency and Conceptual Framework

As described in chapter 2, much is known about college-level pedagogy, and there is less empirical research about how to support the adoption of instructors' equitable pedagogies. In chapter 3, I describe the adaptive equity-oriented pedagogical (AEP) framework and explain how it can be applied to supporting instructors in learning equitable teaching.

In service of this goal, I have 3 overarching objectives for this chapter:

1. Provide a definition of AEP competency.
2. Explain how AEP, as a conceptual framework, can be applied to professional development in the design of a pedagogy course intervention.
3. Articulate different cultural lenses, or perspectives, that can help me understand whether instructors adopt or do not adopt AEP.

In particular, I define AEP competency elements in order to provide context for my conceptual framework that I will use to analyze how instructors learn AEP. Incorporating learning sciences theories into educational development research, my conceptual framework hypothesizes that the treatment promotes greater reflection, which I expect to increase instructor learning in the treatment relative to the control as measured by AEPC assessments defined below.

3.1 Adaptive Equity-Oriented Pedagogical Competency Definition

In this study, I examine treatment and control pedagogy courses that teach instructors about the key AEP principles. AEP is a framework for adapting teaching to address equity barriers to learning based on student data (e.g., formative assessment, observations, surveys) (Phuong et al., 2017b; Phuong et al., 2021a). Effective AEP practice is characterized by instructors' use of the competency elements listed below, which build on each other and reflect a mastery learning model. Consequently, the AEP pedagogy courses focus on equipping instructors with AEP competency elements that enable them to:

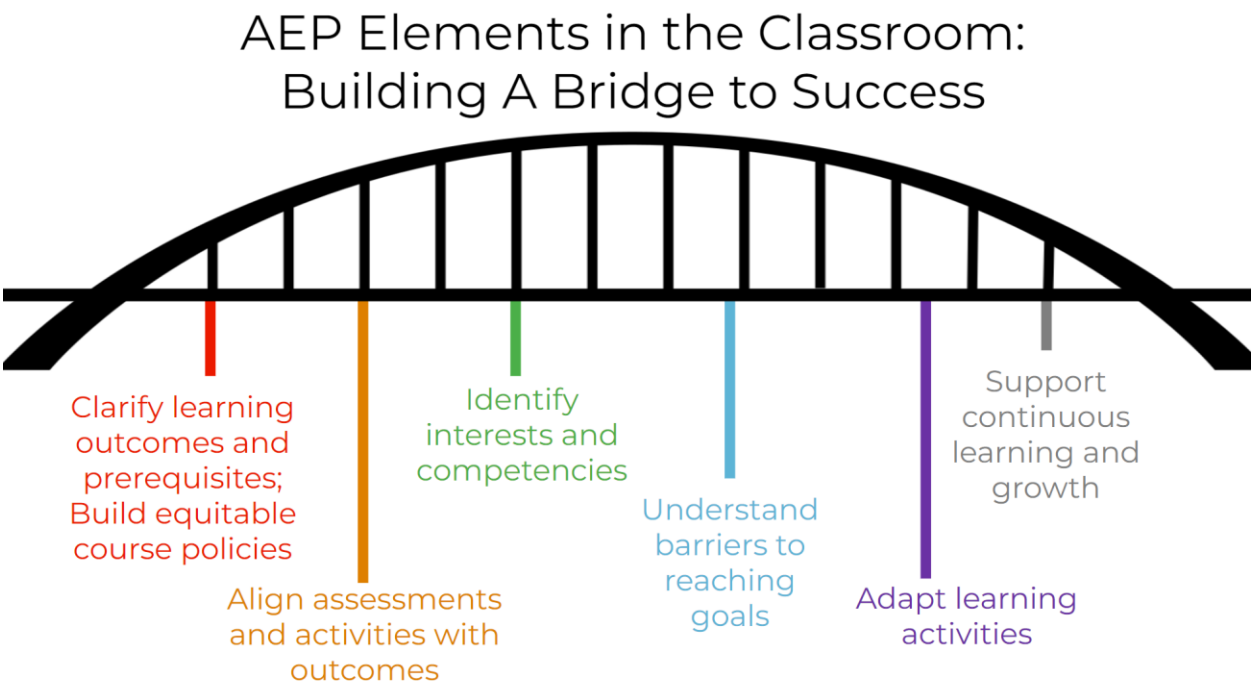
1. **Clarify** learning outcomes, prerequisite knowledge, and equitable course policies (e.g., demonstrate cultural humility by grading for equity, offering flexibility with deadlines, and adjusting policies based on student needs and contextual circumstances)
2. **Align** formative assessments and teaching activities (e.g., modeling key skills, fostering collaborative learning, deliberate practice) with outcomes.
3. **Identify** students' competencies, interests, and needs
4. **Understand** students' equity barriers and contextual challenges (e.g., imposter phenomenon, stereotype threat, basic needs access) to meeting outcomes
5. **Adapt** teaching activities based on students' needs and barriers
6. **Iterate**: Reflect upon pedagogy to support continuous learning, adaptation, and growth

Elements 1-2 typically before and during instruction. Instructors often apply elements 3-5 when they teach a course. Element 6 usually occurs throughout and after an instructor teaches a class.

Figure 3.1 presents an illustration of the AEP competency elements.

Figure 3. 1

Adaptive Equity-Oriented Pedagogical Competency Elements



3.1.1 Applying AEP in Educational Development

Several of the AEP elements draw from established research on equitable teaching practices presented in chapter 2. In addition, research has already identified reflection, collecting data on pedagogical innovations, and discussing harmful psychosocial effects as effective elements of educational development (Harrison-Bernard et al., 2020; Metevier et al., 2010; O'Leary et al., 2020; Parker et al., 2016; Rooney et al., 2020). However, current educational development efforts thus far lack a focus on adapting learning experiences based on instructors' perspectives through written reflections.

The AEP framework offers educational developers a strategy to adapt professional development based on instructors' reflections (Phuong, 2021a). Applying the AEP elements above, educational developers can apply the following when designing learning experiences:

1) Work in partnership with instructors by gathering reflection and formative assessment data to understand their perspectives, experiences, motivations, strengths, and barriers to learning;
2) Leverage these insights to adapt learning experiences in ways that increase instructors' engagement and success (Phuong, 2021b).

3.2 Design of Pedagogy Course Interventions

Based on the literature in Chapter 2 on communities of practice and teacher noticing, I designed pedagogy course interventions that draw on the AEP conceptual framework. Accordingly, the pedagogy course facilitators apply AEP to their practice to promote greater

reflection and instructor learning of equitable pedagogies. Within a community of practice, the pedagogy course facilitators apply teacher noticing to learn more about instructors' lenses (i.e., their values, perspectives, and biases) by observing them and reviewing their written reflections. By adapting to these lenses, facilitators can address instructors' perspectives and model AEP practices for instructors, so they can apply these practices with their students. By engaging in teacher noticing, facilitators can support instructors in becoming more reflective in examining and addressing equity issues.

Consequently, I designed the pedagogy courses as communities of practice to incorporate a sociocultural perspective of learning (Vygotsky, 1978; Wenger, 1998). When learning about equitable teaching in a CoP, I imagine that instructors reflect on their prior learning experiences and interactions with their students and peers to learn equitable pedagogies. Through this reflection, instructors can interact with students more equitably. Drawing on a sociocultural perspective on learning, both pedagogy courses used small group discussion, collaborative technology, and reciprocal teaching to raise, examine, and reflect on issues around equity, student success, instructor experiences, and STEM cultural norms.

I use Table 3.1 below to describe the treatment and control conditions in the study.

Table 3. 1

Comparison of Control and Treatment Pedagogy Courses

Pedagogy course activity	Control	Treatment
Same AEP curriculum content, reflection prompts, assignments/ assessments, and opportunities to redo assignments	✓	✓
Same planning activities: examine course syllabi, learning outcomes, student data	✓	✓
Facilitator provides dialogue and activities based on instructors' expressed interests (e.g., conversation about pedagogical challenges/ needs) for 30-40 minutes each week	✓	✓
Instructors write weekly reflections on teaching	✓	✓
Instructors discuss reflections with each other	✓	✓
Using weekly instructor reflection data, pedagogy course facilitators adapt activities to instructors' learning needs, thereby modeling and applying 6 AEP elements explicitly		✓

Shared element of treatment and control: Reflection as a tool for instructor learning. Both conditions shared guiding principles around reflection (see Table 3.1 above) and sought to create a reflective community of instructors who learned about AEP. I examine how instructors use reflection and data artifacts as tools to think about, learn about, and interact with their students. By engaging in teacher noticing (van Es & Hand, 2017), instructors can become reflective practitioners (Schön, 1987) by working with one another and learning through dialogue and engagement with continuous learning and problem-solving. Tools like reflection can support instructors and facilitators in understanding a) pedagogical goals and challenges and b) ways in which instructors can design next steps to address these problems. These activities and extensive dialogue on pedagogy promote reflection-in-action and reflection-on-action (Schön, 1987). Throughout this process, instructors can engage in content-specific and non-content specific noticing which can inform their pedagogical decisions.

Understanding what behaviors and attitudes are normalized through reflections can enable facilitators to create more targeted professional development that promotes the unlearning of harmful practices and a greater learning of equitable practices. In particular, I argue that guided reflection can be beneficial for facilitators when it elicits the following from college instructors:

- Reflection on values, experiences, learning history, routines, motivations, resistance to equity.
- Reflection on teaching goals, what works, and what needs to be done next.

It is important to collect reflection data to understand a) how instructors perceive the organizations' values and b) how their prior experiences impact the ways they think about teaching and learning.

As instructors reflect and engage with teaching, they will invoke different lenses or perspectives in their written reflections. In this study, the facilitators in the treatment seek to understand instructors' different cultural lenses, or perspectives, by reviewing reflections and adapting to instructors' needs using the AEP framework. This approach enables the facilitators to understand instructors' values and perspectives and how these may or may not motivate pedagogical change.

Like Duckor and Holmberg (2019), I examine changes in instructors' competencies. This study, however, focuses on equity-oriented teaching competencies. This study's examination of instructors' developmental progression is grounded in Vygotsky's (1978) work, which posits that learning is inherently sociocultural. I studied instructors' written and multimodal artifacts (e.g., teaching philosophies, evidence of AEP competencies in teaching portfolios, student gain score visualizations) to examine their AEP competency over the course of a semester. The focus on a developmental learning progression (Duncan & Rivet, 2013; Wilson, 2009) of equitable pedagogies enables me to examine instructors' growth trajectories of AEP competencies (AEPC) in a sociocultural context. In Table 3.2, I offer a summary of the AEPC learning progression that advances reflective and equitable teaching with supporting literature. This learning progression serves as the AEPC construct map under Wilson's (2005) four building blocks for constructing measurement instruments. The construct map is the foundation for the AEPC assessment. A more detailed construct map can be found in Appendix B.

Table 3. 2

AEPC Learning Progression

Learning Progression Level	AEP Elements Used	Description
Level 0: No AEPC Mastery	0	0 AEP elements used, the instructor does not consider learner needs (McCallum, 2013). This level of competency reflects a lack of consideration of the needs and perspectives of the learner.
Level 1: Low AEPC Mastery	1-2	The instructor uses AEP elements 1-2: 1. Clarify learning outcomes and pre-requisite knowledge

		2. Align formative assessments and activities (e.g., clickers, low-stakes assignments and quizzes, discussion prompts) with outcomes Fink (2013), Davis (2009), and Weinstein (2009) highlight the importance of clarifying and supporting students in meeting rigorous learning outcomes. Clarifying expectations and aligning *both* instruction and assessment with outcomes is foundational to helping students equitably navigate course curricula. The lack of alignment engenders equity issues in STEM since it often privileges students who have more background knowledge or those who have more access to financial resources to support their success. Designing for equity requires one to build and clarify learning outcomes and course policies (e.g., grading, assessment, accommodations) that are informed by the equity literature (Feldman, 2018; Phuong et al., 2017). For example, instructors can provide flexibility with deadlines (e.g., extensions), offer opportunities for re-grading, avoid curving down, and adjust policies based on students' contextual circumstances. These policies and practices can foster a growth mindset within a brave space for learning where students can take chances, take risks, make mistakes, and grow in an environment that is focused on mastery learning (Phuong et al., 2017; Phuong et al., 2022).
Level 2: Moderate AEPC Mastery	1-4	The instructor uses AEP elements 1-4. Elements 1 and 2 mentioned above 3. Identify students' competencies, interests, and needs 4. Understand equity barriers/ contextual challenges to meeting outcomes Reaching this level requires that level 1 be satisfied as a prerequisite. This level represents moderate mastery because research shows it is beneficial to identify students' existing competencies, equity barriers (e.g., imposter syndrome, stereotype threat, no internet connection), and learning needs via formative assessments that are aligned with learning outcomes (McCallum, 2013; Phuong et al., 2017).

Level 3: High AEPC Mastery	1-5	The instructor uses AEP elements 1-5. Elements 1-4 mentioned above 5. Adapt teaching practices based on these needs and barriers Drawing on diagnostic data aligned with course learning outcomes, the instructor adjusts instruction, which includes how they engage students in productive struggle, model key skills, provide deliberate practice opportunities, and offer feedback (Phuong et al., 2017). Level 3 represents a higher degree of AEPC because adapting instruction in a manner consistent with AEP requires the prior levels. Adapting instruction without this data-driven strategy can lead to misguided instruction and assessment practices that do not address key learning barriers (Phuong et al., 2021)
Level 4: Very High AEPC Mastery	1-6	The instructor uses AEP elements 1-6. Elements 1-5 mentioned above 6. Iterate: Reflect upon pedagogy to support continuous learning, adaptation, and growth The instructor demonstrates a commitment to adaptation, iteration, and growth with students, improving the application of previous elements recursively. This level suggests that the instructor can recognize breakdowns in their own teaching practices and are flexible enough to employ and adapt culturally responsive and universal design approaches (e.g., video, multimedia) to further their students' success (Hammond, 2014; CAST, 2018). In their reflections, the instructors describe rationales for improving teaching in ways that are grounded in multiple forms of student learning data. This level is the highest because in order to address breakdowns in teaching and learning, instructors need to have informed experiences associated with levels 1-3 (i.e., AEP elements 1-5). Demonstrating a high-level of mastery suggests the instructor continuously adapt practices and policies. It is noteworthy that elevating ongoing, critical reflection (Kendi, 2019), perspective taking, and adaptation to reduce equity barriers advances anti-racist pedagogical practice.

Unique element of treatment pedagogy course: Facilitators modeled and applied the 6 AEP elements explicitly as a tool for instructor learning by using weekly instructor reflection data to continuously adjust pedagogy course discussion and activities. Specifically, the facilitators reviewed instructor reflections to

- Understand instructors' cultural values, degree of deficit thinking, experiences with equity and inequity, motivations, learning history, and what needs to be learned and unlearned. The cultural values can include instructors' and their organizations' values, attitudes, and beliefs (DiMaggio & Powell, 1991).
- Adjust instructor learning experiences to expand their critical consciousness around student learning so that they can interact with students in more culturally responsive ways.

Applying AEP, the facilitators adjusted the treatment based on instructors' reflections on their cultural backgrounds, which include their lived experiences, racial experiences, discourse, and their STEM department culture (e.g., what is valued methodologically). Both conditions prompted instructors to reflect and discuss how they may have experienced stereotype threat and microaggressions based on their intersectional identities (e.g., race, gender, disability). Facilitators in the treatment adjusted their practices by following-up with instructors in subsequent lessons to have instructors *think more critically* about how institutional and psychological barriers impact students from various intersecting identities, and how to adapt their teaching to address these barriers. Facilitators aimed to practice the same critical reflection and equity-oriented practices that they asked instructors to learn and adopt. These AEP strategies were used to further instructors' learning of equitable pedagogies, such as AEP.

To broaden instructors' critical consciousness in the treatment, facilitators used culturally responsive learning activities to enable instructors to expand their interpretation of culturally and linguistically diverse students' learning, reflect on and problematize wider systemic and institutional barriers to learning, recognize triggers around race and culture, be aware of cultural stereotypes and archetypes of individualism and collectivism, examine their biases and limitations, and promote a sense of belonging (Hammond, 2014).

Guiding conceptual framework. My conceptual framework reflects the equitable practices that I believe promote instructor learning. In particular, **AEP adapts culturally responsive pedagogy (CRP) and is modeled in treatment.** This process promotes greater reflection in the treatment, which I expect to increase instructor learning in the treatment relative to the control as measured by AEPC assessments.

Specifically under this framework, in order for learning to be effective, facilitators need to clarify outcomes, assess learner needs, plan instruction, and then adjust based on their own reflection and instructors' learning needs, patterns of behaviors, values, cultural backgrounds, and strengths. By reviewing written reflections, facilitators can understand instructors' perspectives, or lenses, to adapt learning experiences that shift instructors' perspectives and practices.

I leverage this conceptual framework to understand how pedagogy course facilitators adapt their practices to address college instructors' learning needs in the treatment. AEP builds on culturally responsive pedagogy (CRP) by offering a framework to adapt CRP practices and ways to expand instructors' critical consciousness. For context, Ladson-Billings (1995) and Hammond (2014) discuss how CRP involves engaging learners' experiences, performance styles, critical consciousness, and cultural knowledge to make curricula more relevant and effective.

Using an experimental design, I investigate how AEP can be used to offer more targeted culturally responsive professional development based on instructors' weekly reflections.
It is noteworthy that elevating ongoing, critical reflection, perspective taking, and adaptation to reduce equity barriers advances anti-racist pedagogical practice. These are anti-racist competencies that can help instructors understand learners' perspectives, equity barriers, and learning data so that instructors can re-examine their pedagogical decisions, biases, and worldview. Taking a critical and anti-racist lens, the AEP framework is aware that racialization is a part of society, pedagogy, assessment, our educational system, and social fabric (Kendi, 2019). When instructors are initially applying AEP, there may be pedagogical policies, designs, and decisions that privilege certain learners and are racialized even if unintentionally. Acknowledging that there is not a single story and learning pathway for students, the AEP framework therefore focuses on leveraging student learning data and ongoing adaptation to mitigate the impacts that systemic barriers and racialization have on students. Engaging in deeper reflection is key for continuously improving and adapting one's mindset and pedagogy, which is an important step for advancing student success. Such reflection can be further promoted when data help instructors notice and address students' learning needs and equity barriers.

3.3 Lenses on Instructional Change

Further, I am interested in the reasons that drive instructors to learn and implement pedagogical changes. I am inspired by the notion of professional vision (Goodwin, 1994), which refers to developing ways of seeing phenomena that reflect growing levels of expertise. I focus more specifically on *lenses for pedagogical change*, which include the following: instructors' values around the types of data sources used for instructional decision-making, beliefs and goals around equity, the emotions that drive their behaviors, the kinds of content they prioritize, where they place the responsibility for learning, and how departmental and disciplinary cultures impact their pedagogical choices. In this way, these lenses, or mechanisms, allow me to address content-specific and non-content specific noticing skills (e.g., stereotype threat, imposter syndrome). I build on the teacher noticing literature to examine how instructors adapt teaching based on more than unaided classroom observations, but also on information such as written reflections and data on student equity barriers. I define 6 lenses below:

- Authority and Community Lens,
- Learning Goal Lens
- Equity Definition and Perspectives Lens
- Data Lens
- Emotional Lens
- Responsibility for Learning Lens

Authority and Community Lens. The Authority and Community Lens is defined as the knowledge valued by the authoritative figures in the field. Here are questions that one may consider when reflecting on this lens: Do authoritative figures, peers, the scholarly field, the industry, and those I'm serving value the AEP practice? Is there a community of practice around AEP? (Phuong et al., 2021a).

The Authority and Community Lens is important because it can explain why instructors want to be a successful member of a tribe or community they are a part of or want to join. To achieve this goal, instructors look outward to the organization to identify what practices are valued by the tribe. Consistent with neo-institutional theory, Durkheim (2004) highlights that individuals focus on social integration where they pay attention to information about

membership, belonging, and the extrinsic and intrinsic rewards that help them succeed and move forward in the tribe. Therefore, Durkheim (2004) would posit that instructors would have a need to feel valued, fulfilled, supported, and connected to their organization or community, in order to thrive and feel like they belong (Marshall, 2002).

The Authority and Community Lens is relevant to my context, because Communities of Practice (CoPs), Communities of Transformation (CoTs), and Departmental Action Teams (DAT) can offer a space where instructors can receive scaffolding to discuss and reflect on problems of practice and ways to address them. COPs provide a network for instructors to share ideas and expertise so that they can develop equitable practices. I particularly see CoPs and CoTs as a beneficial way for instructors to discuss, interrogate, and re-examine norms, values, beliefs, and practices within an organizational context, such as an academic STEM department. The STEM department can contain authoritative figures (e.g., faculty, deans) and peers who reinforce culturally valued behaviors and practices that instructors might mimic.

Through the Authority and Community Lens, I hope to build on the teacher noticing literature by investigating how instructors are influenced by departmental and disciplinary cultures as they learn to notice equity issues and act on this information. For example, in STEM disciplines, instructors may value making decisions based on quantitative data more than qualitative data due to departmental and disciplinary cultural norms and epistemologies.

The Authority and Community Lens is also inspired by the four frames of leadership (Bolman & Deal, 2008). Leveraging the "symbolic frame" of leadership, professional development programs and efforts can be symbolic gestures that can motivate instructors to sustain pedagogical changes because their environment is signaling the importance of making these changes. Professional development efforts like training and reflection assignments about teaching can provide structures that support the adoption of new practices (Reinholz et al., 2019). Instructors expressing these perspectives in their reflections would be invoking the Authority and Community Lens. I hope to examine if instructors who frequently invoke this lens would be more motivated to adopt equitable teaching practices.

Learning Goal Lens. The Learning Goal Lens is defined as the goals, concepts, or ideas that one foregrounds in their mind. These goals, concepts, or ideas may include a focus on academic content and/or equity issues. This lens speaks to the purpose and beliefs that instructors have for their teaching. For example, instructors may focus on learning outcomes that emphasize academic concepts. Instructors may also have goals focused on addressing equity, such as addressing stereotype threat and improving student sense of belonging in the classroom. In the teacher noticing literature, van Es and Hand (2017) would argue that content-specific (e.g., academic concepts) and non-content specific noticing focused on equity (e.g., stereotype threat) is important for promoting a more inclusive classroom. Building on this work, I am interested in examining whether high-growth instructors trained in AEP competency have more equity-driven and academic content-driven goals than low-growth instructors.

Here are questions that one may consider when reflecting on this lens: Is the AEP practice or knowledge relevant to my teaching, goals, and priorities that I am foregrounding in my mind? Are these AEP practices relevant to the course content and concepts that my organization and I prioritize? Do I prioritize addressing equity issues and students' success with learning academic content?

Equity Perspectives and Definitions Lens. The equity perspectives and definitions lens posits that instructors' definitions of equity can motivate them to adopt or not adopt AEP. One definition of equity may include providing all students with the same learning opportunities.

Another definition may focus on adapting learning opportunities based on students' learning needs. Building on TCSR and teacher noticing, I hope to examine how pedagogy courses can utilize reflection and experiential learning to address instructors' perspectives on equity and inclusion. I am interested in how instructors' perspectives and definitions of equity motivate their use of equitable pedagogies.

Here are questions that one may consider when reflecting on this lens: How do I define equity? Is equity responding to all students' needs or providing everyone with the same thing? What is the connection between equity and student success? Is equity only about academic achievement?

Data Lens. The Data Lens is defined as the kinds of knowledge and data that are important and relevant in an academic and socio-organizational culture (Phuong et al., 2021a). Here is a question that one may consider when reflecting on the Data Lens: Do I value certain kinds of practices because they are tested using data sources and methods that are important to me or my field (Hammer, 2005)? The Data Lens considers the sources from which instructors can make determinations about student competencies. An instructor can use the Data Lens to understand student perspectives and learning processes. This lens can be useful for diagnosing student needs and adapting teaching.

All the pedagogical change theories in my literature review highlighted that data is an important mechanism for promoting instructor pedagogical change. In particular, Birt et al. (2019), Reinholz et al. (2019), and Kezar and Bernstein-Serra (2020) describe how organizational culture plays a role in the adoption of active learning pedagogies. In thinking about organizational culture, I plan to obtain a better understanding of how disciplinary epistemological traditions impact how instructors learn and adopt equitable pedagogies. Therefore, in my research, I argue it would be important to focus on how instructors value and use data to motivate pedagogical decision-making. For example, I am interested in whether and how instructors value and use different data sources (e.g., observations, clickers, formative assessments, surveys, summative assessments) to understand and address equity barriers and patterns in student learning. Thus, the Data Lens can help instructors notice more aspects of their students' perspectives and less visible equity barriers. I hope to examine whether the use of this lens motivates the adoption of equitable pedagogies.

Emotional Lens. The Emotional Lens is defined as the emotions that arise from engaging in a practice, experience, or reflection. I contend that the norms of a socio-organizational culture impact how individuals think they should feel and how they should express their feelings. From organizational behavior theory, natural systems thinking perspectives would support this lens by stating that instructors' intrinsic motivations, experiences, and emotions can motivate behavior (Scott, 2006). Here are questions that one may consider when reflecting on this lens: Does the data from implementing these practices help me feel more valued, fulfilled (i.e., energized), successful, and connected to my organization and those I'm serving? Are these results and feelings worth my time and effort? Is the teaching experience energizing me or burning me out? The Emotional Lens can explain motivations for adapting teaching (Phuong et al., 2021a).

Responsibility for Learning Lens. The Responsibility for Learning Lens refers to how the instructor views their role and the student's role in learning-- whose responsibility is it if a student doesn't learn a concept? Are the "good" students the ones who teach themselves or don't need help? (Phuong et al., 2022). This lens contributes to scholarship on reflective college teaching since it accounts for the systems, structures, and/or beliefs that foreground

individualistic notions of learning that have contributed to STEM students leaving the major because they felt under-supported and isolated in a competitive environment (Jackson & Cobb, 2010; Phuong et al., 2022; Seymour & Hewitt, 1997). This issue can be intensified for URM students who have less support from home.

The Responsibility for Learning Lens would account for the logics that can support and prevent adoption of equitable pedagogies. Identifying these logics in reflections would be beneficial in professional development programs to understand instructors' beliefs, values, and the ways they make sense of the world. For example, instructors may hold a faulty logic that the "good students" are the ones who teach themselves, do not need help, and/or do not ask for help. This kind of logic or belief could inhibit the adoption of equitable teaching practices because instructors may not believe that their role is to be responsive to students' learning needs because students are supposed to teach themselves. Curating learning experiences that enable instructors to interrogate these beliefs is important for shifting instructors' logics and promoting pedagogical changes. I hope to introduce a logic on how instructors can collect student data to design, iterate their pedagogies, and become more responsive to their students' learning needs. I hope to examine the extent to which the mechanisms above can explain the differences between low- and high-growth instructors on AEP competency.

Identifying these lenses in reflections is beneficial for supporting instructors' learning of equitable pedagogies. Facilitators can focus on expanding instructors' critical consciousness so that they recognize the importance of using multiple forms of data and sharing the learning responsibility with their students. Using multiple forms of data and considering various factors that impact learning can enable instructors to have a more holistic view of their students, which is important for adapting instruction to respond to student needs.

3.4 Bringing a Learning Sciences Perspective to the Field of Educational Development

Through my dissertation, I contribute a Learning Sciences approach to the field of educational development. I focus on mechanisms of learning, sensitivity to culture and context, and methodological depth at the level of how instructors learn and adopt inclusive pedagogies, such as AEP. The educational development research that I reviewed in chapter 2 does not focus on adapting to instructors' cultural perspectives of the world. In this study, the facilitators in the treatment engage instructors' different lenses, or perspectives, by reviewing their reflections and adapting to their needs using the AEP framework.

At the same time, I am extending the teacher noticing literature within the Learning Sciences. Within a community of practice, pedagogy course facilitators apply teacher noticing to learn more about instructors' values, perspectives, and biases by observing them and reviewing their written reflections. By adapting professional development practices based on this information, facilitators can address instructors' perspectives and model AEP practices for instructors, so they can apply these practices with their students. My research identifies lenses that instructors can invoke when they notice patterns in student data (e.g., observations, performance, experiences with equity issues). By contributing these lenses, I offer specific mechanisms to understand college STEM instructors' learning processes, values, beliefs, and motivations to adopt equitable pedagogies, such as AEP. For example, the Data Lens can help instructors notice more aspects of their students' perspectives and less visible equity barriers (e.g., stereotype threat, imposter syndrome).

I further explore how pedagogy course facilitators can adapt professional development interventions based on instructors' reflection data to improve instructor competency with

equitable teaching. In the teacher noticing literature, expert teachers learn to notice and be more sensitive to differences in student learning. To analyze teacher learning, I leverage written reflections to examine instructors' noticing and thinking over time based on evidence of student learning or lack thereof. As in the teacher noticing literature, I examine how instructors attend to and make sense of verbal, non-verbal, and written artifacts from students. In particular, I focus on improving how instructors can notice more as they collect data on student learning like one would with the scientific method. By being able to notice patterns in student learning and in their own biases and perspectives, I argue that instructors can more holistically understand their student's learning needs and equity barriers. By having this greater, holistic understanding and awareness, instructors can make more data-informed pedagogical adaptations that target students' needs. I anticipate that this process can improve learning of equitable pedagogies, where instructors interact more equitably with students. To answer research question 2, I hope to build on the teacher noticing literature by leveraging these lenses to explain high and low-growth instructors on AEP competency. Through my dissertation, I contribute a conceptual framework on how AEP can guide facilitators in adapting professional development practices based on instructors' lenses that are identified through written reflections. This approach promotes a model of teacher learning that focuses on increasing competency with equitable pedagogies, such as AEP.

Within a higher education STEM context, my conceptual framework prioritizes equity in the teacher noticing process, which is grounded in a sociocultural perspective of learning. In particular, the AEP framework adds an equity perspective to teacher noticing as it seeks to create a culturally responsive community of reflective practitioners. The AEP framework further contributes to the teacher noticing literature, because it provides a framework with different pedagogical practices that encourage instructors to reflect and respond to their students' socio-cultural context and equity barriers. Phuong et al. (2017) argue that taking students' perspectives is important for promoting student success for all students. Accordingly, part of being an equitable instructor entails using data to reflect on one's own pedagogical frameworks and student perspectives in order to challenge unconscious biases and design for all students.

CHAPTER 4: Methodology

In chapter 4, I describe my quantitative and qualitative research methodology. My methodology was designed to address some of the limitations in the current educational development literature. As described in chapter 2, the current literature has identified that educational development programs can improve instructors' adoption of equitable beliefs and practices based on self-reported surveys. However, since these studies contained no control group or randomization, these improvements could have been due to maturation and a practice effect. In addition, none of the five empirical studies mentioned in chapter 2 controlled for any instructor-level variables when determining the effect of educational development programs. Below, I discuss how my mixed-methods research design builds on the existing literature and addresses some of the limitations.

To test the efficacy of the conceptual framework that I present in chapter 3, I examine if it is worthwhile to invest additional time and energy into applying AEP to educational development programs. To achieve this goal and uphold research ethics, I therefore created an experimental control group that provides exemplary educational development activities such as reflection and dialogue because these practices have been shown to improve instructors' adoption of equitable teaching practices (Harrison-Bernard et al., 2020; Metevier et al., 2010; O'Leary et al., 2020; Parker et al., 2016; Rooney et al., 2020). I designed the control condition to be effective and to have the same number of class sessions as the treatment, because I want to isolate the effect of adapting practices based on weekly instructor reflections. In my study, instructors were randomized to treatment and control conditions to reduce bias, mitigate the impact of confounding variables, create comparable groups, and support causal inference. To control for instructor-level variables and account for dependency in repeated AEPC measures, I leverage hierarchical and longitudinal modeling to examine instructors' learning progression, adoption, and development of equity-oriented teaching practices over time. I also applied hierarchical and longitudinal modeling to analyze growth in instructors' AEPC while accounting for instructor-level characteristics.

In order to provide insight into instructors' learning processes, I complement these analyses with qualitative methods that include exploring teaching reflections, teaching materials, and evidence of equitable teaching. This approach enables me to understand instructors' lenses (i.e., their values, motivations, perspectives) for adopting AEP, as well as to what extent instructors plan, teach, assess, and reflect on data to improve student success in over a dozen STEM courses in disciplines such as computer science, data science, and statistics. Identifying and examining these lenses enables me to characterize the mechanisms that distinguish low- and high-growth instructors on AEP competency. Through this mixed-methods research, I build on the literature in chapter 2 by bringing together learning sciences and organizational behavior theories to analyze the mechanisms that a) motivate instructors to learn equitable teaching practices and b) support greater equity in student learning outcomes.

4.1 Study Context and Experimental Methods

This experimental study examines semester-long treatment and control pedagogy courses that have the same content and measures for learning AEP. The courses focused on supporting competency with AEP practice elements. As mentioned in chapter 1, the term "facilitators" refers to those who teach the pedagogy courses, and the term "instructor" refers to student-instructors enrolled in the pedagogy courses. The pedagogy courses primarily served instructors (e.g.,

graduate student instructors, undergraduate student instructors, tutors) who were concurrently teaching under faculty direction within university STEM courses, such as those in computer science, data science, and statistics. This population of instructors is important to study because they represent the future educators for our diverse undergraduate and graduate STEM student populations. The instructors in my study often have more interactions with students than the faculty of record due to the large numbers of students enrolled in these STEM courses. Moreover, Ebert-May et al. (2011) and Ebert-May et al. (2017) found that educational development can be particularly effective when it focuses on inexperienced college STEM instructors, because they are more willing to learn and adopt non-traditional pedagogies (e.g., inquiry-based learning). The authors found educational development to be impactful when these instructors have opportunities for practice and reflection over a sustained period of time (e.g., multiple months).

In my study, first-time STEM instructors at an R1 public university were randomly assigned to either a treatment or control AEP pedagogy course. The treatment and control courses were co-led by the same facilitators (i.e., a faculty member and myself jointly) who employed exemplary active learning activities and provided an AEP curriculum. These pedagogy courses met weekly, and instructors taking these courses were asked to review student data related to their teaching practices. As described in further detail below, the treatment adapted the pedagogy course curriculum based on instructors' weekly written teaching reflections and the control did not. To ensure delivery of the treatment did not impact the control, instructor reflections from the treatment were reviewed weekly by pedagogy course facilitators only after they taught the control. Table 3.1 compares the 2 conditions. The two conditions are the same in all respects, except for the use of adaptation based on weekly reflections. As described in my conceptual framework, adaptation is a fundamental aspect of AEP and the purpose of this study is to isolate and examine its potential effect. The close resemblance of the two conditions does involve the risk that the treatment effect may be small and may not be detected in this study.

Table 3.1

Comparison of Control and Treatment Pedagogy Courses

Pedagogy course activity	Control	Treatment
Same AEP curriculum content, reflection prompts, assignments/ assessments, and opportunities to redo assignments	✓	✓
Same planning activities: examine course syllabi, learning outcomes, student data	✓	✓
Facilitator provides dialogue and activities based on instructors' expressed interests (e.g., conversation about pedagogical challenges/ needs) for 30-40 minutes each week	✓	✓
Instructors write weekly reflections on teaching	✓	✓
Instructors discuss reflections with each other	✓	✓
Using weekly instructor reflection data, pedagogy course facilitators adapt activities to instructors' learning needs, thereby modeling and applying 6 AEP elements explicitly		✓

In the treatment, facilitators employed the 6 AEP elements in teaching the pedagogy course, thereby modeling how to use student learning data to improve instruction. For example, the facilitators reviewed how instructors used student data to improve teaching as articulated in their weekly reflections. After reviewing reflections and identifying instructors' learning needs,

facilitators explicitly noted making adjustments and demonstrated in the treatment condition how to interpret and use assessment data more effectively for improving student success. For instance, in response to reflections, facilitators expressed why and how to implement AEP elements regarding assessment data, modeled these AEP elements explicitly, and discussed data collection methods to empower instructors to do the same and apply AEP with their students. Furthermore, instructors reported that they valued experimental, quantitative, and replicated research that illustrated the efficacy of specific teaching practices they were asked to adopt. Responding to instructors' perspectives, treatment facilitators highlighted additional randomized, quasi-experimental, or quantitative studies that validated the impact of inclusive and evidence-based pedagogies. In particular, addressing skepticism and resistance to education and qualitative research was addressed with a follow-up discussion of additional replication studies. In the treatment, facilitators also showed the value of qualitative research and emphasized asking instructors to collect their own qualitative and quantitative data to see what kinds of teaching practices worked. The instructors learned from each other and the facilitators to see and experience how using multiple forms of data was helpful for understanding student learning barriers and adapting teaching.

By contrast, the control condition's discussions for using student data were not informed by instructors' weekly reflections.

4.1.1 Participants

Table 4.1 shows that instructors in the treatment (n=65) and control (n=64) groups have similar gender and URM characteristics. The control has about 31% women and 11% URM. The treatment has 28% women and 9% URM.

Table 4. 1

Descriptive statistics for instructors' gender and URM status by control and treatment condition

Covariate		**Control** (n=64) Frequency (%)	**Treatment** (n=65) Frequency (%)
Gender	Female	20 (31.25%)	18 (27.69%)
	Male	42 (65.63%)	46 (70.77%)
	Gender Non-Conforming	2 (3.13%)	1 (1.54%)
URM status	URM	7 (10.94%)	6 (9.23%)
	Not URM	57 (89.06%)	59 (90.77%)

4.1.2 Methods and Data Sources

I collected multiple data sources described below. I collaborated with graduate students, staff, and faculty researchers who supported me with collecting, rating, and analyzing these data

sources. I will refer to this group of individuals as "researchers" throughout my dissertation. The faculty member and I were aware of the study hypotheses at the beginning of the study.

To answer research questions 1 and 2, I used the AEPC assessments. The pedagogy courses required instructors to complete repeated measures of adaptive equity-oriented pedagogical competency (AEPC): a baseline, a midterm, and a final AEPC assessment. These assessments were course requirements. The AEPC assessment contained reflection questions that asked instructors to provide evidence demonstrating AEP competencies. AEPC assessments elicit evidence about teaching philosophies, written reflections on teaching and assessment practices, and evidence of student learning (e.g., student gain score visualizations) and teaching effectiveness. As described in chapter 5, the AEPC assessment was developed and validated using Wilson's (2005) validity framework.

For both conditions, the pedagogy course started in week 2 of the 16 week semester. The pre-assessment reflection questions were divided into 3 parts. Part 1 was assigned and due in week 2. Part 2 was assigned in week 4 and was due in week 5. Part 3 was assigned in week 5 and was due around the beginning of week 6. The pre-assessment reflection questions were administered across weeks 2-5, because I sought to examine how instructors had been teaching in the beginning of the semester. The midterm was assigned in week 9 and was due in week 11. The final was assigned in week 13 and was due at the end of week 15. Final AEPC assessments were still allowed to be submitted in week 16 without penalty to accommodate the demands and contextual circumstances instructors experienced.

To answer research question 2, I drew from instructors' post-AEPC and non-AEPC teaching reflections to understand the presence of qualitative mechanisms that motivate instructors to learn and adopt Adaptive Equity-Oriented Pedagogy (AEP) in higher education.

Specifically, 10 low-growth and 10 high-growth AEPC instructors, as determined by the quantitative analysis, were selected from the treatment condition. To accomplish this goal, the bottom 25% and the top 25% of the 65 instructors on AEPC growth in the treatment group were identified by the magnitude of the change in their AEP competency from pre- to post-semester. First, all URM instructors in either of these groups were selected. Then, additional instructors were randomly sampled to create sets of size 10 in both the low- and high-growth groups. The participants assented to the use of their data. The 10 low- and high-growth instructors were not notified that they were designated as low- and high-growth instructors, because this analysis was conducted post-hoc.

Appendix C provides the selected reflection questions that constituted the dataset for RQ2. I elected to focus on the post-semester AEPC assessment because these items ask instructors to describe their teaching philosophies and values, what they have done, the evidence of demonstrating AEP competency, and their reflections on how they have adjusted their perspectives and pedagogical practices to advance student success. In addition to reflections used to derive AEPC scores, additional (non-AEPC) reflections were solicited on pedagogically relevant topics. For example, I drew on non-AEPC reflections regarding who is responsible for learning (e.g., the student, instructor, etc.), to further understand the values and motivations to learning AEP. This approach enables me to answer research question 2 because it offers insights on the motivations and characteristics of low- and high-growth instructors.

Observation notes were taken during every pedagogy course meeting in both conditions to determine fidelity of implementation for both conditions, though these notes were not directly used for analysis in this dissertation.

4.2 Variables and Notations for Multilevel Regression Models

In this section, I describe the multilevel regression models that I use to answer research questions 1 and 2. I use hierarchical and longitudinal modeling to account for dependency in repeated AEPC measures. Model 1 enables me to examine whether instructors in the treatment condition grew more over time, on average, than those in the control condition, controlling for instructor-level characteristics. As mentioned in chapter 1, I control for instructors' gender and URM identities because instructors with non-male and URM identities may find it easier to learn and adopt AEP, since STEM is often a toxic environment for women and students of color (Sax et al., 2018). As a result, non-male and URM instructors may be more motivated to learn equitable pedagogies due to their experiences in STEM, where there is a lack of equity and inclusion. I also control for the number of students of each instructor, since adapting instruction based on student learning data may be more feasible for smaller classes. I also control for years of teaching experience, because Ebert-May et al. (2011) and Ebert-May et al. (2017) found that educational development can be particularly effective when it focuses on inexperienced college STEM instructors, since they are more willing to learn and adopt non-traditional pedagogies. For my quantitative analyses, I examine the following variables:

4.2.1 Model 1 Response Variable

AEPC variable (y_{ij})

The continuous outcome AEPC variable (y_{ij}) on a scale from 0-22, refers to assessments (i = 0, 1, 2) for the pre-semester (pre-intervention, timepoint 1), mid-semester (timepoint 2), and the final (timepoint 3), where j is an index for the student. Every assessment measured AEPC. The AEPC assessment questions ask students to write a teaching philosophy and practice statement and discuss examples and evidence of their teaching effectiveness. There were open-ended items that asked instructors about their teaching practices in the classroom and application of equity-oriented practices. See Appendix D for examples of the AEPC assessment questions. Appendix E contains a scoring guide.

As explained in chapter 5, the AEPC measure was validated using Wilson's (2005) validity framework. As a part of that process, a Partial Credit Rasch model was fit and the AEPC latent variable was estimated using Warm's Weighted Likelihood Estimate (WLE). The reliability of the instrument based on WLEs is 0.94, suggesting that the instrument has high consistency in measuring AEPC. The AEPC assessment shows no gender assessment bias and is positively and significantly correlated with researchers' review of teaching and pedagogical materials. More details about AEPC assessment validation will be described in chapter 5. In order to incorporate pre- and post-AEPC assessment scores, I use raw scores rather than WLEs for the multilevel regression analysis in chapter 6.

For additional scoring validation and to estimate inter-rater reliability, another researcher independently scored a random sample of 36 AEPC assessments. When analyzed for this purpose, the AEPC assessments were de-identified by name and condition and were assigned a numerical code. The results of this inter-rater reliability study are presented at the end of chapter 5.

4.2.2 Model 1 Covariates

Time dummy variables(t_{1ij} , t_{2ij})

For the statistical model, I created dummy variables for each time point except the first one (i =0). The first time-point is the reference category. The first dummy variable is t_{1ij}, which is 1 at the second occasion (i = 1) and 0 otherwise. The second time variable is t_{2ij}, which is 1 at the third occasion (i=2) and 0 otherwise.

Treatment Group (E_j)

The treatment group variable is a time-invariant dichotomous variable indicating whether instructors were randomly assigned to the treatment or control condition. The value of the variable is 1 for instructors in the treatment condition and 0 for instructors in the control condition. The treatment and the control conditions have the same pedagogy course facilitators who employ active learning strategies. However, in the treatment condition, these facilitators adjusted instructional practices based on instructors' weekly reflections on teaching. The treatment's adaptation started after the first month of the semester. The control condition is the reference group.

Underrepresented minority status (URM) (U_j)

The underrepresented minority status variable reports whether an instructor is an underrepresented minority (URM). I use a time-invariant dichotomous variable indicating the URM status of an instructor (0 = non-URM; 1 = URM). Non-URM instructors are the reference group.

Gender (G_j)

G_j is a vector of two dummy variables for the gender of the instructor. I have a dummy variable for female and for gender non-conforming with male as the reference category.

Years of teaching experience ($Teach_j$)

The years of teaching variable reports the number of years the instructor has taught a course before at either the current or previous institution. The variable is continuous.

Number of students (N_j)

The number of students variable reports the number of students that the instructor taught during the semester of the study. The variable is continuous.

Years of tutoring experience ($Tutor_j$)

The years tutoring variable reports the number of years the instructor spent tutoring students. The variable is continuous.

4.3 Data Analysis

In this section, I first describe my quantitative data analyses, including the multilevel regression models that I use to answer research question 1. Then, I provide my qualitative data analysis approach that I use to answer research question 2.

4.3.1 Quantitative Data Analyses

Model 1. To answer research question 1, I use multilevel modeling to examine whether there is a greater increase in mean AEPC scores across time in the treatment pedagogy course than in the control pedagogy course, controlling for gender, URM status, years of teaching experience, and number of students. I use hierarchical linear modeling with random intercepts for longitudinal data. Regarding research question 1, for measurements of instructor *j* (level 2) at time *i* (level 1), I regress y_{ij} on dummy variables for timepoints 2 and 3 (t_{1ij} , t_{2ij}), treatment group (E_j), interactions between time and treatment ($E_j t_{1ij}, E_j t_{2ij}$), URM status (U_j), a vector of two dummy variables for instructors' gender (G_j), years of teaching experience ($Teach_j$), number of students (N_j), and years of tutoring experience ($Tutor_j$).

I am examining growth in AEPC by studying the increase from pre-intervention AEPC scores to mid-semester AEPC scores and from pre-intervention AEPC scores to post-semester AEPC scores, and how much the growth differs between the treatment and control. I am taking this approach instead of controlling for pre-intervention AEPC because of Lord's paradox. Here is my multilevel regression model:

$$y_{ij} = \beta_0 + \beta_1 t_{1ij} + \beta_2 t_{2ij} + \beta_3 E_j + \beta_4 E_j t_{1ij} + \beta_5 E_j t_{2ij} + \beta_6 U_j + \beta_7 G_j + \beta_8 Teach_j + \beta_9 N_j + \beta_{10} Tutor_j + \zeta_j + \epsilon_{ij}$$

Where $\epsilon_{ij}$is the individual and time-specific residual and ζ_j is the individual-specific random-intercept both assumed to be normally distributed, with zero means and variance parameters θ and ψ, respectively.

4.3.2 Qualitative Data Analysis

To answer research question 2, I examined 10 high-growth and 10 low-growth AEPC instructors in the treatment as determined by model 1. Among these instructors, I identified frequent themes invoked for each group. I qualitatively coded each of six lenses (listed below) separately to protect against bias when coding and testing conjectures. When coding the data, I was blind to the instructor's name, condition, and whether they are in the low- or high-growth group. I deductively code reflections using a codebook, tracking frequencies of codes and exploring self-reported mechanisms for learning AEP (Saldana, 2011). I describe different phases below of how I developed the codebook.

Phase 1: I developed an initial codebook based on my knowledge of the educational development literature and my expectations given that I taught the pedagogy course. The lenses, informed by the literature as described in chapters 2 and 3, formed the super codes. I list these super codes below:

- Authority and Community
- Learning Goal
- Equity Definition and Perspectives
- Data
- Emotional
- Responsibility for Learning

Phase 2: By qualitatively coding written reflections from two instructors from another semester that is not the focus of this dissertation, I was able to revise the codebook, test it for external (proximal) validity, and introduce new subcodes as needed.

Phase 3: I revised the subcodes in consultation with dissertation committee members. The codebook below represents the final set of codes.

Phase 4: I developed conjectures that would be examined through application and comparison of code patterns to the new corpus of data that is the focus of this dissertation. I present those code conjectures later in this chapter.

Phase 5: I systematically applied codes to the corpus of 20 participants' teaching reflections. The unit of analysis is at the reflection question level for instructor written reflections. Responses to reflection questions are my unit of analysis because they enable me to understand how instructors plan, teach, assess their teaching, and reflect on data to improve student success.

Codes were applied for each reflection question. More than one code can be applied to the same unit of analysis. Examples of applying codes will be described in the next section. My goal was to analyze the presence, frequency, and contingency of selected codes to determine similarities and differences in what motivates instructors to learn AEP.

I tagged each written reflection response for a given participant with evidence of each of the codes. This provided me with frequency and code co-occurrence data about whether particular lenses are invoked repeatedly for a given participant. This approach also allowed me to compare self-reported mechanisms for learning AEP across the treatment and control conditions.

Another researcher who is familiar with the AEPC construct was provided with the codebook and some examples. I had a detailed orientation and overview of the coding strategy. At this orientation, we together scored reflections from other semesters that were not used as part of qualitative analyses in this dissertation. The researcher practiced coding during this time and I gave feedback. The researcher independently coded written reflections, being blind to instructors' condition and whether they were in the low- or high-growth group.

We qualitatively coded each of these lenses separately to protect against bias when coding and testing conjectures. Using the codebook, we performed this coding in *MAXQDA 2012* (VERBI Software, 2021), since the software allows us to examine the cross-comparison and co-occurrences of codes.

4.3.2.1 Code Book

I use qualitative codes to understand the differences between low-growth and high-growth instructors from pre- to post-semester on the AEPC measure. In Tables F1-F6 in Appendix F, I describe the super codes and subcodes that guide my qualitative analyses for examining the mechanisms that motivate instructors to learn and adopt AEP. Below, I provide a brief description of the super codes which correspond to the lenses I defined in chapter 3. These lenses are beneficial for interpreting the instructors' written reflections.

Authority and Community Lens. The Authority and Community Lens involves instructors who mimic and value practices from other instructors, their department, or academic field. An example would be: "I am following or am heavily influenced by someone's pedagogical model." One of the key subcodes is "I mimic esteemed equity-oriented instructors".

Data Lens. The Data Lens describes how valued forms of data and analytical methods affect pedagogical practice. The Data Lens is defined as the kinds of knowledge and data that are important and relevant in an academic and socio-organizational culture. The Data Lens considers

the sources from which instructors can make determinations about student competencies. An instructor can use the Data Lens to understand student perspectives and learning processes. This lens can be useful for diagnosing student needs and adapting teaching practices. Some of the subcodes include: “I value quantitative data”, “I value qualitative data”, and “it depends on the context”.

Learning Goal Lens. The Learning Goal Lens is defined as the goals, concepts, or ideas that one foregrounds in their mind. These goals, concepts, or ideas may include a focus on academic content and/or equity issues. This lens speaks to the purpose and beliefs that instructors have for their teaching. For example, instructors may focus on learning goals that emphasize academic concepts. Another example is that instructors may also have goals focused on addressing equity, such as addressing stereotype threat and improving student sense of belonging in the classroom. The subcodes for this lens include the “academic lens” and “equity lens”.

Responsibility of Learning Lens. The Responsibility for Learning Lens refers to how the instructor views their role and the student’s role in learning-- whose responsibility is it if a student doesn’t learn a concept? Are the “good” students the ones who teach themselves or don’t need help? This lens contributes to scholarship on reflective college teaching since it accounts for the systems, structures, and/or beliefs that foreground individualistic notions of learning that have contributed to STEM students leaving the major because they felt under-supported and isolated in a competitive environment (Jackson & Cobb, 2010; Phuong et al., 2021; Seymour & Hewitt, 1997). Some of the subcodes include: “It’s on my student”, “It’s on me”, and “It’s on something else”.

Emotional Lens. The Emotional Lens is defined as the emotions that arise from engaging in a practice, experience, or reflection. I contend that the norms of a socio-organizational culture impact how individuals think they should feel and how they express their feelings. From organizational behavior theory, natural systems thinking perspectives would support this lens by stating that instructors’ intrinsic motivations, experiences, and emotions can motivate behavior (Scott, 2006). “Emotion tied to personal learning experiences” and “Emotions tied to how students did in class” are some of the subcodes for this lens.

Equity Perspectives and Definitions Lens. The equity perspectives and definitions lens posits that instructors’ definitions of equity can motivate them to adopt or not adopt AEP. One definition of equity may include providing all students with the same learning opportunities. Another definition may focus on adapting learning opportunities based on students’ learning needs. “Equity is giving everyone the same thing (i.e., equality)” and “Equity is adapting to address student needs and their socio historical context” are two different subcodes under this lens.

4.3.2.2. Summary of Conjectures

I conjectured that the treatment’s effectiveness may be related to the lenses that instructors revealed in written reflections. I provide my conjectures below:

1. **Authority and Community Lens differences.** Compared to the low-growth group, I expected to see the authority and community subcode, “I mimic esteemed equity-oriented instructors”, occurring more often in the high-growth group. Therefore, I hypothesized that high-growth instructors will be inspired by other instructors and peers who use adaptive data-driven practices. I anticipated that this inspiration would motivate high-

growth instructors to adopt these data-driven practices to improve equitable outcomes because they have been socialized by other instructors who have done the same.

2. **Data Lens differences.** I expected instructors in the low-growth group to primarily focus on quantitative data over qualitative data, not responding to equity barriers and verbal/ non-verbal cues. Whereas, with the high-growth group, I expected to see the "I value quantitative data", "I value qualitative data", and "it depends" subcodes to co-occur more often. Therefore, I hypothesized that high-growth instructors will use and acknowledge the validity of both quantitative and qualitative student data when adapting teaching.
3. **Learning Goal Lens differences.** Among the low-growth group, I expected to see an overemphasis on academic content without a focus on equity barriers. Whereas, with the high-growth group, I expected to see more emphasis on both academic content and equity barriers. In other words, I expected to see both Learning Goal Lens subcodes (i.e., "academic goal" and "equity goal") co-occurring more often in the high-growth group.
4. **Responsibility of Learning Lens differences.** In the low-growth group, I expected to see a higher presence of the subcode "It's on my student" relative to the subcodes: "It's up to me" or "it's something else". I expected to see a higher co-occurrence of these codes in the high-growth group. I anticipated that the high-growth group would be more motivated to identify learning barriers and adapt teaching to address them because these instructors put the responsibility for learning on themselves, the student, and other contextual factors. In addition, they believe their students' success can and should be improved by adapting teaching and assessment practices.
5. **Emotional Lens Differences.** I anticipated that the high-growth instructors would be more motivated by their emotional response to how students perform in their class because they would be adapting their teaching based on students' social-emotional and academic needs. I would anticipate that high-growth instructors would be more adaptive and responsive to emotions than the low-growth instructors. For example, in the low-growth group, I expected to see a lower presence of the subcode "[My] emotions are tied to how students DID in my class (measured via their performance)". I expected to see a higher presence of this subcode in the high-growth group.
6. **Equity Perspectives and Definitions Lens Differences.** I anticipated that definitions of equity focused on adapting instruction to address students' needs would be more prevalent among high-growth instructors than among low-growth instructors. For example, in the low-growth group, I expected to see a lower presence of the subcode "Equity is adapting to address student needs and their socio historical context". I expected to see a higher presence of this subcode in the high-growth group. By contrast, I conjectured that the low-growth instructors would have a higher presence of the following subcode than the high-growth instructors, for instance: "Equity is giving everyone the same thing (i.e., equality)". I state this conjecture because I imagine that low-growth instructors would be less willing to adapt their teaching to students' needs if they believe that everyone should have the same learning experience.

Explanation of core hypotheses. Regarding the Learning Goal Lens, a qualitative driver for lower AEPC growth could be that the instructor primarily values academic content and success, and tends to ignore other humanistic elements including equity barriers that the student faces. In addition, the Authority and Community Lens could explain growth in AEPC because higher-growth instructors may mimic the equitable teaching practices and norms in the academic

department. By contrast, this lens could explain how low-growth instructors might mimic toxic and inequitable pedagogical norms that can be common in STEM departments. The Data Lens can also explain a lack of AEPC growth because the instructor does not value the validity of qualitative data that focus on the human experience and equity barriers. Not considering multiple forms of data could hinder AEPC growth because the instructors might not diagnose and respond to equity barriers and interests. In addition, being responsive to verbal and non-verbal cues from observation (qualitative data) can help an instructor adapt to better support student success. I anticipated that the Responsibility for Learning Lens may be a qualitative driver for growth on AEPC because instructors who share the responsibility with students will be more likely to adapt their teaching to address student learning needs. In addition, I conjectured that the high-growth instructors on AEPC would notice and address students' equity and learning barriers.

Moreover, I anticipated that the high-growth instructors would be more motivated to adapt their teaching based on their emotions of how students did in their class, because these instructors would be more attentive and responsive to students' social-emotional and academic learning needs. Similarly, I conjectured that more high-growth instructors would have a definition of equity that focuses on adapting to students' learning needs and sociohistorical context, because such adaptation would be consistent with their beliefs around equity.

Nevertheless, I recognize my existing biases within this research study and provide a positionality statement below to contextualize a part of my worldview.

4.3.2.3 Positionality

My research is influenced by my positionality as a scholar and practitioner. My positionality stems from my identities as a low-income historically underrepresented student who comes from a Latino and Vietnamese immigrant household whose first language is not English. My underrepresented student identity and learning disability fuel my passion to research and ask questions about equity-oriented teaching, since I want to support students who struggle with learning. I have experienced how the K-16 pipeline can feel like a maze without a map, especially with access to limited resources. I am committed to restructuring this maze into a highway towards greater access, equity, and inclusion.

Professionally, I have taught the pedagogy courses I am studying and am engaged in educational development efforts. The team I am working with has collected pilot data and has been engaging in experimental research for many years. The AEP framework is embedded in the pedagogy courses in this study. As an instructor and a learner, I have experienced how instructors adjusting their practices in the classroom and during office hours can improve my performance and other students' achievement. Therefore, I believe that teaching can impact student learning. These experiences contribute to my post-positivist lens, which values multiple forms of evidence and perspectives, since each form of evidence alone can be limited in understanding learners (LeCompte & Schensul, 1999). Therefore, I believe that valid knowledge needs to rely on data that is triangulated in ways that elevate multiple perspectives, which include the researcher, instructor, and student among others. I feel that triangulation creates valid knowledge because individual data sources or perspectives have explicit and invisible limitations.

CHAPTER 5: Prior Validation of AEPC Measure

5.1 Introduction and Measurement Context

In order to rigorously examine how instructors progress in the development of equity-oriented teaching practices and the degree to which a particular intervention might support such development, this research requires a reliable and valid measure of equity-oriented pedagogical competence. Having found no such measure in existence, I developed, empirically field tested, and validated a measure specifically for this research. This chapter explains how my Adaptive Equity-Oriented Pedagogy Competency (AEPC) measure was developed and validated, including the measurement framework used, the psychometric methods applied, and evidence of reliability and validity collected. The description of the measure provided centrally in this chapter is intended to make the explanation of research results in subsequent chapters clearer and more succinct.

AEPC evaluates the extent to which instructors are practicing elements of the Adaptive Equity-Oriented Pedagogy described in chapter 3. In this sense, AEPC can be used as a measure of college instructors' effectiveness with applying equitable pedagogies. In chapter 6, I draw on the AEPC measure to address research question 1. I utilize this measure for investigating the efficacy of the pedagogy courses on instructors' adoption of equitable teaching practices. In this context, the collection of validity evidence is essential.

This measure includes materials that are often included in portfolios used in part to evaluate college teaching, such as teaching philosophy and practice statements, learning outcomes, formative assessment, and reflections on teaching (Seldin et al., 2010). This study is the first, to my knowledge, that uses a Partial Credit Rasch model and item response theory to validate a measure of equitable teaching at the R-1 university level. In chapter 6, I draw on the AEPC measure for researching the efficacy of interventions on instructors' adoption of equitable teaching practices.

5.2 Methodological Framework: The BEAR Assessment System

I used Wilson's (2005) validity framework to guide the development and psychometric analysis of the AEPC measure. The fundamental elements of Wilson's framework serve as the foundation for the work of the Berkeley Evaluation and Assessment Research (BEAR) Center and the BEAR Assessment System (BAS; Wilson & Sloan, 2000) in particular. Those fundamental elements are referred to as "four building blocks" of an assessment system: construct map, item design, outcome space, and measurement model (Wilson, 2005). I used the BEAR Assessment System to support the AEPC measure's development.

The construct map lays out the definition of what it means to be competent in a trait, such as AEPC. The items are designed to enable individuals to demonstrate their competence by successfully answering questions with difficulties that vary along the construct. The participants respond to items in a well-defined outcome space (e.g., through selection of multiple choice, writing an open-ended response that is scored according to a rubric). The measurement model defines how responses in the outcome space relate to levels in the construct.

I created a construct map for AEPC to support sound assessment design of instructors' adaptive equity-oriented practice. See Appendix B for the construct map. I have been engaging in the four building blocks of the BAS assessment model (e.g., construct map, item design, outcome

space, and Rasch model) (BAS; Wilson & Sloane, 2000; Wilson, 2005), and have been developing a thorough literature review to support the construct's development. I have defined an outcome space by creating a rubric to score each item in order to relate it back to the construct map. As I engaged in the four-building-blocks process, I collaborated with experts in the areas of educational development, learning sciences, equity and inclusion, and college teaching in item design. I also discussed the outcome space and rubrics for these items with these experts.

If this construct map is well-researched and validated, then the construct map can effectively inform instructional decisions for instructors. Accordingly, I have developed items that are grounded in the AEP elements (described in chapter 3) and the course objectives. I designed 6 items that aligned with various levels of my construct map. The 6 items performed well as a set and served as the basis for the analyses for the remainder of this report. I present details about these items in Table 5.1.

Table 5. 1

Items and Levels

Item number	Item	Score Levels	Construct Map Levels
i1a	Question 1a. Teaching Philosophy and Practice statement. See Appendix D for question prompts.	0-7	0-4c
itc	What is the pre-assessment activity that you used to determine whether or not the learning outcome has been achieved? Explain how this assessment aligns with a homework, midterm, and/or final assessment. Consider how you scored this assessment to compare it to the post-assessment. Ideally, this assessment should not take more than 10 minutes. What is the post-assessment activity that you used to determine whether or not the learning outcome has been achieved? Explain how this assessment aligns with a homework, midterm, and/or final assessment. Consider how you scored this assessment to compare it to the pre-assessment. Ideally, this assessment should not take more than 10 minutes. Note: These pre- and post-assessments were created by the instructors. The instructors administered these pre- and post-assessments to their students during class.	0-1	0-1
i8a	What is a list of the teaching practices and activities you will implement to reach your learning outcome? How do these practices and activities incorporate the core elements of Adaptive Equity-Oriented Pedagogy? In particular,	0-7	0-4c

	please describe how your teaching practices respond to your students' existing competencies, interests, and content area knowledge.		
i10a	How did the lesson go and what patterns of student behavior and learning did you notice? Based on data collected from your students, did your teaching practice from your lesson work? If so why? If not, why not? In your response, describe all the specific AEP teaching practice elements you applied and how students responded to these practices.	0-7	0-4c
i11a	Based on the data you collected from your students, what would you do next to improve your students' success?	0-7	0-4c
i12a	What did you learn today that you did not know about your students? What else would you like to know about your students to help you further support their learning? What new or refined practices would you employ to achieve this goal and why?	0-7	0-4c

I defined the outcome space by creating scoring rubrics for the items. In particular, for the items scored from 0-7, the following rubric was used:

Table 5. 2

Scoring Rubric for 0-7 Item Scores

Construct Map Level	**Item Score**	**Description**
4c: Continuous commitment to adaptation and improvement (Very High)	7	Educator describes all AEP elements with an example and explains how their practice impacts the students' success (Citing student and instructor evidence) or Describes all AEP elements with an example and explains rationale for how practice impacts students' success (Citing student and instructor evidence)
4b: Continuous commitment to adaptation and improvement (Very High)	6	Describes application of all AEP elements with an example (Citing instructor and student evidence)

4a: Continuous commitment to adaptation and improvement (Very High)	5	Describes application of all AEP elements
3: Adaptation based on diagnosis (High)	4	Adapts based on full diagnosis, not only based on diagnosis of learning competency
2b: Full Diagnosis (Moderate)	3	Diagnoses learning competency and interest (Full level 2) **Or** Diagnoses learning competency (gets the concept right) and learning barriers (equity barriers, lack of interest, or misconceptions) (Alternative version of Full Level 2)
2a: Partial diagnosis (Moderate)	2	Diagnoses student learning competency **Or** Diagnoses interest
1: Clarifies learning outcomes; Aligns formative assessment and instruction with learning outcomes (Low)	1	Demonstrates application of element 1 and/or 2 Element 1: Communication and clarification of learning outcomes Element 2: Alignment of formative assessment with learning outcomes
0	0	Provides an irrelevant answer, does not apply any elements

For item itc which is scored from 0-1, scorers focused on assessing whether or not instructors met level 1. The following collapsed rubric below was applied:

Table 5. 3

Scoring Rubric for 0-1 Item Scores

Construct Map Level	Item Score	Description
1: Clarifies learning outcomes; Aligns formative assessment and instruction with learning outcomes	1	Demonstrates application of element 1 and/or 2 both questions for item itc. Element 1: Communication and clarification of learning outcomes Element 2: Alignment of formative assessment with learning outcomes. For example, instructors would need to answer the question as to whether the pre- or post-assessments align with the learning outcomes reflected in graded assessments (i.e., homework, midterm, final).
0	0	Does not fully demonstrate elements 1 and 2 for both questions for item itc. For example, the participant only answers one of the questions for item itc correctly and does not answer both itc questions correctly. Provides an irrelevant answer, does not apply any elements

To support consistent scoring using these rubrics, a detailed scoring guide was created and provided to the scorers. See Appendix E for this detailed scoring guide.

5.2.1 Measurement Models

5.2.1.1 Model 1: Partial Credit Model (PCM)

There was an ordinal scoring strategy for item responses, ranging from 0-7. As shown in Table 5.2 above, the point values align with specific levels from my construct map. There are 6 items scored in my model.

I fitted a Partial Credit Model (Masters, 1982) and estimated person locations (θ_n) and the item difficulty parameters (the step difficulties, δ_{ij} of the item *i* and of the levels *j* within the item). This model places items and respondents on the logit scale. I provide an equation below that shows the Partial Credit Model's model-implied probability of a respondent at location θ_n responding to item i at level j. Item *i* has step difficulties $\delta_{i0}, \delta_{i1}, \delta_{i2}, \ldots, \delta_{iM_i-1}$ for steps $0 \ldots M_i - 1$, where M_i is the number of score categories for item *i.* The probability of respondent *n* receiving score *m* for item *i* is

$$P\left(y_{in} = m \mid \theta_n, \delta_{ij}\right) = \frac{\exp \sum_{j=0}^{m} \left(\theta_n - \delta_{ij}\right)}{\sum_{k=0}^{M_i - 1} \exp \sum_{j=0}^{k} \left(\theta_n - \delta_{ij}\right)} \qquad (1)$$

where appropriate constraints are applied to make the model identifiable. In particular, it is assumed that the mean person location is constrained to zero. In the special case where $M_i = 2$, Equation 1 is the Rasch model for dichotomous items.

5.2.1.2 Model 2: Partial Credit Model (PCM) with Latent Regression

For the PCM, De Boeck and Wilson (2004) present latent regression as a person-explanatory extension. I conducted a latent regression to account for measurement error when examining the effects of covariates such as fidelity of implementation scores and gender.

In this latent regression PCM model, the person location θ_n in Equation 1 is decomposed into a sum of terms that depend on instructors' covariates. In this model, Z_{nj} is the value on covariate j for instructor p.

$$\theta_n = \sum_{j=1}^{J} \beta_j Z_{nj} + \epsilon_n \qquad (2)$$

It is assumed that $\epsilon_n \sim N(0, \sigma_\epsilon^2)$. β_j is the regression coefficient for covariate j, β_0 is the intercept, and ϵ_n is the residual of the instructor's location, controlling for J covariates. The regression model's coefficient β_j may be interpreted as the effect of a one-unit change in covariate j on the instructors' AEPC, which in turn affects their probability of scoring at higher levels. As applied to my data set, Z_{n1} is the instructor's gender and Z_{n2} is the instructor's fidelity of implementation score, $J = 2$, β_1 is the effect of gender being female, and β_2 is the effect of a one-unit difference in fidelity of implementation scores on the participant's AEPC location.

5.2.1.3 Model 3: Differential Item Functioning (DIF)

Paek and Wilson (2011) provide the general form of the Rasch DIF model, as presented below:

$$P\left(x_{ni} = 1 \mid \theta_n, G_n\right) = exp(\eta_{ni}) / (1 + exp(\eta_{ni}))$$

and

$$\eta_{ni} = \theta_n - \delta_i + \gamma_i G_n$$

It is assumed that θ_n is normally distributed with a mean that differs by group, i.e., $\theta_n | g \overset{iid}{\sim} N(\mu_g, \sigma)$. As noted by Paek and Wilson (2011), this population modeling accounts for the overall group ability differences, which enables the DIF comparison to be based entirely on γ_i.

This PCM DIF model adds a term, γ_i, the DIF parameter, which allows the difficulty of item i to be different for the reference group (versus the focal group) by an amount γ_i. G is an

indicator for the focal group. The DIF parameter measures the interaction of the group and the item. This parameter also represents the item difficulty difference between the focal and reference group (i.e., $\gamma_i = \delta_F - \delta_R$). The PCM DIF model has a similar form as the Rasch DIF model for dichotomously scored items.

ConQuest (Adams et al., 2012) was utilized to fit my item response model. Statistical analyses were conducted in Excel, Stata, and ConQuest. The DIF analysis was also run in ConQuest. I used Stata to produce figures such as scatter plots.

5.3 Validation Data and Procedures

Through multiple assignments, I collected assessment data regarding instructors' adaptive equity-oriented pedagogical competency (AEPC). For the study, the assessment was administered to instructors in the graduate-level pedagogy and data analysis courses that I have researched. These instructors represent my target population. The sample of instructors (n=124) included in the validation analysis is slightly smaller than the sample (n=129) described in detail in chapter 4. For the validation dataset, the gender non-conforming instructors were excluded because the number was too small to support stable estimation of parameters related to this gender category. Two other instructors were not a part of the validation analysis because the scoring of their responses was not fully determined at the time the validation analyses took place. Differences between the various data sets are further elaborated in chapter 6 below. Respondents wrote their responses in a Word or text processing document. All instructors were asked to write their responses in complete sentences. Below, I describe how I examine reliability and validity in this study.

5.3.1 Reliability

Internal construct consistency. After the model was fit, person locations, which represent the individual's position on the logit scale, were estimated using both Warm's Weighted Likelihood Estimate (WLE) and Expected a Posterior (EAP) (Wilson, 2005). EAP are empirical Bayes estimates and are likely to be biased towards the mean. On the other hand, WLEs are bias-corrected maximum likelihood estimates. The reliability of these WLE estimates is reported and represent the proportion of the observed variance attributable to variance in the underlying trait (i.e., AEP competency) (Wilson, 2005). In particular, I estimated the person separation reliability for WLE estimates, which indicates how precise my person location estimates are in measuring my construct in relation to the between-person variability. High reliability coefficients indicate that my instrument is performing well in separating respondents with different proficiency levels.

Inter-rater reliability. Because the AEPC instrument and AEPC scoring guide were created by me, I needed independent scoring of AEPC assessments to provide more critical validity evidence, such as inter-rater reliability (IRR).

Thirty-six instructors were randomly selected for an interrater reliability study. I conducted IRR using the raw item scores from AEPC midterm assessments of these 36 instructors used for research question 1. As described in chapter 4, there are 4 items for this measure. Below, I discuss the procedure.

To examine whether the measure could be reliably scored by independent raters, I conducted an inter-rater reliability analysis. I trained a researcher who was familiar with the construct on the rubric for each item. I had an orientation and overview of scoring strategy with this researcher.

We scored exemplary sample papers that were not used as part of IRR analyses. He practiced scoring and I gave him feedback.

Then, he independently rated the 36 papers that I had previously rated. While scoring, the researcher was blind to my scores. We did not discuss each other's scoring before or during inter-rater reliability analysis.

I then estimated inter-rater reliability using Cohen's Kappa with a quadratic weight. This approach is appropriate because:

1. The AEPC measure is an ordinal variable.
2. The same two researchers rated each instructor's AEPC assessment responses.
3. The two researchers conducted the ratings independently of each other.
4. For each assessment response, the researcher provided a rating based on the same number and definition of categories.
5. The ratings are considered paired observations on the same measure.
6. It is appropriate to penalize larger discrepancies more than smaller ones (Sim & Wright, 2005).
7. It is a common approach to quantifying rater variability among human raters (Shermis, 2015).

5.3.2 Validity

Response processes. For summative instrument evaluation purposes, I conducted 10 think-alouds with instructors who completed the assessments. I examined the response processes that students expressed while they completed the assessment.

Here is a script for my think-alouds: "I want you to read the question out loud or by yourself. I would like you to explain your understanding of the questions and your thought process as you answer the questions. Please tell me what you're thinking. If you're quiet for a while, I may remind you to think out loud. I may also ask follow-up questions, such as, "May you please elaborate on what you said?" I may also ask, "What prompted you to answer the question in this way?" and "What prompted this specific answer?".

I also administered exit survey questions to ensure that items are not ambiguous and do not have multiple interpretations that could interfere with one demonstrating their competency with applying the AEP elements. These questions were adapted from Wilson's (2005) text and included:

1) On average, how long did you take for each question?
2) What parts of the assignment are confusing or difficult to answer? Why?
3) Which questions best allowed you to demonstrate your competency with applying adaptive equity-oriented pedagogy? Why or why not?
4) Which questions, if any, did not allow you to demonstrate your competency with applying adaptive equity-oriented pedagogy? Why or why not?
5) Were the instructions sufficiently clear? How might the instructions be improved?
6) Did you go back and change any of your answers? If so, why?
7) If you were writing this assignment, how would you change specific questions? Why would you change these questions?
8) What questions might you add to this assignment and why?

Internal structure. I gathered evidence that the items individually and as a set provide a coherent measure of the AEPC construct. I have aligned the scoring levels from my items to the

levels of my construct map. To gather evidence to help validate these hypothesized relationships, I examined the location of item thresholds using the Wright Map (these results are presented in the next section). I also looked at the mean proficiency increase across score categories within each item. I hoped to see higher average thetas (i.e., proficiency scores) in relatively higher categories across all items, meaning, people who tend to score high on a particular item, tend to do so across all items.

Evidence based on relations to other variables. To demonstrate relationships to other variables, I examined convergent validity with my construct. Evidence of convergent validity is established when the construct is associated with a variable one would want it to be associated with. I expected AEPC to be associated with variables such as AEP fidelity of implementation.

I measured AEP fidelity of implementation (scale 0-7) using an AEP element checklist that a researcher would use when observing an instructor's teaching and looking through teaching materials. I expect to see a positive relationship between AEP fidelity of implementation and the instructor's proficiency with the construct because I anticipate that my measure can capture instructors' AEPC.

For fidelity of implementation, researchers filled in a questionnaire. Researchers were asked the following questions regarding the instructors' teaching:

> Question 1. Based on observations of this student-instructor's teaching and review of this student-instructor's teaching materials (e.g., evidence of teaching effectiveness), please mark all of the following practices that characterize this student-instructor's teaching.
>
> _____ A: Communicates and clarifies learning outcomes
> _____ B: Aligns formative assessment with learning outcomes
> _____ C: Diagnoses student learning relative to learning outcomes.
> _____ D: Diagnoses student interest and how it's relevant to course content.
> _____ E: Adapts based on diagnosis
> _____ F: Demonstrates a continuous commitment to adaptation and improvement
> _____ G: Advances a coherent pedagogical rationale in both teaching practice and materials
>
> Question 2. Please elaborate on the reasons for the marks you provided above.

Consequences of use. I am also interested to examine if the instrument's usage led to some desirable and/or undesirable outcomes. I solicited data about the consequences of use that I report below.

Fairness. I examined if the instrument's usage is fair for particular groups (Wilson, 2005). I hope that my instrument's items are fair for instructors of different genders, since research suggests that the field of STEM can be toxic for women (Sax et al., 2018). I used differential-item functioning to make sure that male instructors who have the same θ , or proficiency score, do not have significantly higher item scores than female instructors.

5.4 Results

5.4.1 Wright Map

The Wright Map has a logit scale with respondents on the left and item thresholds on the right. The left panel shows a distribution of instructors' estimated locations. The right panel

shows the locations of levels within items (Wilson, 2005). Respondents who are aligned with an item-level threshold have a 50% chance of scoring at the designated level or above. Respondents who are above the threshold have more than a 50% chance of scoring at or above the designated level.

According to my Wright Map in Figure 5.1, the items appear to cover the range of respondent competencies. Nonetheless, I have some respondents who are above the highest items. In other words, the most competent respondents need more challenging items for optimal measurement.

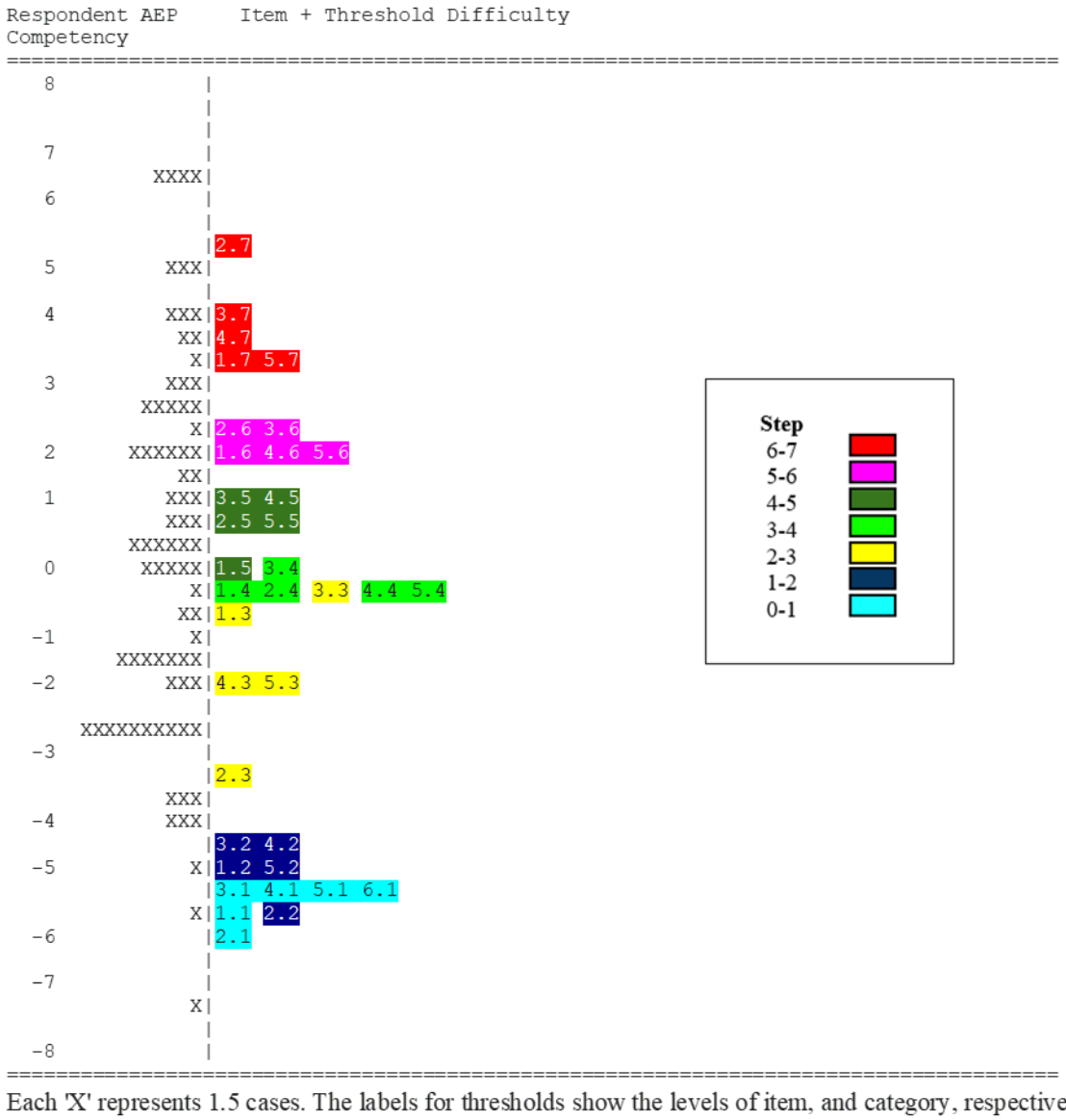
Overall, the item category locations on the Wright Map are similar for item levels that map onto the same construct levels. In the legend in Figure 5.1, I indicate that the bottom threshold in light blue refers to level 1 and the top threshold in red refers to level 4c (i.e., scores of 7). I see banding for each level. I notice that the banding is ordered in the way I hypothesized in the construct map. Respondents who tend to score high on the construct tend to score high across items.

Based on my Wright Map, respondents going from the second to the third threshold required a larger step compared to the step required from level 1 to 2. In addition, respondents going from threshold 3 to 4 to 5 do not require sizable increases in the construct since these thresholds are close to each other.

I notice that the level 1 and level 2 item locations are close to each other and have some degree of overlap. I also noticed that the level 3 and 4 item locations are close to each other. In the next iteration of scoring, it might make sense to collapse these categories into one level. For level 3, it seems that the item locations for items i11a, i12a, and i8a are lower than the item locations of items i1a and i10a. It is worth conducting future research on why these item locations are different.

Figure 5. 1

Wright MAP of AEPC



Each 'X' represents 1.5 cases. The labels for thresholds show the levels of item, and category, respectively.

5.4.2 Item Level Results

Item level internal structure. The mean location increase and banding of thresholds of my Wright Map support construct validity. I expected to observe a monotonic increase in mean WLE as levels increase within each item. A monotonic increase means that as the item level increases, the mean θ (i.e., M WLE) for instructors achieving the item level also increases (Wilson, 2005). I checked and noticed that the mean WLE is increasing for each item. Table 5.4 below shows the typical pattern for item i1a responses. For example, for item itc, I noticed that the mean WLE increases from -2.05 for level 0 to 0.054 for level 1.

Table 5. 4

Instructor Mean Location Increase for each item level

Item	Response Categories	Count	Percent (%)	Mean Location	Std. Dev. of Locations
i1a	0	3	2.42	-6.19	0.93
i1a	1	4	3.23	-4.60	1.48
i1a	2	38	30.65	-2.14	1.01
i1a	3	8	6.45	-0.93	0.78
i1a	4	9	7.26	0.37	0.65
i1a	5	24	19.35	1.56	1.09
i1a	6	17	13.71	2.69	1.13
i1a	7	21	16.94	4.32	2.05
i10a	0	4	3.23	-5.95	0.90
i10a	1	4	3.23	-4.44	1.16
i10a	2	44	35.48	-1.84	1.15
i10a	3	6	4.84	0.06	1.47
i10a	4	18	14.52	0.66	0.64
i10a	5	14	11.29	1.79	0.75
i10a	6	19	15.32	2.98	0.79
i10a	7	15	12.10	5.33	1.29
i11a	0	4	3.23	-5.95	0.90
i11a	1	3	2.42	-4.73	1.17
i11a	2	27	21.77	-2.46	0.75
i11a	3	18	14.52	-1.32	0.71
i11a	4	21	16.94	0.63	0.59
i11a	5	15	12.10	1.79	0.78
i11a	6	18	14.52	2.77	0.84
i11a	7	18	14.52	5.00	1.45
i12a	0	4	3.23	-5.95	0.90
i12a	1	2	1.61	-5.36	0.63
i12a	2	26	20.97	-2.59	0.70
i12a	3	21	16.94	-1.23	0.78
i12a	4	19	15.32	0.63	0.63
i12a	5	16	12.90	1.86	0.91
i12a	6	16	12.90	2.62	0.93
i12a	7	20	16.13	4.78	1.53
i8a	1	2	1.61	-6.61	0.83
i8a	2	17	13.71	-3.19	1.49
i8a	3	35	28.23	-1.62	1.13
i8a	4	16	12.90	0.69	1.12
i8a	5	23	18.55	1.85	1.02
i8a	6	24	19.35	3.28	1.40

i8a	7	7	5.65	6.18	1.18
itc	0	6	4.84	-2.05	4.58
itc	1	118	95.16	0.54	2.89

Item parameter estimates and item fit statistics. I examined how the data fit the model on an item by item basis using weighted mean square statistics (wMNSQ), which are also known as infit statistics (Wilson, 2005). Infit statistics characterize the degree to which the observed randomness agrees with the statistical expectation based on the model. Items with infit statistics above 1.33 are more random than expected. Items with infit statistics below .75 are less random than expected.

Furthermore, item-level categories with infit statistics above 1.33 suggest that these item-level categories are more random than the model implies. In other words, there may be some high achieving students who score low on the category and vice versa. More specifically, instructors with high thetas tend to answer the item incorrectly at that level and instructors with low thetas tend to answer the item correctly at that level. In the ConQuest output, I did not see any items or item levels that had infit statistics above 1.33 (See Table 5.5).

I have provided tables below to show individual item difficulty parameter estimates, item category difficulty parameter estimates, and their associated infit statistics. I flag items as "L" (i.e., low) with fit statistics below 0.75, and I flag items as "H" (i.e., high) with fit statistics above 1.33.

Table 5. 5

Item Difficulty Parameters and Fit Statistics

Item	$\hat{\delta}$	**std.err**	**wMNSQ**	**CI**	***t*-statistic**	**Flag**
i1a	-0.882	0.118	0.88	(0.69, 1.31)	-0.7	
i8a	-0.939	0.129	0.68	(0.74, 1.26)	-2.6	L
i19a	-0.399	0.117	0.78	(0.70, 1.30)	-1.6	
i11a	-0.778	0.118	0.44	(0.73, 1.27)	-5.3	L
i12a	-0.854	0.118	0.54	(0.73, 1.27	-4.0	L
itc	-5.347	0.498	1.17	(0.28, 1.72)	0.6	

Note. *An asterisk next to a parameter estimate indicates that it is constrained
Separation Reliability = 0.985
Chi-square test of parameter equality = 332.28 df = 6, Sig Level = 0.000

Table 5. 6

Item-category difficulty parameter estimates and fit statistics

Item	category	$\hat{\delta}$	std.err	wMNSQ	CI	*t*-statistic	flag
i1a	0			0.49	(0.00-2.04)	-1.1	L
i1a	1	-4.133	0.333	0.95	(0.15-1.85)	0	
i1a	2	-4.544	0.303	0.77	(0.67-1.33)	-1.4	
i1a	3	0.808	0.260	0.94	(0.47-1,53)	-0.1	
i1a	4	0.390	0.273	0.85	(0.52-1.48)	-0.5	
i1a	5	0.560	0.263	0.92	(0.79-1.21)	-0.7	
i1a	6	2.986	0.293	0.97	(0.73-1.27)	-0.2	
i1a	7	3.934*		1.19	(0.62-1.38)	1.0	
i10a	0			0.35	(0.00-2.00)	-1.6	L
i10a	1	-4.162	0.346	0.97	(0.14-1.86)	0.1	
i10a	2	-4.934	0.314	0.91	(0.65-1.35)	-0.5	
i10a	3	1.180	0.268	0.99	(0.33-1.67)	0.1	
i10a	4	-0.530	0.274	0.81	(0.74-1.26)	-1.5	
i10a	5	1.810	0.268	0.89	(0.67-1.33)	-0.6	
i10a	6	2.332	0.296	0.81	(0.75-1.25)	-1.6	
i10a	7	4.305*		0.64	(0.58-1.42)	-1.9	L
i11a	0			0.35	(0.00-2.02)	-1.6	L
i11a	1	-3.614	0.337	1.01	(0.00-2.03)	0.2	
i11a	2	-4.741	0.317	0.81	(0.75-1.25)	-1.5	
i11a	3	-0.718	0.250	0.80	(0.74-1.26)	-1.6	
i11a	4	0.197	0.255	0.70	(0.76-1.24)	-2.7	
i11a	5	1.984	0.260	0.88	(0.69-1.31)	-0.7	
i11a	6	2.635	0.296	0.83	(0.74-1.26)	-1.3	
i11a	7	4.257*		0.71	(0.61-1.39)	-1.6	L
i12a	0			0.36	(0.00-2.07)	-1.5	L

i12a	1	-3.199	0.336	1.05	(0.00-2.33)	0.3	
i12a	2	-5.138	0.322	0.83	(0.74-1.26)	-1.3	
i12a	3	-0.895	0.247	0.82	(0.77-1.23)	-1.7	
i12a	4	0.501	0.253	0.74	(0.74-1.26)	-2.1	
i12a	5	1.873	0.260	0.89	(0.71-1.29)	-0.7	
i12a	6	2.839	0.304	0.92	(0.72-1.28)	-0.6	
i12a	7	4.019*		0.73	(0.62-1.38)	-1.5	L
i8a	0			0.65	(0.00-2.23)	-0.5	L
i8a	1	-1.874	0.424	0.01	(0.0-6.28)	0.0	L
i8a	2	-7.976	0.424	0.93	(0.66-1.34)	-0.4	
i8a	3	-2.254	0.282	0.82	(0.80-1.20)	-1.9	
i8a	4	1.052	0.270	0.87	(0.69-1.31)	-0.8	
i8a	5	1.482	0.274	0.96	(0.79-1.21)	-0.4	
i8a	6	3.294	0.289	0.94	(0.75-1.25)	-0.4	
i8a	7	6.276*		0.69	(0.43-1.57)	-1.2	L

Note. An asterisk next to a parameter estimate indicates that it is constrained

I identified 2 potential reasons below for my low fit statistics, both of which may merit further investigation.

Potential reason 1 for low fit statistics: Repetition between items. I hypothesize that item i8a has slightly low fit statistics because it asks for similar content as item i1a. In fact, both items ask participants to describe the AEP practice elements they have applied. Similarly, I hypothesize that i11a and i12a have low fit statistics because these items were next to each other and were asking for similar content (i.e., what instructors would do next based on data). This might explain why the item responses follow a similar pattern. In the next iteration, I should improve the items to make them less repetitive to each other, in order to reduce the local dependence between these items. This may improve the fit statistics. In chapter 6, I describe how I address this issue by creating one item score for i11a and i12a when answering research question 1.

Potential reason 2: Items differentiate low- and high-performers in deterministic ways. I also examined items i1a, i10a, and i11a because the lower categories of these items have low infit statistics. The low infit statistics could be present, since it is possible that these items are discriminating between low and high competency instructors in a way that is more deterministic than the model implies. This phenomenon could be the case because the item level ordering and scoring strategy follow approximately a Guttman structure, which is a deterministic model.

5.4.3 Reliability

Person separation reliability. I estimated person separation reliability, which describes the extent to which a measure was able to separate people with different levels of proficiency (Wilson, 2005). In ConQuest, the WLE person separation reliability of the instrument is estimated as 0.94. Similarly, in ConQuest, the MLE person separation reliability is estimated as 0.94. Although I was unable to obtain an EAP reliability estimate in ConQuest, these WLE and MLE reliability estimates suggest that the instrument has relatively high consistency in measuring AEPC.

Inter-rater reliability. Tables 5.7-5.10 contain the inter-rater agreement data for each of the 4 items that I studied. Below each table are statistics indicating the percent of exact agreement, the percent of exact or adjacent agreement, Cohen's weighted Kappa, and Cohen's Kappa. These statistics are most commonly used to measure inter-rater consistency in assessments. In Tables 5.7-5.10, the diagonal numbers are the frequencies for exact agreement. The numbers above the diagonal values are the frequencies of adjacent agreement.

Table 5. 7

Inter-rater agreement for item 1

		Rater 1								
		0	1	2	3	4	5	6	7	Total
	0	1								1
	1		1							1
	2			4						4
	3				1					1
Rater 2	4					5				5
	5						1			1
	6						1	10		11
	7							2	10	12
	Total	1	1	4	1	5	2	12	10	36

% exact agreement = 91.67%
% exact or adjacent agreement = 100%
Estimated Quadratic weighted Kappa = .99
Estimated Quadratic weighted = .99

Estimated unweighted Kappa = .891

Table 5. 8

Inter-rater agreement for item 2

		Rater 1		
		0	1	Total
Rater 2	0	1		1
	1		35	35
	Total	1	35	36

% exact agreement = 100%
Estimated Quadratic weighted Kappa = 1
Estimated Kappa = 1

Table 5. 9

Inter-rater agreement for item 3

		Rater 1								
		0	1	2	3	4	5	6	7	Total
Rater 2	0	2								2
	1		1							1
	2			2						2
	3				2					2
	4					4				4
	5					1	12			13
	6						2	9		11
	7								1	1
	Total	2	1	2	2	5	14	9	1	36

% exact agreement = 91.67%
% exact or adjacent agreement = 100%
Estimated Quadratic weighted Kappa = .99
Estimated Kappa = .889

Table 5. 10

Inter-rater agreement for item 4

		Rater 1								
		0	1	2	3	4	5	6	7	Total
Rater 2	0	4								2
	1		0							1
	2			1						2
	3				5					2
	4				1	16				4
	5					2	0			13
	6						6	1		11
	7								0	1
	Total	2	1	2	2	5	14	9	1	36

% exact agreement = 91.67%
% exact or adjacent agreement = 100%
Estimated Quadratic weighted Kappa = .99
Estimated Kappa = .879

The quadratic weighted kappa captures the degree of agreement between two raters above the agreement that would be expected to occur due to chance. The kappa statistic ranges from 0 (i.e., no agreement) to 1 (i.e., perfect agreement). Kappa is the measure of chance-corrected agreement. Kappa relates the agreement to the variation in the population. In my sample of scores, zero occurs as well as seven. There is a great heterogeneity in the responses. The chance correction becomes more serious when there is less variation. Therefore, it is easier to obtain a high kappa when there is a larger range of responses like in my sample.

In Table 5.11, Altman (1999) and Landis and Koch (1977) provide guidelines for interpreting strength of agreement captured by quadratic weighted Kappa:

Table 5. 11

Guidelines for interpreting the strength of agreement with Kappa

Kappa range	Altman's (1999) interpretation	Landis and Koch's (1977) interpretation
< 0.20 Poor	Poor	Poor
0.21 - 0.40 Slight	Fair	Slight
0.41 - 0.60 Moderate	Moderate	Moderate
0.61 - 0.80 Good	Good	Substantial
0.81 - 1.00 Very good	Very good	Almost perfect

Based on these guidelines for interpreting the strength of agreement, the range of quadratic kappas on the AEPC inter-rater agreement reliability study across items is "very good" or "almost perfect" (see Tables 5.7-5.10). In each case, the Kappa statistics are statistically significant (p<0.0001), indicating that the agreement is much higher than due to chance alone.

The level of inter-rater reliability is very high for several reasons. First, the rubric is very analytic rather than holistic in nature. For example, raters are counting the number of elements with evidence. In addition, the training of the second rater was extensive. Consequently, the second limitation is that a brand-new researcher who was not as well-trained might have a lower agreement rate. Moreover, the small sample of 36 papers from 1 university and pedagogy course on AEP may be similar in ways that make them easier to score consistently in comparison to papers collected in a different context. The instructors were also encouraged to bold the AEP elements in their responses, which allowed for easier scoring. The measure when administered to a new population could yield more ambiguous responses, which could reduce IRR.

5.4.4 Validity

I constructed my argument for validity based on the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 2014). The standards have five strands that include instrument content, internal structure, response processes, evidence related to other variables, and consequences. Although the items are being used in assignments in real courses, I am researching the psychometric properties of the instrument to better understand the construct and to improve the instrument in the future. This instrument, its items, and the Partial Credit Rasch model can also inform how course grades are assigned and how feedback is provided to instructors. Another intended use of my assessment is for formative and summative evaluation of the instructors. An additional purpose for the instrument is to evaluate the course curriculum's effectiveness in supporting instructors' mastery of adaptive equity-oriented pedagogy competency. I hope to use data from this study to revise the measure and curriculum of the graduate-level pedagogy and data analysis course.

Content validity: Evidence related to instrument content. As mentioned above, I designed my construct based on the competencies needed to demonstrate proficiency with

adaptive equity-oriented pedagogy. I created the items based on the construct map. In particular, score levels for each item are associated with a construct map level.

For formative instrument development purposes, I have conducted think-alouds and item paneling with the instructors and the teaching assistants who are familiar with the construct. They have provided me with feedback on improving the items and my construct map. Therefore, my construct map and items have gone through multiple revisions.

Evidence based on response process. The evidence on response process validity gathered is encouraging. The instructors indicated that the assessment items were very easy to understand. The instructors interviewed also stated that the questions were easy to understand. For instance, one instructor wrote in an exit survey, "The instructions were very clear. Thanks for being so clear! Other instructors can learn from y'all." Another instructor stated, "I feel like everything was straightforward." When describing the instructions, an instructor said, "The instructions were very clear and concise." Interestingly, one instructor expressed: "I think the written instructions, and those given in class were very clear in explaining the goals of the assignment to us." This quotation prompted me to consider investigating in the future the degree to which clarifying instructions in-class impacts the ease of understanding items on the instrument.

I also asked instructors about which questions enabled them to best demonstrate their competency with applying AEP. One instructor stated, "Questions 1, 2, 3, 4, 6, 9, 10, 11, 13, 15, and 16 all allowed me to demonstrate my competency with applying AEP practices mainly because they either had m[e] reference them or demonstrate the outcomes of using them." Another instructor expressed: "Specifically, the questions in the first section were good questions that helped stimulate our minds about how we use AEP in the classroom. These questions directly asked us how these principles help our teaching and how we can incorporate these principles into our teaching."

I also asked how instructors might change the assignment to improve it. One instructor stated that she would recommend requiring more data visualizations to be produced. Another instructor expressed, "I would not change any question and the overall assignment is designed in a reasonable way." Similarly, when I asked instructors about adding questions, many instructors said that they would not add more questions to assignments. I hypothesize this may be the case because the assignments already had many questions. It is also possible that many instructors did not want to have to answer more questions.

When responding to questions during think alouds, the instructors drew on experiences, student data, course frameworks (e.g., adaptive equity-oriented pedagogy (AEP)), and concepts to motivate their responses. For example, one instructor, "I drew on adaptive equity-oriented pedagogy the most, as I am rapidly adapting my teaching practices based on feedback. I believe this framework should be applied to every student as it is the fairest and help students improve the quickest." When describing the concepts that influenced their thought process as they responded to questions, a tutor expressed: "I wanted to use AEP the most because it gave me a step by step process on how to effectively tutor each of my tutees. The five core concepts are easy to follow and have shown a lot of success when I used them. I think every student should be able to have a takeaway for the concepts that they want to learn from a class. In order to achieve this, I used AEP and wanted to get better at this concept."

During interviews, many instructors described how they cycled through the different elements of the AEP framework to adapt teaching practices that resonated with their philosophy and that they thought would improve student learning, such as wise criticism, modeling, inquiry-based learning, and diagram mapping. One instructor stated that he "relied heavily on the AEP

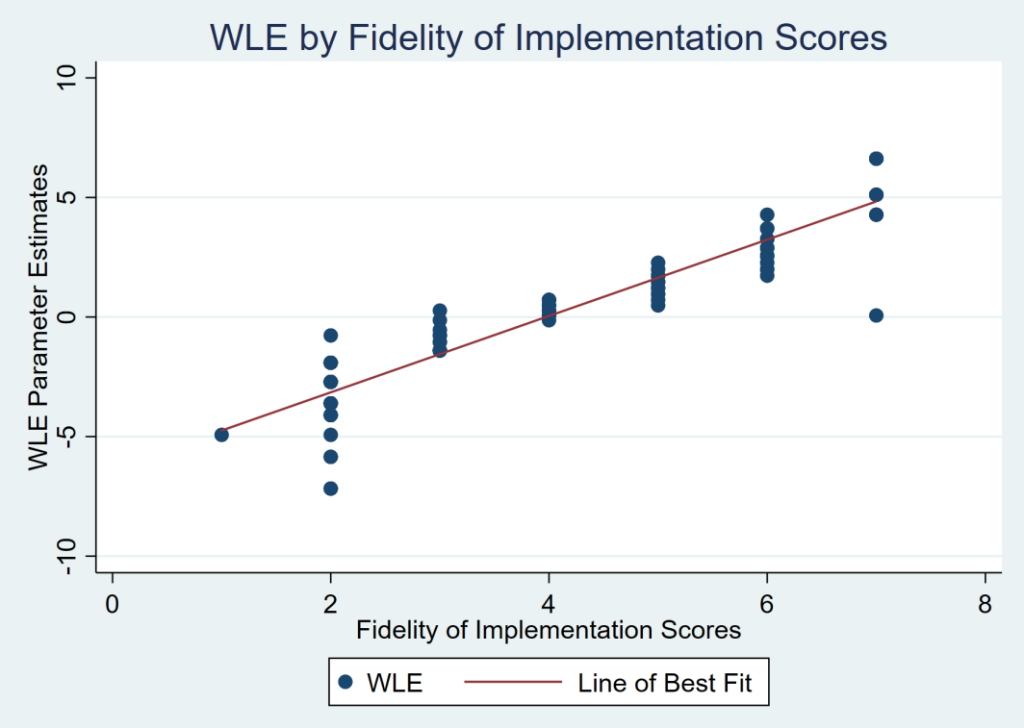
methods because many of the principles of AEP agreed with [his] own teaching philosophy and [he] had found that based on [his] experiences with the students the AEP methods were the most effective."

One instructor mentioned that a few questions seemed similar to each other. For instance, this instructor said, "A lot of the questions are a bit repetitive, so maybe cutting it down and adding more variation in questions would be good." However, this instructor provided correct answers to these questions and expressed a solid understanding of what the question was asking. In the future, I will try to see if I can consolidate some of my items to make the assessment sound less repetitive.

Evidence based on relations to other variables: Fidelity of implementation. My hypothesis about convergent validity was met between instructors' WLE parameter estimates and the researcher's AEP fidelity of implementation score. The correlation between the WLE parameter estimates and the researcher's AEP fidelity of implementation score is 0.94. This correlation may be spuriously high because some of the materials reviewed as a part of the fidelity evaluation were also submitted by the instructors and scored as item responses for determining the measure. Nonetheless, this correlation was statistically significant at the 5% level. See Figure 5.2 for the scatter plot between these two variables, which portrays a positive relationship between these two variables.

Figure 5. 2

Scatterplot of WLE vs. AEP Fidelity of Implementation



In order to further assess external validity, I conducted a Latent Regression with the PCM. I regressed instructors' AEPC on their fidelity of implementation scores, when controlling for gender. For each unit increase in the fidelity of implementation score, the instructors' location in the logit scale is estimated to increase, on average, 1.571 logits when controlling for gender.

This increase was statistically significant at the 5% level. I present these results in Table 5.12 below. This result suggests that when an instructor scores higher on the fidelity of implementation score, there is a higher likelihood that the instructor will have demonstrated relatively higher degrees of AEPC.

When controlling for fidelity of implementation scores, gender does not significantly predict an instructors' AEPC at the 5% level. This result is promising because gender is not expected to be associated with AEPC competency.

Based on these findings, researchers' review of instructors' teaching and their pedagogical materials tended to align with instructors' competency estimates. These results can further support the external validity of my measure, since I hypothesized that instructors' WLE parameter estimates and the researcher's AEP fidelity of implementation score would be positively correlated.

Table 5. 12

Latent Regression

Predictor	**Coefficient**	**Std. Err**
Fidelity of Implementation Score	1.571*	0.040
Gender	0.248	0.150

Note. $*p < .05$

Fairness: Differential Item Functioning (DIF). ConQuest was used to fit a model for DIF. In addition, when the DIF model was run in ConQuest, I noticed that the standard errors were large for item itc, which can be explained by the fact that ConQuest tends to provide a larger standard error to the final item in the estimation process. Therefore, the DIF analysis was run again in ConQuest where item itc was the first item in the list with the intention of obtaining a more accurate estimate of the item's standard error. I used this standard error for itc, which I utilized to produce the standard error of $\hat{\gamma}$ in the table below. I used the estimates from the first DIF model run in ConQuest for producing the values in Table 5.13 below.

Table 5. 13

DIF Item Analyses for Gender

Item	**Focal Group**	$\hat{\gamma}$	**StdErr of** $\hat{\gamma}$
i1a	Female	-0.112	0.236

i8a	Female	-0.216	0.258
i10a	Female	-0.074	0.236
i11a	Female	0.220	0.236
i12a	Female	0.114	0.236
itc	Female	0.15	0.998

Note. $*p < 0.05$ None of the $\hat{\gamma}$ are statistically significant

I apply Paek and Wilson (2011)'s criteria for analyzing DIF. Paek and Wilson (2011) state that a logit difference value less than 0.426 is "negligible", a value between 0.426 and 0.638 is considered "intermediate", and a value over 0.638 is considered as a "large" DIF (p. 1028).

No items show statistically significant DIF. For example, the first item, i1a, is estimated to be 0.112 logits less difficult for the females. In other words, female students found the item to be easier than male students. This difference is not statistically significant because a 95% confidence interval [-0.575, 0.351] contains zero.

In addition, the magnitude of the DIF is negligible because they are always less than the Paek and Wilson (2011) threshold of 0.426 logits. For instance, item i1a has $\hat{\gamma}$ = -0.112. Since |-0.112 logits| < 0.426, this means the observed level of DIF, even if it were statistically significant, would be negligible (Paek & Wilson, 2011).

Item i11a has the largest magnitude in favor of males ($\hat{\gamma}$ = 0.220), but this value is negligible in magnitude and is not statistically significant. However, due to the size of the data set, the statistical power to detect DIF is low. Therefore, DIF analysis should be performed again when additional data is collected.

In sum, I found no evidence of Differential Item Functioning (DIF), or item bias, between males and females because no item by gender interactions (i.e., $\hat{\gamma}$) were both statistically significant and above 0.426 in magnitude (see Table 5.13) (Paek & Wilson, 2011). In other words, for all levels of item difficulty, male instructors do not have significantly higher item scores than female instructors who have the same AEP Competency score, or vice versa.

The absence of gender DIF in the measure is evidence that the measure is equally fair for males and females. This enables for the meaningful interpretation of any gender effects that may appear in multilevel regression analyses that address research questions 1 and 2.

Consequences of use: Benefits. One of the goals of this assessment is to help instructors document equity-oriented practices and demonstrate teaching effectiveness in ways that go beyond course evaluations. In particular, there is evidence that instructors were able to document and improve their AEP practices: instructors stated that the instrument helped them improve their

instruction and student learning. For example, one instructor stated, “I draw upon Adaptive Equity-Oriented Pedagogy because I have seen firsthand how using these strategies benefit[s] not only myself, but also the students I teach.” In addition, an instructor expressed, “I relied heavily on the AEP methods because many of the principles of AEP agreed with my own teaching philosophy and I had found that based on my experiences with the students the AEP methods were the most effective.” Similarly, another wrote: “I thoroughly enjoyed this assignment. I think it made me think a lot more during my tutoring session and made me focus solely on helping the student. I also think it was super impactful to visualize my impact on my students. Before, I just tutored and didn’t think much about the outcome. After deploying a pre and post assessment, the improvement is illustrated vividly during tutoring. Thank you for this opportunity.”

Many instructors also discussed how the AEP framework helped them promote inclusive learning environments. When describing how examining student learning data and feedback helps check biases, an instructor said, “I drew on Adaptive Equity-Oriented Pedagogy because whenever I teach, I aim to be as inclusive as possible. I also want to be able to understand and check my biases, and to have my students feel comfortable in their classroom settings.”

Multiple instructors also expressed that this assessment helped them apply AEP to adjust teaching practices they were passionate about (e.g., inquiry-based learning, providing examples to show problems, wise criticism, growth mindset-oriented praise).

Figures 5.3 and 5.4 provide examples of how instructors documented improvements in student learning based on identical pre- and post-assessments on the same content and items.

Figure 5. 3

Student Pre- and Post-assessment Gains

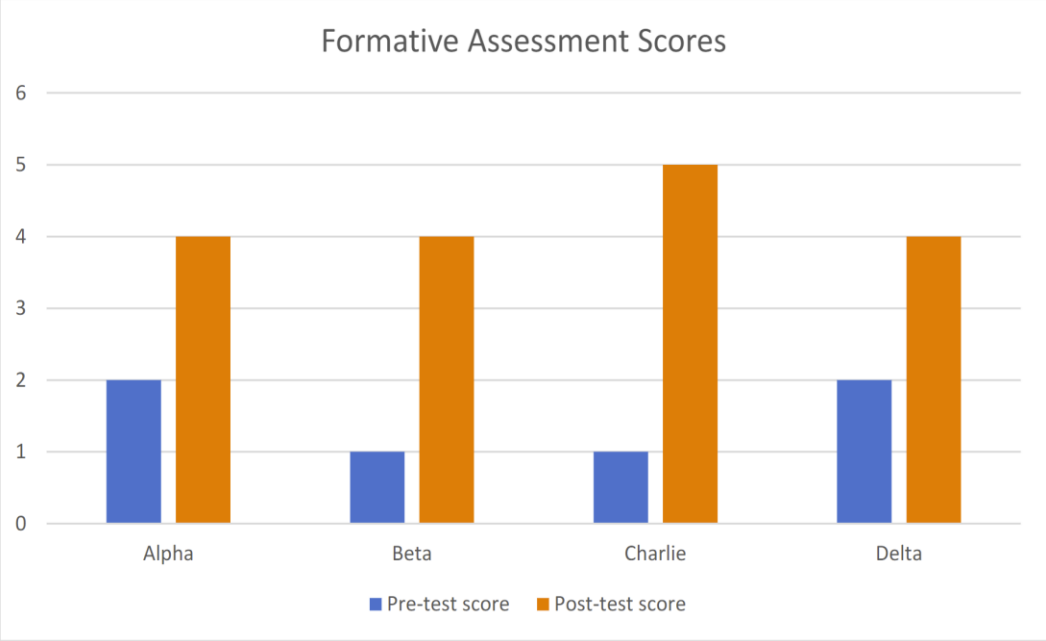
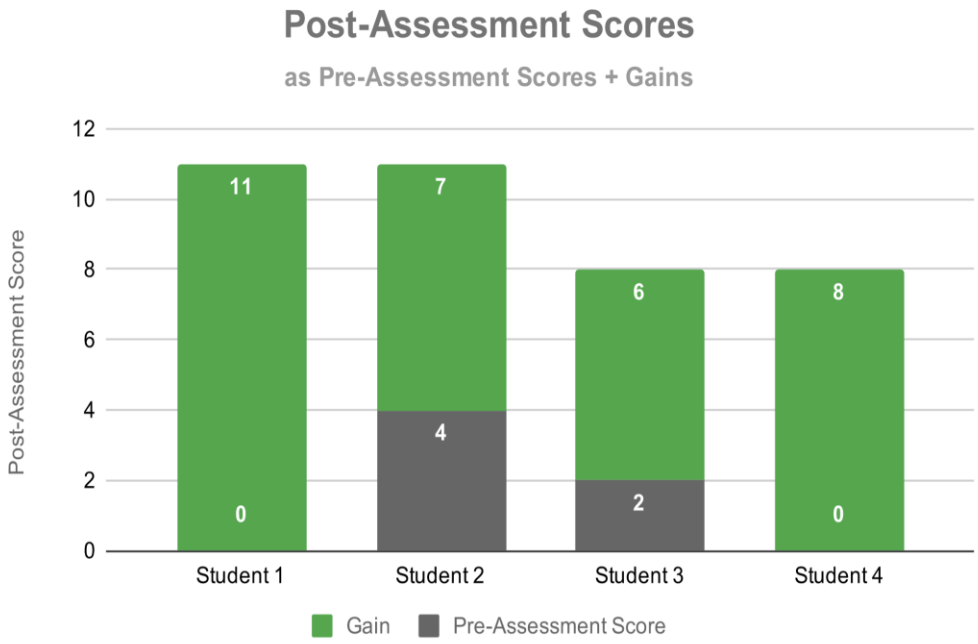


Figure 5. 4

Student Pre- and Post-assessment Gains



This assessment could also be used for research purposes to evaluate instructor's pedagogical impact on student learning from pre- to post-semester.

Potential risks. This assessment needs to be validated on a larger sample before it is used for employment decisions. A potential negative consequence could arise if this assessment were used prematurely for evaluation, hiring and promotion purposes. If this measure is not more adequately and thoroughly validated, negative consequences may arise, which include a lack of fairness and potential litigation. These are important considerations when one seeks to create an equity-oriented metric for evaluating and hiring instructors who can demonstrate a track record of improving student success.

5.5 Discussion and Lessons Learned

Based on prior research (Fink, 2013; McCallum, 2013; Phuong et al., 2017), I theorized that the highest AEPC level would be the most difficult to reach because it required instructors to demonstrate and act upon all lower-level elements. In theory, reaching the highest level requires ongoing planning, observation, active reflection, and adaptation based on data, which are the traits an instructor with high AEPC mastery would demonstrate. Based on my Wright map, it appears that the highest level of the AEPC construct was the most challenging to reach, because instructors needed a lot of the construct to endorse that level. The levels and ordering of my construct are not only supported by my Wright map, but they are also supported by mean location increases within all the items of the instrument.

I crafted multiple questions that tap into low, middle, and high levels on the AEPC construct map. I expected these items to target students at different competency levels. Based on my Wright map, it appears that I was able to accomplish this goal. After examining multiple data

sources, results from this study's reliability and validity evidence suggest that the development of this AEPC measure is moving in a positive direction.

5.5.1 Fidelity of Implementation Implications and Next Steps

The statistical significance of the fidelity of implementation predictor is an important finding, suggesting that an instructor's score on a self-report measure is significantly associated with an external researcher's assessment of their competency. One could expect this finding because instructors need to provide and cite evidence of meeting their competency in the AEPC measure, and some of this same evidence is considered in the fidelity of implementation evaluation. These results, if generalizable, suggest that the AEPC measure could potentially be a good proxy for observation and review of teaching materials. Further research needs to be conducted with a more fully independent evaluation process.

5.5.2 Practical Constraints and Limitations

I acknowledge that this validation of AEPC experienced practical constraints and limitations. For example, this study faces challenges with generalizability since this instrument was administered to a specific sample in a very specific context. The instrument was administered in a letter grade pedagogy and data analysis course that provided instructors with opportunities to learn, apply, and reflect on their adaptive equity-oriented pedagogical practice each week. In addition, the instructors had little to no classroom teaching experience. They also decided to become instructors, and a majority of them were interested in teaching. This course was intended for computer science and data science disciplines, which have a unique socio-academic culture at the R-1 university in this study. Therefore, I recommend examining the reliability and validity of this instrument in other contexts. Examination of these limitations requires further research.

5.6 Significance

This measure can help us reduce some of the problems associated with course evaluations in higher education. Instructors often cite course evaluation responses as a reason for not changing their teaching significantly (Lederman, 2018). Reasons include students rating lecture higher than active learning (Phuong et al., 2017) and instructors deciding to teach to the middle when students report a wide range of preferences. Examining and validating an equity-oriented teaching measure is significant because many instructors feel that course evaluations are emotional responses from students that are often harder to control without inflating grades or offering other incentives (e.g., cookies) (Hessler et al., 2018; Stark & Freishtat, 2014). Additionally, unlike this AEPC measure, course evaluations have been found to be biased against female instructors (Boring et al., 2016; MacNell et al., 2015; Stark & Freishtat, 2014).

Therefore, if validated with a larger sample and elevated in the teaching evaluation process, this measure could incentivize instructors to focus on high-impact equity-oriented teaching practices that they have control over, drive learning, and reduce instructor evaluation bias.

Using Item Response Theory, this study presents a clear and rigorously validated teaching measure that goes beyond course evaluations. This measure draws on questions that are legitimate within higher education teaching evaluation. These questions tap into one's teaching philosophy and practice statement and teaching materials (e.g., syllabi, formative assessment, summative

assessment, evidence of student learning). Consequently, the AEPC measure asks respondents to draw on multiple data sources and materials because teaching effectiveness cannot be adequately evaluated based on one data source (Stark & Freishtat, 2014). Furthermore, this measure can help diagnose where instructors are developmentally in a mastery model of learning based on AEPC elements, thereby informing professional development and elevating equity in conversations around effective teaching practices for all students.

CHAPTER 6: Quantitative Findings

The purpose of the chapter is to examine whether and how educational development interventions can impact instructors' equitable teaching competencies. In my conceptual framework that is described in chapter 3, the facilitators model AEP explicitly to train instructors in the treatment condition so that they can apply AEP with their students. I hypothesize that this process of modeling AEP increases instructor learning over time in the treatment relative to the control, as measured by AEPC assessments. This hypothesis is consistent with research demonstrating that instructor reflection, which is a core component of the AEP treatment, can improve learning (Phuong et al., 2017; Phuong et al., 2022). I investigate this hypothesis by conducting an experiment that enables me to isolate the effect of the treatment, where I also control for multiple instructor-level covariates. This investigation addresses research question (RQ) 1, which is presented below:

1. **Research question 1:** Is there a greater increase in mean AEP competency scores across time in the treatment pedagogy course than in the control pedagogy course, controlling for gender, URM status, years of teaching experience, years of tutoring experience, and number of students?

In this chapter, I present descriptive statistics on the continuous and categorical variables that I use for RQ 1. I also summarize the results of regression analyses that address this RQ.

Before answering this question, I collected data and produced dataset 1 to validate the AEP competency (AEPC) assessment, which was described in chapter 5. Dataset 1 holds the data on the subset of mid-semester AEPC assessment items used for validating the uni-dimensional AEPC measure. To then address RQ 1, dataset 2 includes items used for analyzing instructors' AEPC scores over time by including pre-semester, mid-semester, and post-semester AEPC assessments. Drawing on these longitudinal AEPC data, I use multilevel modeling to analyze mean differences between the treatment and control pedagogy courses, controlling for the variables mentioned in RQ 1.

Dataset 1 description (i.e., Validation Dataset). In validating the AEPC instrument using item response theory, I used a dataset with 6 items (raw scale 0-36) on the midterm as described in chapter 5. These items supported the unidimensional IRT-based AEPC score. For dataset 1, items i7a and i9a were combined into one item called item itc. Item itc is scored as 0 or 1. Table 5.1 provides the items again below that were discussed in chapter 5.

Table 5. 1
Items and Levels

Item number	Item	Score Levels	Construct Map Levels
i1a	Question 1a. Teaching Philosophy and Practice statement. See Appendix D for question prompts.	0-7	0-4c
itc	What is the pre-assessment activity that you used to determine whether or not the learning outcome has been achieved? Explain how this assessment aligns with a homework, midterm, and/or final assessment. Consider how you scored this assessment to	0-1	0-1

	compare it to the post-assessment. Ideally, this assessment should not take more than 10 minutes. What is the post-assessment activity that you used to determine whether or not the learning outcome has been achieved? Explain how this assessment aligns with a homework, midterm, and/or final assessment. Consider how you scored this assessment to compare it to the pre-assessment. Ideally, this assessment should not take more than 10 minutes. Note: These pre- and post-assessments were created by the instructors. The instructors administered these pre- and post-assessments to their students during class.		
i8a	What is a list of the teaching practices and activities you will implement to reach your learning outcome? How do these practices and activities incorporate the core elements of Adaptive Equity-Oriented Pedagogy? In particular, please describe how your teaching practices respond to your students' existing competencies, interests, and content area knowledge.	0-7	0-4c
i10a	How did the lesson go and what patterns of student behavior and learning did you notice? Based on data collected from your students, did your teaching practice from your lesson work? If so why? If not, why not? In your response, describe all the specific AEP teaching practice elements you applied and how students responded to these practices.	0-7	0-4c
i11a	Based on the data you collected from your students, what would you do next to improve your students' success?	0-7	0-4c
i12a	What did you learn today that you did not know about your students? What else would you like to know about your students to help you further support their learning? What new or refined practices would you employ to achieve this goal and why?	0-7	0-4c

Dataset 2 description (Experimental Dataset). In the second dataset, the AEPC score at each time point (i.e., pre, mid, and post-semester) is the sum of 4 items (as described below) that ranges from 0-22. The stems of each item are common across the time points, but the particular questions are unique to each time point.

Differences between dataset 1 and dataset 2. Three of the four items from dataset 2 result from a revision in the scoring approach applied to the 6 items from dataset 1. Items and rubrics were combined because I found that they were providing similar information. I included a new item about learning outcomes in dataset 2 (Item 2). For the mid-semester time point, I kept dataset 1's item i1a in dataset 2. I combined dataset 1's items i8a and i10a into one item named item 3 for dataset 2. I also combined dataset 1's items i11a and i12a into another item called item 4 for dataset 2. The combined items 3 and 4 were each scored from 0-7 based on the rubric for the original items. I dropped item itc in dataset 2 because it did not make sense to ask instructors to provide pre- and post-assessments at the beginning of the semester, as they had not served in their role long enough to design these assessments. These instructors were taking a pedagogy course for first-time instructors. See Appendix C for the items. In Table 6.1, I show the relationship between the items for the mid-semester AEPC assessment in the two datasets:

Table 6. 1

Comparison of items and scales for mid-semester AEPC assessments for 2 data sets

Dataset 1: AEPC Initial Validation		**Dataset 2: Instructor Growth in AEPC over time**	
Item Label	**Score Range**	**Item Label**	**Score Range**
Item i1a	0-7	Item 1	0-7
N/A	N/A	Item 2	0-1
Item itc	0-1	N/A	N/A
Item i8a	0-7	Item 3	0-7
Item i10a	0-7		
Item i11a	0-7	Item 4	0-7
Item i12a	0-7		
Total			
6 items	0-36	4 items	0-22

6.1 Dataset 2 Summary Statistics

In the treatment and control pedagogy courses, I collected data on the characteristics of the instructor, which is also a part of dataset 2. There are 129 instructors in dataset 2. Each instructor has 3 repeated AEPC measures. There are no missing data on the assessments. This creates a total of 387 observations.

Associated with the instructor are the following covariates: instructors' gender, URM status, assignment to treatment or control, years of teaching experience, years of tutoring experience, and number of students. Associated with each observation for each instructor is a time point (i.e., pre, mid, post). The variables for these analyses were defined in chapter 4. I will provide the summary statistics of these variables below.

6.1.1 Descriptive Statistics

Table 4.1 shows that treatment (n=65) and control (n=64) groups have similar gender and URM characteristics. The control has about 31% women and 11% URM, and treatment has 28% women and 9% URM. There are no missing data on these variables.

Table 4. 1

Descriptive Statistics for Gender and URM Status by Control and Treatment Condition

Covariate		**Control** (n=64) Frequency (%)	**Treatment** (n=65) Frequency (%)
Gender	Female	20 (31.25%)	18 (27.69%)
	Male	42 (65.63%)	46 (70.77%)
	Gender Non-Conforming	2 (3.13%)	1 (1.54%)
URM status	URM	7 (10.94%)	6 (9.23%)
	Not URM	57 (89.06%)	59 (90.77%)

In Table 6.2, I present summary statistics, such as means and standard deviations (SD), for other continuous predictor variables by condition: years of teaching experience, years of tutoring experience, and number of students. Both conditions are comparable on these covariates.

Table 6. 2

Average years of teaching experience, years of tutoring experience, and number of students by control and treatment conditions

Covariates	**Control (n=64) Mean (SD)**	**Treatment (n=65) Mean (SD)**
Years of teaching experience	0.92 (1.00)	0.86 (0.95)
Years of tutoring experience	2.23 (1.67)	2.47 (2.05)
Number of students	4.83 (2.59)	4.08 (2.36)

6.1.2 AEPC Assessment Variable

As described in chapter 4, I developed an AEPC assessment on a scale of 0-22 for answering research question 1. In Table 6.3, I present summary statistics, such as means and standard deviations (SD), for my AEPC test variable which includes the AEPC assessment scores from all 3 time points (i.e., all pre-, mid-, and post-AEPC assessment scores).

Table 6. 3

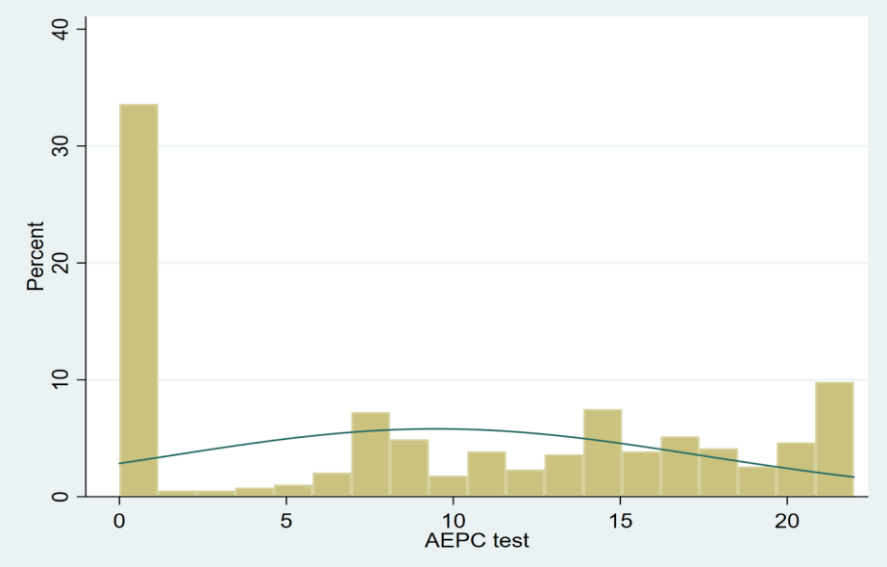
Summary Statistics for the AEPC Assessment Variables

Variable	Frequency	Mean	Standard Deviation
AEPC test	387	9.49	7.96 (Overall)
			3.19 (Between)
			7.29 (Within)

There were 129 instructors in the sample. The test outcome variable has a mean of 9.49 points with an overall standard deviation of 7.96. This time-varying variable has a between-instructor standard deviation of 3.19 and a within-instructor standard deviation of 7.29. The lowest score on the test outcome variable was 0 points, and the highest score was 22 points. Figure 6.1 displays a histogram of the AEPC test variable with AEPC assessment scores from all three time points. One can see a significant portion of scores near zero due to the floor effect on the pre-AEPC assessment and the second-highest frequency of scores near 22, which come from the post-assessment. Most of the remaining scores were distributed fairly evenly across a range from 7 to 21. These remaining scores come from the mid-semester and post-semester AEPC assessments. For additional context on the floor effect, the high number of zeroes may be because many instructors were not ready to perform well on the pre-AEPC assessment, since this was a pedagogy course designed for first-time instructors. To address the lack of normality in the distribution of the AEPC test variable, I use robust standard errors in my regression models for answering RQ 1. A regression with robust standard errors based on a sandwich estimator was performed, since they are more robust to violations of the model assumptions.

Figure 6. 1

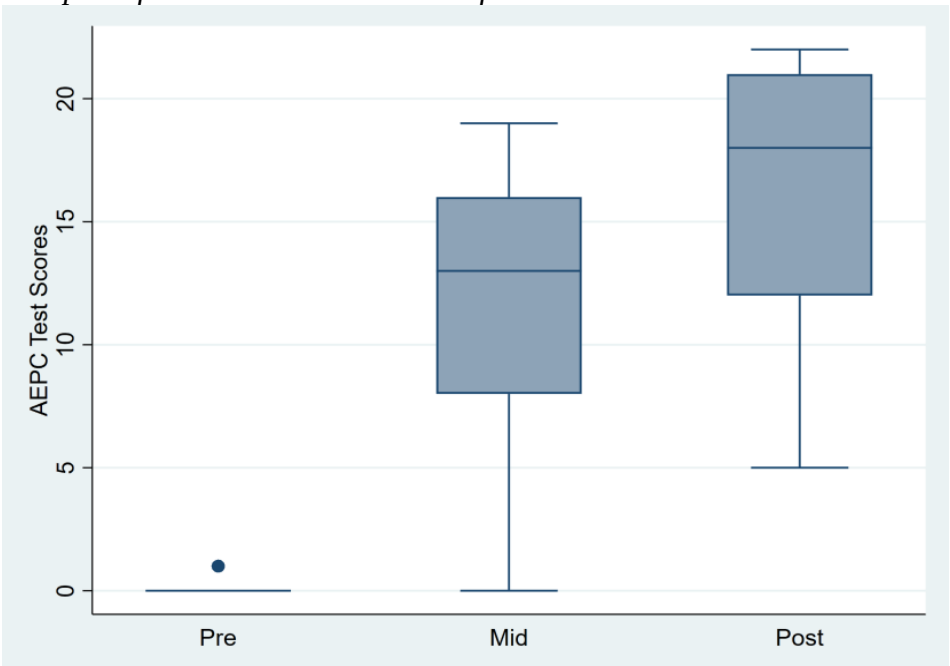
Histogram of AEPC Test Variable for Combined Treatment and Control Groups



In Figure 6.2, I provide box plots (i.e., displaying the minimum, lower quartile, median, upper quartile, and maximum) to represent the distribution of AEPC test scores in the combined treatment and control groups at pre-, mid-, and post-semester assessments. As described above, there is less variability at the pre-semester due to the high number of instructors who scored a 0 on the AEPC pre-semester assessment. Variability in the AEPC test scores is larger at the mid-semester and post-semester time points.

Figure 6. 2

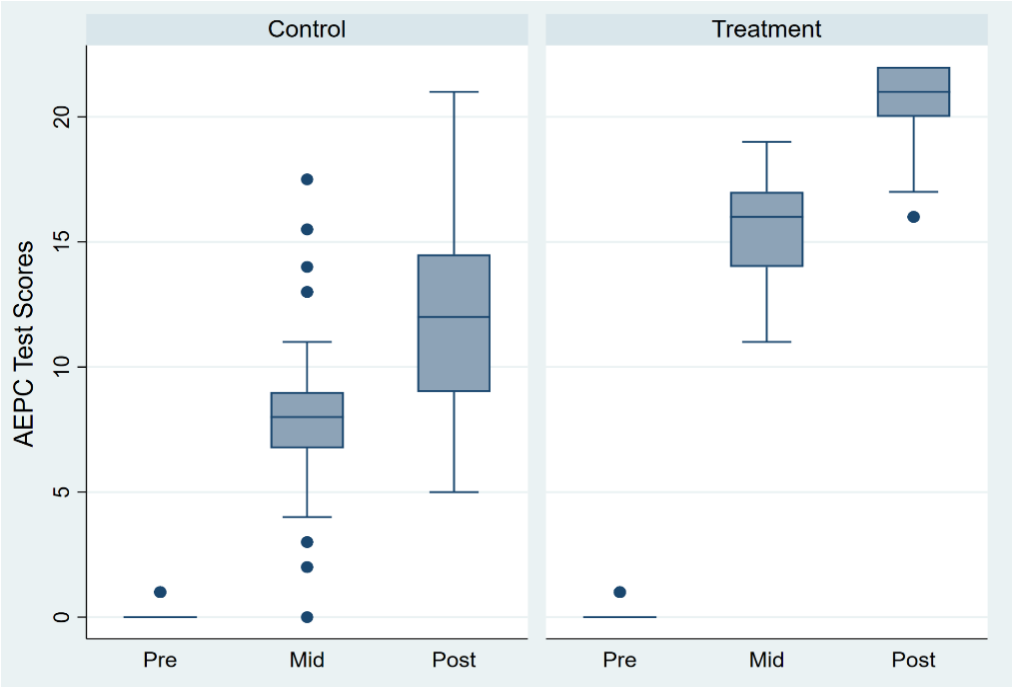
Boxplot of AEPC Scores Over Time for Combined Treatment and Control Groups



In Figure 6.3, I provide the corresponding box plots of the AEPC test scores separated by control and treatment condition. This figure illustrates how the treatment condition has AEPC test scores that have generally higher values and less variability at the mid-semester and post-semester time-points when compared to the control condition.

Figure 6. 3

Boxplot of AEPC Scores Over Time by Condition



In Table 6.4, I present statistics on AEPC measures over time for treatment and control conditions. The treatment and control conditions were comparable because they had no statistically significant differences on baseline AEPC (p>0.05). As shown in Table 6.3, these results indicate that although both conditions experienced AEPC test score increases over time, the treatment had greater AEPC increases, on average. The treatment outperformed the control on the post-AEPC assessment by about 8.38 points.

Table 6. 4

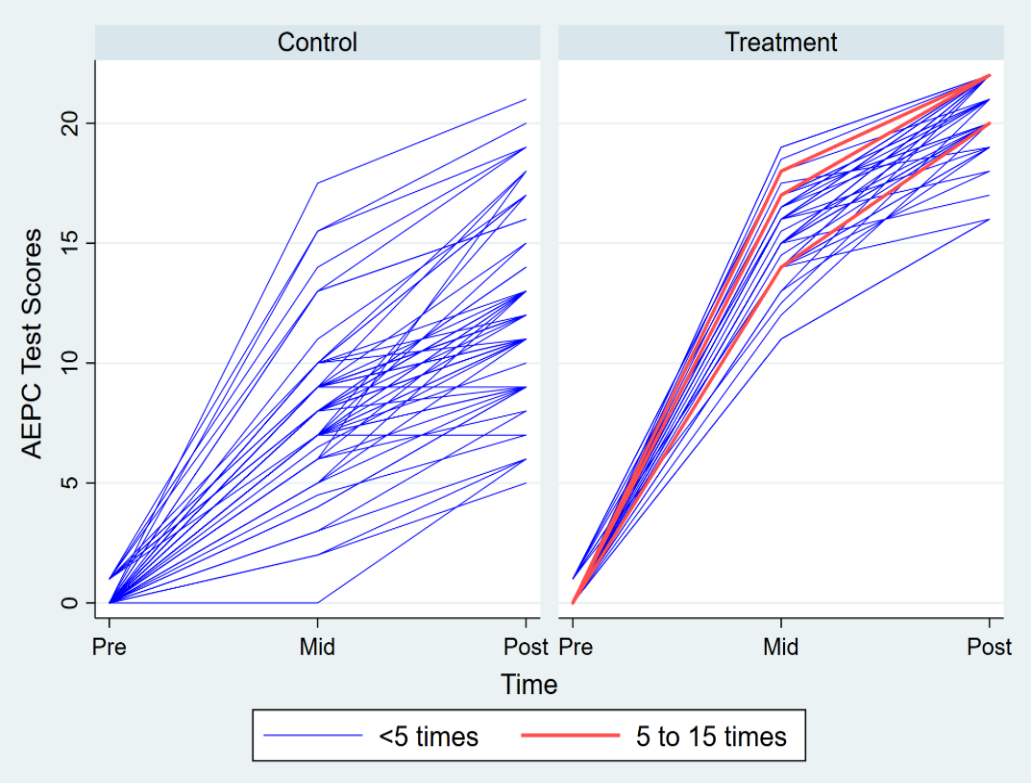
Average AEPC Scores and Standard Deviations (SD) Across Time by Control and Treatment Conditions

AEPC Assessment (0-22 scale)	**Control (n=64) Mean (SD)**	**Treatment (n=65) Mean (SD)**
Pre (Time 0)	0.16 (0.37)	0.12 (0.33)
Mid (Time 1)	7.99 (3.20)	15.65 (1.87)
Post (Time 2)	12.27 (3.81)	20.65 (1.61)

I provide a graphical illustration of these AEPC scores over time by condition in figure 6.4. In Figure 6.4, the lines represent individual instructors' AEPC test scores over time. Since the AEPC scores are discrete on a 22-point scale, any line segment may represent multiple instructors. Each blue line depicts a pattern across pre-, mid-, and post-AEPC that occurs less than 5 times. Each red line represents a pattern that occurs between 5 and 15 times.

Figure 6. 4

AEPC Test Scores Over Time by Condition



6.2 Regression Analyses

To more formally assess the research questions in light of the evidence I collected, I performed a regression analysis for each. I fit the models with random intercepts described in my methods section to the datasets described above, resulting in parameter estimates that quantify each effect and *p*-values that indicate their statistical significance. In Table 6.5, I present my model estimates and interpret the findings at the 5% significance level to address RQ 1.

To aid in the assessment of practical significance, I also compute the corresponding coefficients when the response variable is expressed in standard deviation units. Specifically, I divide the regression coefficient by the standard deviation of the control group's mid-semester AEPC scores from RQ 1. These scores are denoted SD_{midC}. I selected the standard deviation of the control group's mid-semester AEPC scores (SD_{midC}) because I am interested in how much the treatment group improved relative to the control group. I did not select the pre-AEPC

assessment's standard deviation because there was not much variance and doing so may overstate the practical significance.

6.2.1 Research Question 1

Table 6. 5

Regression Table for Model with Robust Standard Errors for Research Question 1

Model 1 Parameters	**Estimate (Robust Standard Error)**	**Estimate in SD_{midC} Units (Robust Standard Error)**
Fixed Part		
Time Point 1	7.84*** (.39)	2.45*** (.12)
Time Point 2	12.11*** (.46)	3.78*** (.14)
Treatment Group	-.12 (.11)	-.04 (.03)
Treatment Group × Time Point 1 ($\hat{\beta}_4$)	7.69*** (.45)	2.40*** (.14)
Treatment Group × Time Point 2 ($\hat{\beta}_5$)	8.41*** (.51)	2.63*** (.16)
Female (Male= reference)	-.46 (.34)	-.14 (.11)
Gender non-conforming	-1.89 (1.24)	-.59 (.39)
URM status	-.14 (.38)	-.04 (.12)
Number of students	-.06 (.05)	-.02 (.01)
Teaching experience	.15 (.15)	.05 (.05)
Tutoring experience	-.01 (.08)	-.00 (.02)
Intercept	.56 (.34)	.17 (.11)

Random Part		
$\sqrt{\psi}$	1.36	.42
$\sqrt{\theta}$	1.76	.55

Robust standard error in parentheses
*p<0.05, **p<0.01, ***p<0.001

Research question 1. Table 6.5 presents parameter estimates with standard errors and asterisks indicating significance at the 5%, 1%, and .1% levels. I examine the difference in growth between the two conditions from pre to mid ($\hat{\beta}_4$), and pre to post ($\hat{\beta}_5$) AEPC. On a scale of 0-22, the increase in mean AEPC test scores from pre- to mid-semester is estimated to be 7.69 points (95% CI [6.80, 8.57]) higher in the treatment than the control, controlling for the instructors' gender, URM status, number of students, years of teaching experience, and tutoring experience (z=17.03, p<0.001). As shown in Table 6.5, this estimate can be expressed as 2.40 standard deviation units. Moreover, the increase in mean AEPC test scores from pre- to post-semester is estimated to be 8.41 points (95% CI [7.42, 9.40]) higher in the treatment than the control, controlling for variables mentioned above in the model (z=16.65, p<0.001). As shown in Table 6.5, this estimate can be expressed as 2.63 standard deviation units. Overall, the treatment by time interaction is statistically significant in the model (X^2(2)=327.80, p<0.001).

6.2.2 Discussion for Research Question 1 Findings

In chapter 3, I described how my conceptual framework reflects the equitable practices that I believe promote instructor learning. In my conceptual framework, the facilitators model AEP and adapt culturally responsive pedagogies (CRP) to train instructors in treatment. I hypothesized that this process would promote greater reflection in the treatment, which I expected to increase instructor learning over time in the treatment relative to the control as measured by AEPC assessments. In this chapter I found evidence that supports this hypothesis. The findings for RQ 1 suggest that there is a very large effect of the treatment on the AEPC scores over time.

Moreover, the instructors in the control condition, on average, experienced large growth on AEPC from pre- to post-semester that is consistent with prior literature, because the control group employed exemplary educational development activities (e.g., reflection, experiential learning). As described in chapter 4, the experimental control group included educational development activities such as reflection and dialogue because these practices have been shown to improve instructors' adoption of equitable teaching practices (Harrison-Bernard et al., 2020; Metevier et al., 2010; O'Leary et al., 2020; Parker et al., 2016; Rooney et al., 2020).

In chapter 2, I identified five studies that examined pre- to post-gains in college instructors' equitable teaching competencies within professional development programs in STEM settings (Harrison-Bernard et al., 2020; Metevier et al., 2010; O'Leary et al., 2020; Parker et al., 2016; Rooney et al., 2020). Comparing my research findings with the findings of these five studies faces several limitations. First, these studies conceptualized and measured equitable teaching competencies differently from my study and come from different contexts. Second, none of the studies had a control group and therefore none provided a comparison for the standardized effect sizes that I calculated for my study.

6.2.3 Limitations

I would like to note the modest sample size of URM and gender non-conforming instructors in both conditions. In addition, the number of students is low in this sample because discussion sections are optional. As a result, detecting statistical significance for main effects and interactions involving these covariates would be more difficult. Having a larger dataset would address this issue in the future.

In addition, there are limitations to generalizability of the research due to the specific population of instructors and the context in which the study was conducted. For example, the pedagogy courses served first-time instructors who a) had relatively little teaching experience, b) demonstrated low average baseline competency with equity-oriented practices, and c) primarily taught in computer science, data science, statistics, and computational disciplines, which have a unique culture at the R-1 university in this study. Therefore, I believe that conducting further research in other contexts, colleges, and universities would be warranted.

6.2.4 Conclusion

In this chapter, I found that educational development interventions using AEP can significantly increase instructors' equitable teaching competencies, on average. In particular, I observed that instructors achieved significant growth, on average, in developing equitable pedagogies even though they all happened to start with a low baseline competency in AEPC. These findings suggest that the interventions using AEP can support instructors who are initially resistant and are not already competent with equitable teaching practices. Such interventions are worthwhile since they can promote instructors' learning and adoption of equitable teaching practices over time. Therefore, training instructors in AEP can be effective for creating more welcoming and inclusive learning environments.

CHAPTER 7: Qualitative Findings

In chapter 6, I observed variation in the extent to which instructors learned or adopted AEP elements and techniques. Next, I turn to qualitative evidence to illuminate the mechanisms that may help to explain this variability.

In chapter 7, I focus on answering research question 2: What mechanisms for instructional change characterize differences between low- and high-AEP growth instructors in the treatment? To do so, I address conjectures that I discussed in chapter 4 which are briefly summarized below:

1. **Authority and Community Lens.** I hypothesized that high-growth instructors would cite other instructors, peers, or literature when describing the reasons they adopt adaptive equitable practices.
2. **Equity Perspectives and Definitions Lens.** I imagined that high-growth instructors would articulate definitions of equity that address students' different challenges and backgrounds.
3. **Learning Goal Lens.** I envisioned that high-growth instructors would incorporate both academic and equity goals to guide teaching and learning practices.
4. **Emotional Lens.** I anticipated that the high-growth group would be more motivated by their emotions of how students perform in their class because they would be adapting their teaching based on students' social-emotional and academic needs.
5. **Data Lens.** I hypothesized that high-growth instructors will use and acknowledge the validity of both quantitative and qualitative student data when adapting teaching.
6. **Responsibility for Learning Lens.** I imagined that the high-growth group would be more motivated to identify learning barriers and adapt teaching to address them because these instructors put the responsibility for learning on themselves, the student, and other contextual factors.

An explanation of the coding framework and its application to participant reflections is available in Chapter 4.

7.1 Analytical Approach

As mentioned in chapter 4, I drew from instructors' final teaching reflections to identify the mechanisms that motivate instructors to learn and adopt Adaptive Equity-Oriented Pedagogy (AEP) in higher education. These teaching reflections involved extended responses related to instructors' pedagogical beliefs, practices, and evidence of applying AEP Competency (AEPC). Specifically, 10 low-growth and 10 high-growth AEPC instructors, as determined by the quantitative analysis, were selected from the treatment condition. To accomplish this goal, the bottom 25% (n=16) and the top 25% (n=16) of the 65 instructors were identified by the magnitude of the change in their AEP competency from pre- to post-semester. Next, all URM instructors in both of these groups of instructors were selected for inclusion. Finally, a different member of the research team randomly sampled additional instructors to create sets of 10 in the low- and high-growth groups. The 10 low- and high-growth instructors were not notified that they were designated as low- and high-growth instructors because this analysis was conducted post-hoc; however, they had previously granted consent to use their reflections for research purposes. From these reflections, I selected 11 specific questions for analysis. Appendix C

provides the selected reflection questions that constituted the dataset for RQ2. I coded the instructors' extended responses to these 11 questions related to their pedagogical beliefs and practices.

On a scale of 0-22, the mean low-growth group AEPC growth score is 18.31. One taught statistics and the others taught computer science. Two were female and eight were male. One identified as URM. By contrast, the mean high-growth group AEPC growth score is 22. Six taught computer science, three taught electrical engineering and computer sciences, and one taught data science. Four were female and six were male. Three identified as URM. Recall that instructors' gender and URM status were not significant predictors of AEPC growth in the multilevel regression model for research question 1. I do not provide subcode analysis by instructors' demographic groups or by the course that individual instructors taught, and use pseudonyms in order to protect privacy.

In the descriptive statistics for each subcode, I provide the number of instructors whose responses indicated each subcode at least once, for both the low- and high-growth groups. I also present the total count of the subcode, which refers to how often the subcode appears across all 11 questions and all 10 instructors in each group. I also provide the mean and median presence of subcodes for each group. I provide these descriptive statistics for all the subcodes that are not zero for both the low- and high-growth groups. I also examined, where appropriate, whether the subcodes appeared in questions 1-2 (which ask about instructors' philosophies and practices) and questions 3-18 (which focus on other aspects of teaching) because I want to see if the subcode is mentioned consistently throughout the reflection.

In the findings below, I highlight subcodes for which there were differences of 3 or more in either the number of instructors in each group who cited the subcode, or in the median frequency with which the subcode is mentioned by participants in each group. I use this "relevance criteria" for determining the subcodes that I describe and analyze in more detail below. I also identify subcodes that are highly represented within both groups. I then explore whether these subcodes emerged differently among the groups. For the examples presented for each lens, I select the focal participants by identifying instructors whose patterns of responses were representative of the key differences between low- and high-growth groups.

7.2 Findings

7.2.1 Authority and Community Lens

Research suggests that departmental and disciplinary cultures influence the adoption of equitable and inequitable teaching practices (Kezar & Bernstein-Serra, 2020; Phuong et al., 2021). However, I did not find strong evidence that low- and high-growth instructors exhibited major differences in how they invoked the Authority and Community Lens. In fact, most subcodes were fully absent from the majority of instructor responses on the assessments. Only two subcodes were reflected at all in instructor responses: "[Colleagues] value this teaching practice in my discipline" and "Something in the department constrains my teaching effectiveness." Since the former subcode was only mentioned by one participant, I focus on the latter subcode which meets my relevance criteria for describing a subcode in more detail. Table 7.1 below shows descriptive statistics for the Authority and Community Lens.

Table 7. 1

Presence of Authority and Community Lens Subcodes Between High- and Low-Growth Groups

	High-Growth (HG) (n=10)				Low-Growth (LG) (n=10)			
Subcode	# of Instructors	Total	Mean	Median	# of Instructors	Total	Mean	Median
• [Colleagues] value this teaching practice in my discipline	0	0	0	0	1	1	0.1	0
• Something in the department constrains my teaching effectiveness	9	9	0.9	1	6	6	0.6	1

Six instructors in the LG group and nine in the HG group represented the subcode: "Something in the department constrains my teaching effectiveness." All of the responses that were identified for this subcode were for reflection question 18 which asks, "What features of the course design (e.g., syllabus schedule, teaching strategies, formative assessment, alignment of instruction and graded assessments) do you think hindered student success? Why? Please cite any evidence." I believe this to be the case because this question allows instructors to reflect on the department's constraints to their teaching. It is important to note that many of the instructors are graduate and undergraduate student instructors who have limited control over the syllabus and course policies.

While the quantitative difference between low- and high-growth instructors who reflected this code is not large, there were some qualitative differences in how these two groups described the constraints they encountered. Instructors in the low-growth group reflected mostly on the teaching and assessment practices of the courses they were supporting. They provided examples which include the lack of alignment between instruction and assessment, a rigid syllabus schedule, and inadequate support from course staff (e.g., TAs lack the ability to support students with the course content and assessments). For example, one LG instructor, Joaquin, wrote that "A couple students, for example, expressed frustration with the professor for not providing enough information about two dimensional arrays […] I reminded them that I can only help them with conceptual material that may enable them to better understand the project and that I couldn't even look at their code." Similarly, all of these instructors from the LG group cited faculty members' course practices that were embedded in departmental structures and thus outside of their control. Here is an example of statements from LG instructors Jeremiah and Sadie, respectively: "the rigid syllabus schedule can hinder students, because the difficulty level is not easy to predict" and "long projects hinder student success."

Instructors in the HG group cited many of the same issues, but tended to mention *more* such issues and cited the need to provide more well-structured courses in their responses. For

example, Sally states, “I think that there weren’t enough resources for students struggling with the projects in this course. Many of my students complained about long office hour queues and a lack of support in this area.” Other high growth instructors cited additional structural elements of the course. For example, Maria noted that “The course is broken into three disjoint modules, which can be a little bit confusing for students (and is one of the most common complaints I’ve heard).”

Despite these subtle differences in content, my analysis of responses across groups does not provide much evidence for my conjecture that high-growth instructors would have more responses that reflect the subcode “I mimic esteemed equity-oriented instructors.” I did not find sufficient evidence that the influence of authoritative figures, departmental communities, and disciplinary cultures drove differences between low- and high-growth instructors on AEPC.

7.2.2 Equity Definitions and Perspectives Lens

My emerging research has suggested that instructors' conceptions of equity are likely to influence their adoption of equitable teaching practices (Phuong et al., 2021). I found evidence of differences between low- and high-growth instructors for this set of codes. Three subcodes meet the relevance criteria: “I want all my students to succeed/ I address equity issues,” (10 HG vs 5 LG) “Equity is adapting to address student needs and their socio historical context,” (10 HG vs 5 LG) and “Equity is giving everyone the same thing (i.e., equality).” (0 HG vs 3 LG). The rest of the subcodes for this lens were not represented by any instructors. Table 7.2 provides the descriptive statistics of the Equity Definitions and Perspectives Lens Subcodes between high- and low-growth groups.

Table 7. 2

Presence of Equity Definitions and Perspectives Lens Subcodes Between High- and Low-Growth Groups

	HG (n=10)				**LG (n=10)**			
Subcode	# of Instructors	Total	Mean	Median	# of Instructors	Total	Mean	Median
• I want all my students to succeed/ I address equity issues	10	46	4.6	5	7	23	2.3	1
• Equity is adapting to address student needs and their socio-historical context	10	63	6.3	6.5	6	23	2.3	1.5

• Equity is giving everyone the same thing (i.e., equality)	0	0	0	0	3	4	0.4	0

Two subcodes related to equity definitions demonstrated similar patterns in responses among both High-Growth and Low-Growth instructors. The subcodes **"**I want all my students to succeed/ I address equity issues" and "Equity is adapting to address student needs and their sociohistorical context" were represented in the early responses (Q1 and Q2) of all ten high growth instructors. It was important to examine the presence of the equity definitions subcodes in questions 1 and 2 because these questions ask about these definitions. I then look at whether these definitions are mentioned consistently throughout subsequent responses to see if the definition is acting as a lens on their teaching. Based on my qualitative analyses, these subcodes were represented in half of the early responses of low-growth instructors (six instructors in the low-growth group mentioned adapting to students' needs, and seven mentioned wanting all students to succeed). Both codes were also represented more often in the responses of high growth instructors across all questions (*want all students to succeed*: HG=46; LG=23; *equity is adapting to needs and context*: HG=63; LG=23).

For the subcode, "I want all my students to succeed/ I address equity issues", the ten high-growth instructors had a range of invoking the subcode of 3 to 7 times, where 5 is the median. By contrast, the seven low-growth instructors had a larger range of invoking subcodes of 1 to 8 times with a median of 1.

For the subcode, "equity is adapting to needs and context", the ten high-growth instructors had a range of invoking the subcode of 5 to 8 times, where 6.5 is the median. By contrast, the six low-growth instructors had a larger range of invoking subcodes of 1 to 8 times with a median of 1.5.

Both of these codes also emerged in different parts of the reflections for the high-growth versus low-growth group. Among all high-growth participants, these codes were each represented at least once in later reflection questions about specific classroom practices including formative assessment and evidence of student success (Q11, Q15, or Q16). The low-growth group, in contrast, rarely reflected these codes in those questions. Only three participants in the low-growth group explicitly mentioned wanting all students to succeed in these subsequent questions, and only four participants in the low-growth group explicitly mentioned adapting to address student needs. It is also of note that the same 3-4 participants in the low-growth group were responsible for most mentions of success and equity beyond the initial two questions.

There is a weaker, but related pattern when exploring instances of the subcode "Equity is giving everyone the same thing (i.e., equality)". This code was reflected in the responses of three low-growth and no high-growth instructors. However, even those instructors who defined equity as providing the same opportunities to students, described their practices as adapting their teaching. This apparent contradiction may indicate some confusion in how low-growth instructors conceptualized equity broadly in terms of their definitions and practice. None of the instructors in the low- or high-growth groups mentioned the other subcodes related to equity.

7.2.2.1 Comparing a Low- and High Growth-Instructor

To better understand how these dynamics emerged in instructor responses, here I present two brief examples that contrast the typical responses of a low-growth and a high-growth instructor, with focus on reflection questions 1, 2, 11, 15, and/or 16.

Low-growth instructor. Jeremiah, an instructor from the low-growth group, expresses a commitment to supporting all students' success and addressing equity. In his response to question 1 (i.e., his teaching philosophy statement), he defines "equity as every student having the resources and opportunities that they need to be successful." Jeremiah states, "Equity is important to me, because I believe that every student can learn regardless of their prior circumstances. Thus, it is important for me to engage with students constantly and adapt to their needs." Jeremiah states how he uses formative assessment and adapts instruction based on student progress. He does not explicitly mention the term "equity" in his response to question 2, however, he talks about adapting his teaching to address barriers to learning and address student needs.

Later, Jeremiah responds to question 15, which asks for qualitative evidence for understanding students' success, with a one sentence response: "I observed that they became more confident in answering questions publicly." Jeremiah follows a similar pattern with his response to Question 16 about evidence for how his teaching impacts students' psychosocial outcomes, "Their confidence grew as they believed in the growth mindset." In his early answers, Jeremiah briefly mentions how equity and prior circumstances are important considerations for adapting teaching practices; however, he does not elaborate on definitions of equity as it relates to students' sociocultural and historical contexts. His later statements do not provide much detail on how he thinks about equity to improve students' growth mindset through pedagogical practice.

High-growth instructor. By contrast, Maria, an instructor from the high-growth group, echoes both the *wants all students to succeed* and *adapting to needs and context* subcodes more consistently than Jeremiah. Similar to Jeremiah's response to question 1, Maria defines equity as "ensuring that every student has a fair opportunity to learn and succeed." However, Maria goes into more details on discussing student demographics: "I understand that not all students have the same background in subjects--some may have more exposure to certain topics because of factors like socioeconomic status, gender, race, or country of origin [...]" In response to question 2, Maria writes how she gathers data and adapts her teaching to support student learning.

> When I teach, I solicit feedback and gauge students' understanding and engagement in order to determine what strategies work best for particular students, reflecting the fundamental "adaptive" nature of AEP. [...] In the classroom, I walk around continuously and frequently check in to see if students have any questions and to ensure that they're on the right track; I realize they sometimes won't ask for help unless I prompt them to, so I try to actively remove that particular obstacle.

Committed to supporting all students' success, Maria notices the equity barriers more than Jeremiah by highlighting how some students have less prior experience, experience stereotype threat, and have to balance a job with classwork.

Furthermore, Maria discusses how she sought to address equity barriers in the classroom by adapting her teaching to ensure participation from all students, not only the most vocal ones. For example, she stated "during lab checkoffs, one student will tend to answer more of the

questions, [...] in these situations, I'll specifically direct a question towards the other student (in a way that doesn't make either of them feel bad), and I find that this is often the most telling question in terms of assessing the students' grasp of the material."

Later, in response to question 15, Maria writes how she used formative assessment to adapt her teaching and improve student achievement. Here she echoes the same themes from question 2, where she talked about how she informally asked students questions and used ongoing formative assessment to meet students where they are at. In response to question 16, Maria provides additional evidence of how she addressed equity issues, such as a limited sense of belonging for students in STEM: "My students seem to be comfortable asking me for help (especially as the course has progressed and I've gotten to know them better), which I think is a sign that they feel that they belong in lab and that I care about their learning." This statement shows how Maria promotes a sense of belonging as a way of supporting all students' learning.

These findings suggest that compared to low-growth instructors, high-growth instructors more often articulated definitions of equity that focused on "I want all my students to succeed / I address equity issues" and "adapting to address student needs and their socio historical context."

7.2.3 Learning Goal Lens

Equity-oriented instructors should be able to align their academic goals with equity goals, so that the two are not at odds. I found evidence that the Learning Goal Lens reflected major differences between low- and high-growth instructors. The "Equity goal" subcode and the co-occurrence of the "Academic goal" and "Equity goal" subcodes meet my relevance criteria. I found that the co-occurrence of the "Academic goal" and "Equity goal" subcodes– which refers to instructors invoking both subcodes in a response to a single question– is different between low- and high-growth groups. Although the "Academic goal" subcode does not meet my relevance criteria, I still provide the descriptive statistics of this code in Table 7.3 below.

Table 7. 3

Presence of Learning Goal Lens Subcodes Between High- and Low-Growth Groups

	HG (n=10)				**LG (n=10)**			
Subcode	# of Instructors	Total	Mean	Median	# of Instructors	Total	Mean	Median
• Academic goal	10	72	7.2	7.5	10	45	4.5	5
• Equity goal	10	49	4.9	4.5	8	12	1.2	1
• Co-occurence of academic and equity subcodes	10	42	4.2	4	6	8	0.8	1

"Equity Goal". Eight low-growth and ten high-growth instructors were identified as mentioning the subcode, "Equity goal". This code was represented more often in the responses of

high-growth instructors across all questions than for the low-growth instructors (*Equity Goal subcode presence:* HG=49; LG=12). The instructors in the HG group provided responses that represented this code about four times more often, on average per participant, than in the LG group (HG mean = 4.9, HG median = 4.5; LG mean = 1.2, LG median = 1).

All ten instructors in HG referenced the subcode "Equity goal" in questions 1-2 and again when talking about assessment practices and using data to inform their teaching in at least one of the following questions: 6, 11, 15, 16. Only three instructors in the LG group similarly referenced the subcode "Equity goal" like the HG group for these later questions. Therefore, HG instructors represent the subcode "Equity goal" in responses throughout the reflection questions, while LG instructors are not as consistent, especially in questions 6-16.

Co-occurrence of Academic and Equity Goals. I also examined the co-occurrence of the "Academic goal" and "Equity goal" subcodes, which refers to instructors invoking both subcodes in a response to the same question. The co-occurrence of the "Academic goal" and "Equity goal" subcodes was present more often in the responses of high-growth instructors across all questions than for the low-growth instructors (*Academic and Equity Goal subcode co-occurrence presence:* HG=42; LG=8). The instructors in the HG group provided responses that represented both subcodes about four times as often, on average per participant, than in the LG group (HG mean = 4.2, HG median = 4; LG mean = 0.8, LG median = 1). All ten high-growth instructors represented both subcodes for questions 1 and 2 as well as for at least one question between questions 3-18. By contrast, five low-growth instructors represented both subcodes for question 1, only two LG instructors represented both subcodes for question 2, and only one LG instructor represented both subcodes for any of the questions 3-18 (only once in question 6).

These findings suggest that high-growth instructors described the integration of academic and equity goals to guide teaching and learning practices more both in their statements about what they notice and believe, as well as in their statements about their actual teaching practices. For high-growth instructors, equity appears to be a consistently applied lens, or frame of reference, and not an add-on or after-thought.

7.2.3.1 Comparing a Low- and High Growth-Instructor

I provide examples of an instructor from the low-growth group and another one from the high-growth group who show differences on the equity goal subcode.

Low growth. In response to question 2 (i.e., the teaching practice statement), Nolan articulates an academic goal for his teaching that focuses on course content from the syllabus. For example, he discusses how he adapts his pedagogy every time he teaches: "I am adjusting my teaching strategy to compensate for [the difficulties associated with large classes] by always asking if there are any questions, and starting off each lesson with me asking what they do/don't know about the topics covered in class."

Later, in response to question 16, Nolan also emphasizes academic goals and content when he describes how a student became more motivated to learn by asking about differences between courses in the computer science program. Nolan states, "One of my students asked me about 61b and how it was different than cs88 because they wanted to try more CS classes so that shows motivation." For his responses to questions 2 and 16, Nolan does not highlight a specific equity goal or barrier.

High-growth. By contrast, Sally calls attention to equity more explicitly, focusing on gender equity in Computer Science and addressing all students' needs. Echoing an equity goal for her teaching in response to question 2, Sally states,

> A majority of the students I've taught this semester have also been women, and some have felt uncomfortable and even looked down upon for being a woman in a course and major dominated by men. This adds an extra layer of self-doubt and insecurity that only negatively impacts their learning. I think being intentional about treating students equitably and with equal respect regardless of their gender, race, religion, sexual orientation, etc., is crucial to overcoming these types of learning barriers.

Sally also echoes both the "Academic goal" and "Equity goal" where she wants students to feel comfortable and adapts her teaching style to address every student's unique background and way of learning the course material. Sally writes that "Every student has a unique background and a unique way of learning and working with material" and that "it often takes working closely with your students and finding ways to assess their progress and adapt to their needs." She describes how she enacts these principles by checking in with students' holistic wellbeing frequently, using formative assessment, and identifying as a specific goal to support students with less preparation who may feel overwhelmed.

In response to question 16, Sally elaborates how she achieved her academic and equity goal for promoting gender inclusion with the academic material. Sally writes about a female student who initially felt like she "wasn't cut out to code", expressed "a lot of self-doubt", and felt like she "couldn't do it." Consequently, Sally adapted her teaching and found that this student became "much more eager to learn and is always willing to try new or challenging problems."

My findings support the conjecture for the Learning Goal Lens in that high-growth instructors tend to report equity and academic goals more often than low-growth instructors, and often together in the same reflection response.

7.2.4 Emotional Lens

My emerging research suggests that emotion can also be a driver for the adoption of equitable teaching in higher education (Phuong et al. 2021). Indeed, the Emotional Lens reflected major differences between low- and high-growth instructors, specifically under the "Negative emotion" subcode for which no low-growth, but five high-growth instructors were represented. Therefore, I focus on the negative emotion subcode since it was the only subcode under this lens that meets my relevance criteria. All other subcodes in this category did not drive significant differences between low- and high-growth instructors, because both groups invoke these subcodes in relatively comparable ways. For example, between one and three participants in each of the high-growth and low-growth groups invoked all but one of the remaining subcodes. These subcodes were all related to positive emotions or the underlying reasons for particular emotions. None of the instructors in either group mentioned the subcode, "Emotions related to teaching online." See Table 7.4 for descriptive statistics of the Emotional Lens.

Table 7. 4

Presence of Emotional Lens Subcodes Between High- and Low-Growth Groups

	High Growth (n=10)				**Low Growth (n=10)**			
Subcode	# of	Total	Mean	Median	# of	Total	Mean	Median

	Instructors				Instructors			
● Positive emotion	3	5	0.5	0	1	1	0.1	0
● Negative emotion	5	6	0.6	0.5	0	0	0	0
● Emotions tied to how students felt in my class or other courses	2	3	0.3	0	2	3	0.3	0
● Emotions are tied to personal learning experience	2	3	0.3	0	1	2	0.2	0
● Emotions related to what I did and how I felt during class	1	2	0.2	0	0	0	0	0
● Emotions are tied to how students did in my class or other courses	1	1	0.1	0	2	2	0.2	0

Negative emotion. The five high-growth instructors expressed themes such as being motivated by their negative emotions or the negative emotions that their students experienced.

High-growth. As shown in the quote from the last section, Sally connects negative emotions to equity and identity. For example, she comments on how women "have felt uncomfortable and even looked down upon for being a woman in a course and major dominated by men." She later discusses how she adapts her teaching to support all of her students' success. She also provides a reflection of how she reduced self-doubt for a female student by validating them and adapting her teaching to ensure they all succeed. These findings suggest that Sally describes how negative emotions are connected to several of the other lenses, including her definitions of equity and her goals for promoting equity in the classroom.

I would argue that my findings cannot fully support or reject the Emotional Lens conjectures. Based on the reflection question responses that I coded, I found evidence that one Emotional Lens subcode, "Negative emotion", reflected notable differences between low-growth and high-growth instructors (according to my relevance criteria).

7.2.5 Data Lens

Equity-oriented instructors draw on multiple data sources to understand and address equity barriers to students' success (Kearns et al., 2018; Phuong et al., 2017; Phuong & Nguyen, 2019; Phuong et al., 2022; Reinholz et al., 2019). I found evidence that the Data Lens reflected major differences between low- and high-growth instructors. Based on my relevance criteria, I

discuss all the subcodes except for the "It depends" subcode. None of the instructors represented this subcode. See Table 7.5 below for descriptive statistics on the presence of Data Lens subcodes between high- and low-growth groups.

Table 7. 5

Presence of Data Lens Subcodes Between High- and Low-Growth Groups

	High Growth (n=10)				**Low Growth (n=10)**			
Subcode	# of Instructors	Total	Mean	Median	# of Instructors	Total	Mean	Median
● I value quantitative data	10	49	4.9	4	7	14	1.4	1.5
● I value qualitative data	10	80	8	8.5	9	33	3.3	3.5
● Formative assessment is important	10	62	6.2	7	9	28	2.8	3
● I believe it is important to use data to measure progress towards learning outcomes	10	38	3.8	3.5	4	6	0.6	0
● Data about socio historical context and student lived experience is important	10	41	4.1	4	5	12	1.2	0.5
● Uses multiple data sources to inform instruction and advance student learning	10	26	2.6	3	1	2	0.2	0
● Co-occurrence of "I value quantitative data" and "I value qualitative data"	10	46	4.6	4	3	3	0.3	0

Nearly every subcode was represented more strongly in the high-growth group (in fact, all Data Lens subcodes were mentioned at least once by all ten participants) versus the low-growth group. The low-growth group has more variability across subcodes. For the two subcodes

"I value qualitative data" and "Formative assessment is important," nine instructors mentioned them but for other subcodes, only one instructor mentioned them.

Three subcodes related to the Data Lens demonstrated similar patterns in responses among both High-Growth and Low-Growth instructors. All ten high-growth instructors represented the subcodes "I value quantitative data", "I value qualitative data", and "Formative assessment is important". These subcodes were also represented by most of the low-growth instructors (7 instructors in the low-growth group mentioned "I value quantitative data", nine mentioned "I value qualitative data", and nine said "Formative assessment is important"). The total presence, mean, and median of the three subcodes were represented more often in the responses of high-growth instructors across all questions (*I value quantitative data*: HG total = 49, HG median= 4; LG Total=14, LG median = 1.5; *I value qualitative data*: HG=80, HG median= 8.5; LG=33, LG median= 3.5; *Formative assessment is important*: HG=62, HG median= 7; LG=28; LG median= 3;. These findings suggest that low- and high-growth instructors agree that quantitative data, qualitative data, and formative assessment are important.

"I believe it is important to use data to measure progress towards learning outcomes". Four low-growth and ten high-growth instructors were identified as mentioning this subcode.

"Data about socio historical context and student lived experience is important". Five low-growth and ten high-growth instructors were identified as mentioning this subcode.

Ten instructors in HG referenced the subcode, "Data about socio historical context and student lived experience is important," in questions 1-2 and seven instructors represented this code in at least two of the following questions: 3-18. By contrast, three LG instructors referenced "Data about socio historical context and student lived experience is important" in both questions 1-2. Only one instructor in the LG group referenced "Data about socio historical context and student lived experience is important" in at least two of the following questions: 3-18.

"Uses multiple data sources to inform instruction and advance student learning". One low-growth and ten high-growth instructors were identified as mentioning this subcode. All ten high-growth instructors represented the subcode for both questions 1 and 2 (like for the academic and equity subcode co-occurrence). Moreover, six HG instructors represented the subcode once for question 3. By contrast, one low-growth instructor represented the subcode for question 1 and 2. The presence of this subcode was not represented anywhere else for LG instructors.

7.2.5.1 Co-occurrence of Subcodes

Co-occurrence of "I value quantitative data" and "I value qualitative data". I also examined the co-occurrence of "I value quantitative data" and "I value qualitative data", which refers to instructors invoking the two subcodes in a response to a given question. Ten high-growth and three low-growth instructors were identified as representing both subcodes. The co-occurrence of these two subcodes was present more often in the responses of high-growth instructors across all questions than for the low-growth instructors (*"I value quantitative data" and "I value qualitative data" subcode co-occurrence presence:* HG=46; LG=3). As most LG instructors did not have the co-occurrence of these two subcodes, the mean and median subcode presence are lower for LG group compared to HG group (HG mean = 4.6, HG median = 4; LG mean = 0.3, LG median = 0).

Five high-growth instructors represented both subcodes for question 1, and all ten high-growth instructors represented the subcodes for question 2. Moreover, all ten HG instructors

represented both subcodes at least twice between questions 3-18. By contrast, one low-growth instructor represented both subcodes for question 1 and two low-growth instructors represented both subcodes for question 2. The co-occurrence of subcodes was not represented anywhere else for LG instructors.

Based on these findings, I found evidence that the Data Lens reflected major differences between low- and high-growth instructors. The subcode, "Uses multiple data sources to inform instruction and advance student learning," represented the largest differences between the LG and HG groups.

7.2.6 Responsibility for Learning Lens

Responsibility for Teaching involves a teacher's responsibility to create a syllabus, assessments, and to teach a class. Instructors who divorce the responsibility for teaching from the responsibility for learning may think they are responsible for teaching and the students are responsible for learning. This approach can promote a fixed mindset around teaching in the sense that there is nothing more that the instructor can do to improve their teaching and students' learning. It is on the student if they failed the exam.

By contrast, an instructor who shares the responsibility for learning with the student has a growth mindset because they believe their teaching practices and students' learning can grow if they work in partnership with their students. Here, sharing responsibility requires an instructor to be responsive to students' learning needs and performance. This approach takes into account students' learning differences and learning in partnership with students to respond with practices that intentionally address students' needs and equity barriers. These instructors seek to notice patterns in student learning, reflect on their practices, identify breakdowns in their teaching, and adapt their teaching accordingly.

My research suggests that instructors who share the responsibility for learning are more likely to adopt equity-oriented practices, since these instructors play a more proactive role in supporting student success (Phuong et al., 2022). Unlike other sections, I share my definition for sharing the responsibility because it constitutes three subcodes with particular definitions that would aid the reader in understanding the rest of this chapter. Sharing the responsibility for learning means that the following subcodes co-exist in an instructor's philosophy and practice:

- The "It's up to me" subcode refers to instructor responses that suggest instructors themselves should take responsibility to mitigate the impact of these barriers on student success. The instructor believes in their responsibility to proactively understand their students' questions, barriers, needs, interests, and experiences. As a part of their role, they believe that it is important to proactively address student learning needs and equity barriers.
- The "It's on my student" subcode refers to instructor responses that suggest the student should play a proactive role in advancing their own learning.
- The "It's something else" subcode refers to instructor responses that indicate another contextual factor (e.g., equity barrier, stress on the community or family) plays a role in impacting student learning.

Thus, to clarify, instructors who represent all three codes of the Responsibility for Learning Lens recognize the importance of playing a proactive role in understanding and addressing students' learning challenges and barriers. One fundamental way in which instructors can enact these responsibilities is by collecting multiple data sources to understand and address

student learning needs. Therefore, the Responsibility for Learning Lens may be related to the Data Lens if instructors are connecting their beliefs to practice, because effectively responding to student learning needs is dependent on understanding and addressing these needs through multiple data sources (e.g., observations, formative assessments, surveys). Accordingly, I am interested in whether instructors who share responsibility (i.e., they represent all three codes) are more likely to be motivated to gather multiple data sources to understand and address learning barriers by adapting their teaching. In this context, I envision sharing the responsibility for learning being more of a mindset and "Using multiple data sources to inform teaching and advance student learning" being an action.

Table 7.6 provides the descriptive statistics on the Responsibility for Learning Lens. "It's up to me," "It's something else," and "co-occurrence of the three subcodes" meets my relevance criteria. Therefore, I focus on these subcodes in my narrative below. Nevertheless, in the table below, I still provide the descriptive statistics for the "It's on my student" subcode.

Table 7. 6

Presence of Responsibility for Learning Lens Subcodes Between High- and Low-Growth Groups

	HG (n=10)				**LG (n=10)**			
Subcode	# of Instructors	Total	Mean	Median	# of Instructors	Total	Mean	Median
• It's up to me	10	86	8.6	9	10	45	4.5	4
• It's on my student	10	53	5.3	4.5	10	55	5.5	5
• It's something else	10	43	4.3	4	8	16	1.6	1
• Co-occurrence of three subcodes: It's up to me, It's on my student, It's something else	10	26	2.6	3	1	2	0.2	0

"It's up to me". All instructors in both groups were identified as mentioning the subcode, "It's up to me". This code was represented more often in the responses of high-growth instructors across all questions than for the low-growth instructors (*It's up to me subcode presence:* HG=86; LG=45). The instructors in the HG group provided responses that represented this subcode about twice as often, on average per participant, than the instructors in the LG group (HG mean = 8.6, HG median = 9; LG mean = 4.5, LG median = 4).

All ten instructors in HG and LG referenced the "It's up to me" subcode when talking about their role in the learning process. Based on Table 7.6, findings suggest that the belief of "It's up to me" is a more consistent perspective or lens for high-growth instructors. In fact, HG

instructors invoke the subcode more often in early questions 1-2 as well as in later questions 3-18.

"It's something else". Eight low-growth and ten high-growth instructors were identified as mentioning the subcode, "It's something else". This code was represented more often in the responses of high-growth instructors across all questions than for the low-growth instructors (*It's on something else subcode presence:* HG=43; LG=16). The instructors in the HG group provided responses that represented this code more often, on average per participant, than in the LG group (HG mean = 4.3, HG median = 4; LG mean = 1.6, LG median = 1).

Like the "Equity goal" subcode, all 10 instructors in HG referenced "It's something else" in questions 1-2 and again in at least two of the following questions: 3-18. By contrast, only one LG instructor referenced "It's something else" in both questions 1-2. This is important because it demonstrates that HG instructors consistently consider other factors that are responsible for and play a role in students' learning, while LG expressed such beliefs less often. For example, three instructors in the LG group referenced "It's something else" in at least two of the following questions: 3-18. This is important because it contrasts how HG instructors consistently invoked "It's something else" in questions 3-18 throughout the reflection questions. The "Equity goal" and "It's something else" subcodes likely have similar statistics because they both address equity topics or issues.

Sharing the Responsibility for Learning: Co-occurrence of "It's up to me", "It's on my student" and "It's something else". I also examined the co-occurrence of sharing the responsibility for learning, which refers to instructors invoking all three subcodes in a response to a given question: "It's up to me", "It's on my student" and "It's something else". The co-occurrence of these three subcodes was present more often in the responses of high-growth instructors across all questions than for the low-growth instructors (*Sharing the Responsibility for Learning subcode co-occurrence presence:* HG=26; LG=2). All ten high-growth instructors represented the three subcodes for questions 1 and 2 (like for the "Academic goal" and "Equity goal" subcode co-occurrence). Moreover, six HG instructors represented the three subcodes once for question 3. By contrast, one low-growth instructor represented the three subcodes for question 1 and 2. The co-occurrence of subcodes was not represented anywhere else for LG instructors.

Co-occurrence between Sharing the Responsibility for Learning and Using Multiple Data Sources to Inform Teaching and Advance Student Learning. For both HG and LG instructors, on a question-by-question level, the three subcodes for the Responsibility of Learning Lens appeared in the same questions in which the Data Lens subcode appeared: "Using multiple data sources to inform instruction and advance student learning". These findings suggest that high-growth instructors have a higher pattern of co-occurrence between sharing the responsibility for learning and "Using multiple data sources" to advance student learning more than the low-growth instructors. This is evident in the high-growth instructors' statements about what they notice and believe as well as in their statements about their actual teaching practices.

I drew from a different reflection task from the beginning of the semester to code responses to the question, "whose fault is it if a student doesn't learn a concept?" Interestingly, none of the LG or HG instructors indicated all three codes that reflect sharing the responsibility for learning at the beginning of the semester when they responded to this question. This finding is consistent with the fact that instructors' pre-AEPC scores were near zero, as mentioned in chapter 6. I am not asserting any causal or predictive claim by pointing this pattern out. However, since the rest of the questions were from the end of the semester, it appears that the

high-growth instructors did adopt the belief of sharing the responsibility for learning, as evidenced by the higher presence of subcode co-occurrence for this lens.

In summary, I found that the low-growth instructors placed the responsibility for learning more often on their students (LG mean = 5.5, median =5) compared to themselves (LG mean= 4.5, median= 4) or contextual factors (LG mean= 1.6, median= 1). By contrast, the high-growth instructors shared the responsibility for learning with themselves, the student, and other factors. The high-growth instructors were more aware, reflective, and responsive to student learning barriers.

7.2.6.1 Comparing a Low- and High Growth-Instructor

I present a brief case study of a low-growth instructor, Derek, whose responses did not include the subcode "Uses multiple forms of data to support student learning" and who does not invoke the "It's on something else" subcode. For contrast, I present a high-growth instructor, Isabella, whose responses indicate they share the responsibility for learning and proactively use multiple forms of data to identify equity barriers and adapt teaching to address them. These cases are typical of the patterns found in the co-occurrence of sharing the responsibility for learning and using multiple data sources between LG and HG instructors. For example, Derek from the LG group had a presence of 0 for the "Sharing the responsibility for learning" and "Uses multiple data sources to inform instruction and advance student learning" subcodes (near the median of 0 and mean of 0.2 for presence in the LG group for both subcodes). Isabella from the HG group had a presence of 3 for the "Sharing the responsibility for learning" and "Uses multiple data sources to inform instruction and advance student learning" subcodes (near the median of 3 and mean of 2.6 for presence in the HG group for both subcodes).

Low-growth instructor. When describing his teaching philosophy in question 1, Derek writes:

> My teaching philosophy is that I always aim to give students the opportunity to come up with solutions on themselves. Computer Science is such a special subject where it is generally hard for students to come up with solutions originally. However, the ability to come up with solutions by themselves is an essential skills not only in dealing with exams but also in the industry. Thus the central idea of my teaching is that I would provide explanations for details and concepts as clearly as possible, but I would only give students hints and guide the students toward solving the problem but keep the actual answer to myself. … Then when it comes to the problems, I would also tell the students that I will only give out the hints but never the code or the exact solution.

In this statement, Derek places the responsibility for learning primarily on the student. He expresses that it is the student's responsibility to succeed in solving original problems. He notes that he does not provide the answer at all to students, but rather provides "hints and guide[s] the students" toward solving problems.

Derek's response suggests that he also assumes some responsibility by indicating that he "provide[s] explanations for details and concepts as clearly as possible." However, providing insufficient scaffolding, or support, can be a particular risk for students who bring negative conceptions about themselves into the learning environment. Although providing time for productive struggle can be effective, keeping "the actual answer to [himself]" and "never" showing "the code or the exact solution" breaks a critical feedback loop for the student, and may

also prevent the student from reflecting on whether their problem-solving strategy was on the right track.

In contrast to many HG instructors, Derek does not mention other sources of learning responsibility ("it's something else" subcode), such as peer collaboration, as a way for students to get the answer. Giving hints and withholding an answer leans towards an individualistic approach of learning, because the instructor places most of the responsibility for learning on students without explicitly mentioning how he targets their learning needs and equity barriers. In the extreme, these types of situations can create a lack of alignment between instruction and assessment, which has been shown to exacerbate attrition as well as equity and opportunity gaps in college STEM courses (Phuong et al., 2022; Seymour & Hewitt, 1997; Weston et al., 2019).

Derek justifies this approach by highlighting the status of Computer Science as a "special subject where it is generally hard for students to come up with solutions originally." He further notes that "the ability to come up with solutions by themselves is an essential skills (*sic*) not only in dealing with exams but also in the industry." Although Derek is not explicitly describing an expectation that students come up with solutions by themselves *without* sufficient support, this expectation is a pervasive part of the STEM teaching and learning culture (Weston et al., 2019). Such an expectation reproduces the social narrative that coming up with solutions by oneself means one is "cut out" for Computer Science and industry, and that students who are not able to fit such a mold "are not cut out to be science majors and should be encouraged to leave the major as early as possible" (O'Leary et al., 2020, p. 10; Weston et al., 2019). Many STEM courses have been historically used to "eliminate less-than-stellar students" which has filtered out many women and minoritized students (Steen, 1987, p. 90; Weston et al., 2019). Weston et al. (2019) found that learning alone without the help needed was one of the top five themes that switchers and persisters in STEM courses most frequently cited to characterize weed-out courses. The most frequently cited themes were assessments that were misaligned with course content and curved grading, which promoted "hostile, competitive classroom cultures" (Weston et al., 2019 p. 201).

While it is true that engineers need to be able to work independently, many engineers in corporate settings collaborate frequently with each other and rely on external resources to succeed (Felder, 2007). In reality, industry teams, engineers, researchers, and other stakeholders collaborate to design impactful technologies and tools for a wide audience, which cannot be fully accomplished by "the ability to come up with solutions by [oneself]" (i.e., working and learning alone). Although Derek mentions communicating hints to students, he does not explicitly describe collaborating with students himself or promoting collaboration among students. Derek's response also did not reference using multiple sources of data to inform teaching. Moreover, in question 1, he did state:

> I am naturally an ally of all of my students because I am not a genius and I have been through exactly the same learning process as they may be experiencing when they study computer science. I personally feel for most of their struggles and concerns. To incorporate more of that in my teaching, I also spend time on chatting with my students about topics ranging from their personal background to their interest in computer science. These conversations will let the students realize that teachers are just people who have been through the learning process a few years earlier than them.

Derek does refer to conversations with students as data points. However, his statements suggest that he perceives his students' learning processes to be the same as his own: "I have been

through exactly the same learning process as they may be experiencing when they study computer science. I personally feel for most of their struggles and concerns." When he hears the students' struggles and their interests, Derek's statement shows how he compares students to himself based on one form of data where he chats with students. He reports that conversations with students are useful for students to understand how teachers like himself "have been through the learning process a few years earlier". In his following statement, he suggests that students follow the previous path of their teachers: "These conversations will let the students realize that teachers are just people who have been through the learning process a few years earlier than them." Derek's statements (e.g., "I have been through exactly the same learning process") do not acknowledge the diverse ways students experience learning, and they suggest that his teaching and conversations with students will reinforce a homogenous learning pathway in STEM.

Based on his written reflection, Derek seems to rely on an assumption of shared personal experiences of students in STEM, rather than collecting additional data and critically reflecting on how students' experiences can differ based on students' background, prior experiences, unique struggles, and equity barriers. I would argue that it is okay to discuss similar or shared experiences. However, when instructors over-identify with students and state they have the same experiences as students, they do not make the space for the possibility that there is variance in students' learning experiences, which is important when addressing student needs. Sharing the responsibility for learning is a reciprocal process and partnership, where instructors work with students to understand how they are learning and what they can do to better support the students' success.

Subsequently, Derek wrote for question 1:

> In the concept breakdown phase of my teaching, there is not much progress assessment involved. However, when it comes to the problem solving part, since students only receive hints along the way, they are making progress by themselves. I constantly recognize and compliment their progress each time they move forward. This serves as a positive feedback when they are making any progress.

This response does indicate that Derek uses qualitative data to adapt teaching. Although Derek compliments student progress, he does not rely on multiple data sources or indicate that his feedback is detailed or tailored to the student.

For question 2, Derek states, "I explicitly ask students to reach out to me when they face problems." This statement corroborates how Derek places the responsibility for learning on the student by asking the student to reach out to him, while also relying mainly on qualitative data sources.

High growth instructor. By contrast, a high-growth instructor, Isabella, was more explicit in her response about how she identifies students' learning barriers and adapts her teaching to address them. She reports drawing on multiple data sources, responding to student learning barriers, and acknowledging the psychological and contextual factors that impact student success.

In her teaching philosophy (i.e., question 1), Isabella, like Derek, shares the learning responsibility with the student by highlighting the role that *both* the instructor and the student play in the learning process: "The instructor's role in supporting learning is to lead the class in her growth mindset, her clear explanations of topics, her feedback towards her students, and her positive attitudes towards helping diverse students succeed. The student's role in the learning

process is to come to the instructor if they need help and to make sure that they keep up with lecture material." Isabella's response emphasizes the notion of partnership and joint work in the language she selects. For example, she talks about "working together with the students to make sure their needs are met."

Like Derek, Isabella also expresses that it is important for students to reach out to the instructor if they need help. However, Isabella's response more explicitly reflects the responsibility for learning by taking an extra step. For example, she collects formative assessment data and administers surveys to proactively identify and address students' learning and equity barriers like stereotype threat. Through this process, she uses multiple forms of data to identify and acknowledge that there is "something else" like equity barriers that plays a role in students' learning. She takes proactive steps to notice and address such barriers.

Furthermore, recognizing her role in addressing other factors influencing learning, she considered equity barriers by stating her stance on equity and being proactive in response to question 1, "I want to be able to know my students on a personal level so that I can be aware of any inequities they face, and be able to support them academically through them all. I am an ally for all groups of students."

In their reflection responses to question 1, Isabella and Derek both place responsibility for learning on the student. However, compared to Derek, Isabella also considers other sources of learning responsibility ("It's something else" subcode) like equity barriers that Derek did not mention explicitly in his response to the same question. She describes leveraging multiple data sources to understand equity barriers and find what the "something else" impacting students' learning might be. Isabella's written response places a higher degree of responsibility on herself than Derek's did. The key difference between Isabella and Derek is that Isabella represents the subcodes "It's on me" and "Uses multiple data sources" in her responses, as she finds barriers and adapts her teaching to the student needs' and context.

Consistent with how she describes sharing the responsibility for learning, Isabella described using multiple data sources to identify the key barriers to learning and what she could do next to further support learning. In response to question 1, she expressed that she used qualitative and quantitative data to identify equity barriers by "[sending] out a survey in the beginning of the school year to ask them about what they wanted to get out of the class, their demographics, and a little bit about their background." To understand her students' learning and backgrounds, she credited formative assessments as one of the impactful strategies to advance her students' success. She recalled that "Consistent with [her] philosophy on adaptive equity-oriented instruction, this approach probably worked because [she] used formative assessment and student interest surveys to diagnose students' existing competencies and interests." Reflecting on how she continuously adapts to address learning barriers in question 1, she wrote that her "teaching philosophy connects to the AEP framework because [she] get[s] to adaptively adjust [her] teaching styles each class to better suit [her] students' needs."

As Isabella collects multiple data sources, she shares the responsibility for learning by addressing other factors (i.e., "It's something else") that play a role in the learning process, such as stereotype threat and collaborative activities. For example, in response to question 2 (i.e., the teaching practice statement), Isabella identifies that her students have experienced stereotype threat. Noticing this equity barrier, she then writes, "I am adjusting my teaching practices by relating to their struggles. Most of my students are females and minorities, and as a female myself I am able to identify with the stereotype threat that they face. I encourage all my students to do the best they can and to embrace and build their analytic backgrounds."

She describes how she played a proactive role in adapting her teaching to mitigate equity barriers, which has reaped positive benefits with student learning gains based on pre- and post-assessments. In response to question 3 (i.e., discussing most effective practices for supporting student success), Isabella cites improvements in learning by "collecting the scores of the pre and post tests." Acknowledging the role of collaborative learning, she states that "My post-test was better than my pre-test and I noticed that students were more engaged in the collaborative activities." She recounted that "The average gain score was about 6 points, since 9-3 = 6. The range was also smaller on the post-test. Therefore, I think the lesson worked because there were improvements in student learning." Here, Isabella discussed how the post-test was better and included collaborative activities, which contrasts with how Derek prioritized students' responsibility to independently find the answer. In sum, Isabella's statements further support how she shares the responsibility for learning and uses quantitative and qualitative data (e.g., formative assessment and observations of student engagement in collaborative activities) to assess and support student learning proactively.

7.3 Discussion

In light of my conjectures, I now discuss the key lenses that drove differences between low- and high-growth instructors. I also highlight how the Responsibility for Learning Lens and Data Lens have emerged as critical lenses for advancing more equitable teaching practices. Then, I challenge existing notions of expertise to be more expansive, social, and collaborative. I end my discussion with implications for faculty development, limitations of the study, and opportunities for future research to build on the Responsibility for Learning Lens.

7.3.1 Review of Findings in Relation to Initial Conjectures

Equity Definitions and Perspectives Lens. High-growth treatment instructors consistently referenced their definitions and perspectives for equity in early questions (i.e., questions 1-2) and the later questions 3-18 to see if these definitions acted as a consistent lens for reflecting on and making pedagogical decisions. As such, these high-growth instructors consistently discussed their rationale and approach for their pedagogical adaptations, which were grounded in their different definitions and perspectives of equity, which appeared to be more entrenched when they reflected and made pedagogical decisions. Therefore, I argue that the Equity Definitions and Perspectives Lens appear to be more salient in motivating the perspectives of the high-growth instructors as they engage with students, adopt equitable pedagogies, and advance success for all learners. Moreover, high-growth instructors defined equity as adapting to students' needs more often than low-growth instructors. A few low-growth instructors defined equity as giving everyone the same thing, but they did not invoke that definition consistently when describing pedagogical practices.

Learning Goal Lens. My findings support the conjecture for the Learning Goal Lens in that high-growth instructors tend to report "Equity goal" and "Academic goal" subcodes more often than low-growth instructors, and often together in the same response. Like the Equity Definitions and Perspective Lens, high-growth instructors' "Academic goal" and "Equity goal" subcodes for student learning are more salient and frequent when they describe their pedagogical decisions in response to reflection questions. By considering both an "Academic goal" and "Equity goal," the instructors are noticing and addressing that psychosocial factors and equity barriers impact students' learning of academic content. Results from examining instructors'

"Academic goal" and "Equity goal" subcodes demonstrate that equity is central to the goals that the high-growth instructors report when discussing their beliefs and practices. Furthermore, I discovered that instructors in the high-growth group tend to follow certain patterns in their reflection responses. For example, the instructors:

1. *Expressed* an equity goal or belief, *noticed* the equity barriers impacting their students, and *adapted* their teaching to address the barrier; or
2. *Noticed* equity barriers impacting their students, *expressed* an equity goal or belief, and *adapted* their teaching accordingly.

Through these approaches, high-growth instructors addressed the exclusionary experiences that many STEM students faced.

Emotional Lens. I would argue that my findings cannot fully support or reject the Emotional Lens conjectures. Based on the reflection question responses that I coded, I found evidence that one Emotional Lens subcode, "Negative emotion", revealed differences between low-growth and high-growth treatment instructors. These findings suggest that compared to low-growth instructors, high-growth instructors tend to be motivated by negative emotions to adapt their teaching and promote a more inclusive environment. It is possible that the high-growth instructors may be demonstrating more empathy by noticing and relating to the negative emotions that students experience. However, more research needs to be conducted to see if this is the case.

Data Lens. Based on the findings, I found evidence that the Data Lens reflected major differences between low- and high-growth instructors. The subcode, "Uses multiple data sources to inform instruction and advance student learning," represented the largest differences between the LG and HG groups, where HG group mentioned this subcode more. Interestingly, both the low- and high-growth groups tended to have a stronger value for qualitative data. I noticed that many low-growth instructors tended to rely more often on their observations and interactions with students when reflecting on their pedagogy. By contrast, the high-growth instructors tended to value and use both qualitative and quantitative data more often when reflecting on and making instructional decisions. Taking a UDL perspective (Rose et al., 2006), it is important to collect multiple data sources since some students may not actively share their barriers in a conversation. Having multiple ways (e.g., surveys, conversation) for students to share their barriers is key to understanding their needs and creating a more humanized form of education that is responsive to their lived experiences.

Responsibility for Learning Lens. The findings suggest that low-growth instructors placed the responsibility for learning more often on their students compared to themselves or contextual factors. On the other hand, high-growth instructors identified and adapted teaching to address learning barriers because they shared the responsibility for learning with themselves, the student, and other factors (such as imposter syndrome, stereotype threat, lack of prior experience, peers). Consequently, the high-growth instructors better noticed their students' equity barriers, and in turn reflected on their practice to make pedagogical adjustments to address students' needs.

7.3.2 Importance of Sharing Responsibility for Learning and Using Multiple Data Sources to Advance Student Learning

I identified an emergent pattern of the co-occurrence between the "Sharing responsibility for learning" and "using multiple data sources" codes. I was interested to understand the individual, departmental, and societal characteristics driving instructors to simultaneously share

responsibility for learning and use multiple data sources. In addition, I hoped to discover associated beliefs and practices that might support equity in STEM teaching. I found that sharing the responsibility for learning and using multiple data sources to inform teaching was more prevalent in the HG group than in the LG group. This pattern suggests that instructors who share the responsibility for learning and use multiple data sources may be more competent at promoting more inclusive and welcoming learning environments, as measured by AEPC.

By contrast, instructors who do not share the responsibility for learning or do not use multiple data sources were observed to be less competent at promoting inclusive environments and less effective at improving student success (Phuong et al., 2022). Placing the responsibility for learning primarily on students can promote a "teach yourself" model. This model can ignore the social nature of learning. For instance, Derek's primary focus on the individualistic aspects of Computer Science and related industries echoes part of the Protestant ethic of "pulling yourself by the bootstraps" (i.e., individual hard work is what leads to success; Bonilla-Silva, 1999; Kinder & Sears, 1981; Markus & Conner, 2013; Sears & Henry, 2003). Goode et al. (2021) have argued that an individualistic nature can be pervasive in Computer Science education. Such an approach can be a one-size-fits-all model that ignores multiple data sources on the individual and collective needs of students in the classroom.

To advance a critical perspective, I contend that a "teach yourself" pedagogy privileges certain populations and does not fully consider the unjust and unequal barriers, obstacles, and sociohistorical contexts of students, including those from minoritized backgrounds. A "teach yourself" model can disenfranchise those with already marginalized or undervalued epistemic, social, and material capital. Adopting such a pedagogy that places the onus of learning primarily on students can be used against them for the social and cultural reproduction of power relations. For example, a "teach yourself" pedagogy privileges the instructor and absolves the instructor from responsibility if a student does not learn a concept and/or does not succeed. Reproducing such a norm reifies power structures where instructors do not need to focus on assessing, improving, or adapting their practice and mindset. This insistence of a lack of shared responsibility for learning can limit instructors' perspective taking and the development of social-emotional competencies, such as reflecting on and addressing their areas for growth. This norm can further reinforce the misguided logic of weed-out courses and the survival of the fittest mentality, perpetuating the notion that students are deficient and not "cut out" for STEM if they cannot rise up to the instructors' expectation with minimal guidance and scaffolding. Applying this logic and mentality to pedagogical practice in higher education has historically excluded and filtered out marginalized students from STEM (Steen, 1987; Weston et al., 2019).

7.3.3 Using Data for Targeted Support in Active Learning and Productive Struggle

The examples of Derek and Isabella are reflective of more general patterns of responses I found throughout the low- and high-growth group, respectively. They illustrate differences in how instructors use data to adapt pedagogical practices like active learning to support students' learning. Furthermore, Derek's approach of lecturing concepts at a basic level and asking students to engage in advanced problem-solving has been a typical teaching approach in higher education STEM courses (Deslauriers et al., 2019). When taught in these lecture-based modes, students in lecture can have an illusion of knowledge thinking they understand the concepts and how to solve problems (Deslauriers et al., 2019). But, research has shown that many students, including URM students, learn less and could not apply more advanced problem-solving

techniques when experiencing such predominantly lecture-based classrooms. As described in chapter 2, Theobald et al. (2020) found that lecture-based classrooms can exacerbate achievement gaps between URM and non-URM students. The lack of alignment between instruction and assessment of more advanced problem-solving has contributed to persistence and retention challenges in STEM (Theobald et al., 2020; Seymour & Hewitt, 1997).

By contrast, active learning (e.g., deliberate practice, productive struggle) can provide opportunities for both students and instructors to play a more proactive role in advancing learning. Pedagogically, productive struggle can be effective for learning (Schwartz & Bransford, 1998). Productive struggle is a part of the AEP framework (see AEPC Learning Progression, Level 3). However, solely or over relying on productive struggle without sufficient targeted support, guidance, and possible answers to problems eventually does not address a wide variety of student needs and support learning (Clark et al., 2012; Phuong et al., 2017). Consistent with the AEP framework, instructors would use multiple data sources to adjust evidence-based teaching practices (e.g., productive struggle, modeling key skills, deliberate practice, feedback) to address students' equity barriers and support success.

While both HG and LG instructors might describe active learning practices, my codes highlight how there can be a spectrum where HG instructors like Isabella use adaptive, data-driven practices more often compared to LG instructors like Derek. Instructors like Derek in the LG group can benefit from providing more active learning (e.g., productive struggle) and formative assessments to proactively understand and address student learning barriers. Further, during this process, instructors can support students by drawing on multiple data sources and considering external factors such as sociocultural contexts of students. For example, during class, instructors can create opportunities for productive struggle in a variety of formative assessment activities where students can grapple with and practice problems individually or in small groups. These activities can enable students to assess whether they understand concepts and then receive support and guidance from instructors (e.g., hints, modeling, mini-lecture(s), eventually providing possible answers). This process can promote students' metacognitive development, where they can reflect on their learning progress after receiving guidance and seeing possible correct answers to solving problems. Moreover, like Isabella in the HG group, instructors can share the responsibility for learning and collect multiple data sources to understand and address students' experiences, sociocultural contexts, and learning barriers (e.g., misconceptions).

7.3.4 Practical Strategies and Recommendations for Adapting Pedagogical Practices

It might appear that adapting to a classroom of students might be difficult. In Appendix A, I provide strategies that these instructors and other instructors have used to adapt to a whole classroom of students. Some of these strategies include using automated formative assessment data, worksheets with problems of varying levels of difficulty, and online surveys that can automatically summarize information. Online learning has incredible potential to customize formative and summative assessment problems. Examples include PraireLearn and intelligent tutors.

Formative assessments, such as ungraded quizzes, surveys, and feedback forms, are key strategies and low-hanging fruit for cost-effective ways to gain context on how to gain insight into students' equity barriers and adapt teaching practices. The assessment data support instructors in what concepts to focus on based on where students are revealing misconceptions or challenges. After collecting this information, instructors should "look at the most frequent kinds of responses that are actionable, such as patterns in student performance, student feedback, and

classroom observations. Then, instructors can consider adjusting teaching methods when a theme is noted by more than 10 percent of the class." (Phuong et al., 2021, p. 1). For issues identified "by less than 10 percent of students, instructors can think about what kind of individualized support and targeted resources could be provided, such as video recordings of class, specific readings and supplemental resources that do not assume prior knowledge." (Phuong et al., 2021, p. 1).

Furthermore, instructors' adaptation can go beyond teaching practices. For example, adaptation can also include grading for equity approaches and flexible course policies (e.g., allowing extensions, letting students redo assessments or other assessments measuring the same learning outcomes to gain points). This holistic adaptation process enables instructors to take the responsibility to mitigate equity and contextual barriers that students face by understanding the external factors impacting students' learning and allowing students to progress in an environment that is responsive to their circumstances.

7.3.5 Expanding Notions of Expertise

The LG instructors' tendency to place the responsibility for learning more heavily on the student echoes literature, suggesting that equating expertise with independence can be harmful for students from minoritized and diverse backgrounds (Phuong et al., 2022). When an instructor places the responsibility for learning primarily on the student and expects students to come up with the solutions by themselves without sufficient support (e.g., the "teach yourself" model), then a problem can arise with a lack of alignment between instruction and assessment. This "teach yourself" model can minimize the important role the instructor plays in guiding the learning process and the feedback loops that can advance student learning. Through guiding the learning process and assessing student feedback, instructors can acknowledge and address the various levels of incoming preparation, equity barriers, and systemic oppression that students face within and beyond the classroom.

Recognizing the legacy of racialized education systems in STEM (Goode et al., 2021), I argue that equating expertise with independence can become racialized in a learning context, may promote resistance to feedback, and can create barriers to learning, perspective taking, and growth. The Protestant ethic of individualism can be harmful to social justice and equity efforts (Hudson & Coukos, 2005), since it reproduces symbolic racism (Bonilla-Silva, 1999; Kinder & Sears, 1981; Sears & Henry, 2003). According to Kinder and Sears (1981), symbolic racism is "a blend of antiblack affect and the kind of traditional American moral values embodied in the Protestant ethic" (p. 416).

As mentioned above, it is a common belief that coming up with solutions independently is an essential skill in industry (i.e., corporate settings). However, this narrative does not match the full reality of many industry settings, which have also been shown to be highly collaborative and interdependent (Felder, 2007). Therefore, it is important for learners to develop skills to work both independently and collaboratively to address complex problems.

In the extreme, adopting a "teach yourself" model is a racialized approach to teaching because it privileges and elevates Eurocentric cultural values (i.e., "pull yourself up by the bootstraps") as the standard for success. This model can reproduce a "hidden curriculum" that assumes a white middle- and upper-class cultural capital for success, which can ignore and disadvantage interdependent cultural values and the diverse ways students learn (Bourdieu, 1973; Goode et al., 2021; Markus & Conner, 2013; Phuong et al., 2017; Phuong et al., 2022). In addition, the "teach yourself" model privileges students who already know the content and have

the financial means and time to hire tutors. The latter can be more difficult for low-income students with limited access to financial resources and who need to work part-time job(s), which limits their time to spend on schoolwork. The myth of self-reliance, bootstrapping, and individual merit do not fully consider the institutional, external, and social resources often "unavailable to urban youth of color, but overwhelmingly accessible to more privileged young adults (Anderson, Turner, Heath, & Payne, 2016)" (Lardier Jr. et al., 2019, p. 492). In this way, the "teach yourself" model and equating expertise with independence promotes and exacerbates opportunity and achievement gaps. Such a belief can reinforce the "lone genius" myth and the idea that people can be self-made (Montouri & Purser, 1995), which could further perpetuate a myth of meritocracy.

By contrast, in an AEPC model, an expert is someone who continues learning, seeks feedback and support (e.g., coaching), pushes themselves, collects multiple data sources to reflect on their own practice and understand those they work with, identifies conscious and unconscious biases, and tries to grow both independently and collaboratively. These were the characteristics that I often found present in teaching reflections among HG instructors. Furthermore, seeking support and feedback is not seen as a lack of confidence or as a deficiency, but rather as a desire to grow and innovate.

As reflected especially in the high-growth instructors, the approach of seeking feedback and checking-in with learners can promote key social-emotional instructor competencies, such as perspective taking, understanding learners' experiences and emotions, and reflecting on areas for improvement (e.g., metacognition) (Nguyen et al., 2022). These skills and competencies are important for practitioners who wish to design inclusive learning environments for everyone (Nguyen et al., 2021).

7.4 Implications for Educational Development

My research also suggests that it can be beneficial to train instructors in sharing the responsibility for learning and using multiple data sources to adapt teaching and advance student learning. These approaches can enable instructors to notice and address students' learning needs and equity barriers. I have found that these competencies can be learned and developed by instructors through educational development programs, such as the ones focused on AEP being researched in this study. Critically interrogating the "teach yourself" model, the myth of meritocracy, and certain beliefs about self-reliance is key to advancing more equitable and anti-racist pedagogies.

7.5 Limitations and Future Research

My sample size of 20 instructors is limited in generalizability, and future research can explore additional reflection responses and a larger number of instructors. It would be worth examining the degree to which these findings replicate in other educational institutions, including more diverse higher education contexts such as community colleges, state colleges, and private colleges. Moreover, I would like to examine additional items to see if my findings for the following lenses still hold: Authority and Community Lens, Equity Definitions and Perspectives Lens, the Learning Goal Lens, Data Lens, and Responsibility for Learning Lens. In conducting subsequent research, I would be interested in designing additional items that gauge some of these lenses more directly. I hypothesize that in more diverse higher education contexts, lenses – especially Equity Definitions and Perspectives, Data, and Responsibility for Learning – would not only be frequently

used by HG instructors, but also necessary to address equity barriers that students face academically, emotionally, and socially. A similar study in different contexts, then, would be invaluable for yielding new insights into how instructors vary in supporting the success of an increasingly diverse student population, and whether they show co-occurrences between codes.

In particular, I am interested in new research directions after seeing a possible relationship between sharing the responsibility for learning and using multiple data sources to advance student learning. With a larger sample, I would like to investigate further how the treatment can encourage instructors to share the responsibility for learning and use multiple data sources to promote a more sociocultural approach to learning rather than an individualistic one. Such research would be intriguing to me because many higher education institutions reinforce individualistic approaches to learning.

Another avenue might be to refine measures for AEPC and sharing the responsibility for learning. In particular, I hope to develop a more robust measure for assessing the degree to which educators share the responsibility for learning. Based on data from these measures, one can then design targeted professional development to promote greater AEPC and the value of sharing the responsibility for learning. The goal of these programs would center on training to support instructors' growth on the spectrum of sharing the responsibility for learning and AEPC. For example, the training could provide reflexive activities for instructors to examine privilege and the strengths and access to learning opportunities that students from low-resourced communities possess. It is also critical for instructors to understand that what worked for them would not work for all. This is why it is essential to use multiple data sources to understand what supports and does not support student learning.

In addition, I would like to further examine whether a lack of shared responsibility may influence resistance to assessment and using multiple data sources to inform practice. Moreover, I hope to investigate whether those who do not share the responsibility for learning may see assessment as more work or taking too much time, and therefore do not want to proactively use and respond to assessment data. I anticipate that not sharing the responsibility for learning may limit the social-emotional competencies of perspective taking and reflecting on areas for improvement.

By contrast, I would like to examine if those who share responsibility might acknowledge it takes more time and effort to understand students, but can see that there is a greater return for student learning. I would expect that sharing the responsibility for learning can promote the social-emotional competencies mentioned above and a commitment to continuous improvement and growth.

I would also be interested in examining how instructors can share the responsibility for learning when supporting student success during inquiry-based and open-ended activities (e.g., discovery learning). This line of research excites me because the process of productive struggle and discovery learning can help learners exercise agency in discovering novel ideas and solutions on their own terms, which can be empowering. However, I argue that discovery learning can be misused if instructors expect students– without sufficient scaffolding– to discover the instructors' answer or way of thinking.

Furthermore, my current and future research plans can support the field of educational development in better understanding different mechanisms (e.g., sharing the responsibility for learning) that impact instructor resistance to using formative assessment and multiple data sources to inform pedagogies that advance student learning.

CHAPTER 8: Conclusion

In this chapter, I first provide a summary of the major dissertation findings and contributions to the literature. I then discuss lessons learned, limitations, directions for future research, and the implications of my research.

8.1 Summary of Major Dissertation Findings and Contributions to the Literature

We have long known that teachers vary in their capacity to invite all students into the learning process. My findings take the next step: Identifying specific mechanisms that drive college STEM instructors' adoption of equitable teaching practices, such as Adaptive Equity-Oriented Pedagogy (AEP). Below, I focus on four specific contributions made by this study: demonstrating the value of RCT in the field of educational development; validating an equitable teaching measure; providing a framework for adaptive instructor professional development; and illuminating mechanisms that explain how high-growth instructors foreground equity, share the responsibility for learning, and leverage multiple data sources.

As shown in chapter 2, the current higher education literature has identified, based on self-reported surveys, that college instructor professional development programs can increase STEM instructors' adoption of equitable beliefs and practices (Harrison-Bernard et al., 2020; Metevier et al., 2010; O'Leary et al., 2020; Parker et al., 2016; Rooney et al., 2020). However, since these studies contained no control group or randomization, the identified improvements could not be tied to a specific intervention and could have been due to maturation, a practice effect, or other confounding factors. Based on my systematic literature review on college instructor professional development programs focused on inclusive STEM teaching, I found that this area of scholarship could benefit by controlling for instructor-level variables when assessing the impact of the elements of these programs on instructor's equitable teaching competencies. Below, I discuss how my mixed-methods research design builds on the existing literature and addresses some of the limitations. As described in chapter 1, I explored the following research questions for my dissertation:

1. **Quantitative research question 1:** Is there a greater increase in mean AEP competency scores across time in the treatment pedagogy course than in the control pedagogy course, controlling for gender, URM status, years of teaching experience, years of tutoring experience, and number of students?
2. **Qualitative research question 2:** What mechanisms for instructional change characterize differences between low- and high-AEP growth instructors in the treatment?

Randomized controlled trials. In my study, 129 instructors[3] were randomly assigned to treatment and control pedagogy courses. While the control course taught student-instructors about AEP, the treatment course modeled AEP explicitly by using weekly student-instructor reflection data to continuously adjust course discussion and activities. I designed the control condition to be effective and to have the same number of class sessions as the treatment. I see my

[3] As mentioned in chapter 1, I use the term "instructor" for graduate and undergraduate students who serve in an instructional capacity under faculty direction and are enrolled in the pedagogy courses.

work contributing to conversations on how studies of college instructor professional development programs focused on equity can leverage randomized controlled trials and multilevel modeling to reduce bias, mitigate the likelihood of confounding variables, create comparable groups, and support causal inference. To control for instructor-level variables and account for dependency in repeated AEPC measures, I leverage hierarchical and longitudinal modeling to examine instructors' learning progression, adoption, and development of equity-oriented teaching practices over time. I also applied hierarchical and longitudinal modeling to analyze growth in instructors' AEPC while accounting for instructor-level characteristics.

Validated equitable teaching measure. In collaboration with faculty and senior administrators, I sought to understand how to develop instructors' equitable teaching competencies and student success to create more welcoming environments in STEM. These competencies were derived from the AEP framework. My dissertation described and validated an Adaptive Equity-Oriented Pedagogical Competency (AEPC) assessment, a measure of college instructors' effectiveness with inclusive teaching, that was used to track instructor growth in equitable teaching competencies in my study. Based on partial credit Rasch modeling (Masters, 2010), the AEPC assessment has high reliability (0.94), has strong validity based on Wilson's (2005) strands of validity framework, shows no gender assessment bias, has high inter-rater reliability, and correlates significantly with researchers' review of teaching and pedagogical materials (Phuong et al., 2022).

My research demonstrates that instructors can use the AEPC assessment to demonstrate competency and progress in applying equitable teaching practices. This assessment approach enables instructors to document evidence of their own inclusive teaching practices, which provides insight on how they might plan, teach, reflect, and alter their pedagogical practice to address students' equity barriers. In fact, the AEPC assessment includes materials that are often found in portfolios used in part to evaluate college teaching, such as teaching philosophy and practice statements, learning outcomes, formative assessment, and reflections on teaching (Seldin et al., 2010). In addition, the AEPC assessment does not exhibit the known gender and racial bias found in course evaluations (Boring et al., 2016; MacNell et al., 2015; Phuong et al., 2022; Stark & Freishtat, 2014). The AEPC assessment provides instructors with clear, measurable criteria that have been shown to significantly improve student success. These clear, measurable, and documentable criteria are needed if we want to increase instructors' adoption of equitable teaching practices at scale.

In my research, I found that it was critical to examine assessment bias because many assessments have historically assumed a white middle-and upper-class cultural capital, exhibited bias against minoritized subgroups, and filtered them out of educational systems (Bourdieu, 1973; Chapman, 1988; Kendi, 2019). Therefore, it is important for me to ensure that assessing the adoption of equitable teaching competencies does not exhibit bias against any instructor subgroup.

Adaptive instructor professional development. Using the AEPC measure, I leveraged multilevel modeling to examine instructors' learning longitudinally to understand their development and adoption of equity-oriented practices. I specifically examine growth in equitable teaching competencies over time between the treatment and control conditions. I sought to understand if the process of modeling AEP would increase instructor learning over time in the treatment relative to the control, as measured by AEPC assessments. This hypothesis was consistent with research demonstrating that instructor reflection, which is a core component of the AEP treatment, can improve learning (Phuong et al., 2017; Phuong et al., 2022). Using

multilevel-regression modeling, I found that the treatment, on average, significantly improved instructors' AEP competencies from pre- to post-semester compared to the control (2.63 standard deviations, $p<0.001$). Compared to other studies (Harrison-Bernard et al., 2020; Metevier et al., 2010; O'Leary et al., 2020; Parker et al., 2016; Rooney et al., 2020), recognizing the methodological differences, this is a strong effect. Compared to the control, more student-instructors in the treatment used multiple data sources to adjust teaching and address equity barriers to advance student learning. In the treatment, I also qualitatively analyzed written reflections of low- and high-growth instructors on AEP competency to identify the beliefs and practices that characterized each group. Compared to low-growth instructors, I found that the high-growth instructors more often shared the responsibility for learning with students and also used multiple data sources to support students' success.

High-Growth instructors foreground equity and Share the Responsibility for Learning. While the quantitative analysis shows the greater adoption of AEP practices in the treatment group, I leveraged qualitative methods to better understand the motivations, beliefs, and values that motivated instructors to adopt AEP. In order to provide insight into instructors' learning processes, I complement the statistical analyses with qualitative methods that include exploring teaching reflections, teaching materials, and evidence of equitable teaching. This approach enables me to understand instructors' lenses (i.e., their values, motivations, perspectives) for adopting AEP, as well as to what extent instructors plan, teach, assess, and reflect on data to improve student success in STEM courses in disciplines such as computer science, data science, and statistics. Identifying and examining these lenses enables me to characterize the mechanisms that distinguish low- and high-growth instructors on AEP competency. To examine the low-and high-growth instructors, I identified a sample of the top and bottom 25% of treatment instructors, based on their growth in equitable teaching competencies. To include non-dominant voices, I first sampled minoritized instructors from each group. Then, I randomly selected other instructors to create sets of 10 for each group. Next, I coded instructors' teaching reflections using a deductive codebook based on the six lenses described in chapter 7. I compared differences in the presence of subcodes to provide insight into whether and how the seven lenses characterize low- and high-growth instructors in the treatment group. Through this mixed-methods research, I build on the literature discussed in chapter 2 by bringing together learning sciences and organizational behavior theories to analyze the mechanisms that a) motivate instructors to learn equitable teaching practices and b) support greater equity in student learning outcomes.

I found that noticing equity barriers in higher education was key for high-growth instructors. In my dissertation, I focused more specifically on lenses for pedagogical change, which relied on the presence codes in instructors' teaching reflections focused on the following: instructors' definitions and goals around equity, the emotions that drive their behaviors, the kinds of content they prioritize, where they place the responsibility for learning, use of multiple data sources to inform teaching, and how departmental and disciplinary cultures impact their pedagogical choices. In this way, these lenses, or mechanisms, allowed me to address content-specific and non-content specific noticing skills (e.g., stereotype threat, imposter syndrome). In the teacher noticing literature, van Es and Hand (2017) would argue that content-specific (e.g., academic concepts) and non-content specific noticing focused on equity (e.g., stereotype threat) is important for promoting a more inclusive classroom. Building on this work, I examined whether high-growth instructors have more equity-driven and academic content-driven goals than low-growth instructors. I also investigated how instructors reported adapting teaching based

on more than unaided classroom observations, but also on information such as written reflections and data on student equity barriers. Building on the teacher noticing literature from learning sciences in these ways, I contributed an adaptive framework for developing instructors and helping them notice and address students' equity barriers.

8.1.1 Considerations for Designing Responsive Professional Development

Considering Vygotsky's notion of ZPD, instructors may be viewed as operating along a learning progression that was supported through scaffolding (e.g., responsive activities). The treatment provided a culturally responsive and adaptive learning experience that strengthened the scaffolding and thus accelerated the learning progression relative to the control. The treatment facilitators reviewed instructors' reflections to understand their values, motivations, and current teaching and learning models. In the treatment, facilitators addressed instructors' cultural values (e.g., commitment to data and quantitative methodologies grounded in disciplinary norms) by communicating and presenting pedagogy research, strategies, and stories that aligned with instructors' cultural values. For example, instructors reported that they valued experimental, quantitative, and replicated research that illustrated the efficacy of specific teaching practices they were asked to adopt. Responding to instructors' perspectives, treatment facilitators highlighted additional randomized, quasi-experimental, or quantitative studies that validated the impact of inclusive and evidence-based pedagogies. In particular, addressing skepticism and resistance to education and qualitative research was addressed with a follow-up discussion of additional replication studies. In the treatment, facilitators also showed the value of qualitative research and emphasized asking instructors to collect their own qualitative and quantitative data to see what kinds of teaching practices worked. The instructors learned from each other and the facilitators to see and experience how using multiple forms of data was helpful for understanding student learning barriers and adapting teaching.

Prior to intervention, the facilitators identified that many instructors in the treatment defined equity as equality and placed the responsibility for learning only on the student. Thus, the treatment had conversations on interrogating long-standing beliefs and inequitable ideologies that the "good students are the ones who teach themselves." This approach was key to expanding critical consciousness since it asked instructors to rethink common cultural norms, values, meritocracy, and larger systems (e.g., structural racism) that reproduce inequities, where students from underprivileged backgrounds may not have equitable access to resources and opportunities that enable them to teach themselves. Understanding that the learning responsibility is shared and that contextual factors influence learning outcomes reflects a higher degree of critical consciousness on equity issues. Addressing equity issues is critical for adapting to student learning needs. Proactively sharing the learning responsibility shifts the pedagogical focus from an individualistic "pull yourself up by the bootstraps" approach to one that is more social, developmental, and culturally responsive. Future pedagogy course iterations should more comprehensively address instructors' lenses to promote equitable beliefs and practices.

Adapting practices to students is critical, since the specific strategies described above may not work in other contexts. For example, as an educational developer, I have met faculty in non-STEM disciplines who do not value randomized controlled trials or studies with regression analyses. These non-STEM faculty cared more about qualitative evidence of the impact on student learning and sense of belonging. In fact, I found in my qualitative research that low- and high-growth instructors valued qualitative data more than quantitative data by the end of the semester.

Being critical of my research, I have also noticed from my own personal experience that STEM faculty and graduate student instructors are not all swayed to adopt equitable pedagogies just because they saw a quantitative study. Smith (2020) found that just showing quantitative studies to faculty was insufficient for influencing adoption; rather, instructors' beliefs and values need to be addressed in educational development programs.

I echo Smith's (2020) argument. I found that applying principles like AEP is critical because we need to understand and address instructors' specific perspectives, values, beliefs, and needs with respect to learning equitable teaching competencies. Within and beyond this study, I met STEM instructors who look for different kinds of evidence when deciding to consider adopting an equitable pedagogy, such as the sample size and scale of the study (e.g., the number of students being taught), the discipline of the content being taught, whether the quantitative regression models accounted for important variables (e.g., baseline achievement, URM status), the use of a randomized controlled trial, and if the teaching practices being researched aligned with their values and suited their pedagogical needs. Identifying what instructors look for and giving them an opportunity to adapt and iterate an evidence-based teaching practice in their context based on student data was a key component of the experiential learning process. Providing assignments that asked instructors to plan teaching, apply an evidence-based practice, collect data, see what worked and what did not, and make pedagogical adjustments and decisions was a key part of the learning process that was supported by a responsive approach like AEP. I would argue that the ways we adapt a pedagogy course and the content of the pedagogy course will need to evolve based on factors, such as the changing nature of higher education ecosystems, trends in the equity and anti-racist literature, and the emerging research on teaching and learning. Therefore, simply spending more time to refine a course without intentionally applying competencies of AEP may not be addressing the underlying beliefs that are needed to create more inclusive and welcoming environments in STEM. Smith (2020) has argued that educational developers need to address college instructors' reasoning, values, and beliefs to support the adoption of evidence-based practices. The AEP framework offers an empirically tested approach for achieving this goal.

I acknowledge that the five studies described in chapter 2 do show gains in equitable teaching competencies. However, these studies measured gains using self-report surveys from instructors about their beliefs and if they would adopt equitable pedagogies. My research required instructors to provide examples and evidence of how instructors applied equitable teaching competencies.

8.2 Lesson Learned

I have learned how my work is situated in conversations about equity in STEM education. In my experience, some have argued that equity and diversity waters down the curriculum and standards. But, my work pushes against this. In my prior research, I found that prioritizing equity and diversity in pedagogy can be much more effective than traditional active learning, in helping students meet rigorous learning outcomes (Phuong et al., 2017; Phuong & Nguyen, 2019). In my dissertation, I took this a step further by using learning sciences theories to build and validate measurable, equitable competencies that foster more inclusive classrooms. These learning theories included Vygotsky's (1978) Zone of Proximal Development (ZPD) and the teacher noticing literature. In the teacher noticing literature, expert teachers learn to notice and be more sensitive to differences in student learning (van Es & Hand, 2017). I drew on the

teacher noticing literature to build AEP competencies that focused on leveraging content-specific (e.g., academic concepts) and non-content specific noticing focused on equity (e.g., stereotype threat). Prioritizing content and non-content specific noticing is important in AEP, since it is important for instructors to identify and address equity barriers as they move a student farther along their ZPD.

Throughout this work, I have witnessed the power of leveraging systematic research strategies to reveal surprising findings. For example, I thought that the Authority and Community Lens would explain differences between low- and high-growth instructors. Overall, my conjecture was not supported. In addition, I thought that low-growth instructors would value quantitative data more than qualitative data. In reality, the opposite ended up being true. These reflections reinforce why I need to continue conducting systematic research. In this study, another researcher randomly selected low- and high-growth instructors to create sets of 10 for each group after the minoritized instructors were selected. This strategy limits the selection of low- and high-growth instructors who best fit a hypothesis. In addition, focusing on instructor cases who are representative of trends in subcodes or the co-occurrence of codes is more systematic for illuminating key differences.

I have learned that my work is not only shifting how instructors think about equity, but also how I interrogate my biases and think about equity. I realized that my perspectives have been biased by my prior research, experiences at middle- to upper-class elite institutions, and the literature that challenges the myth of meritocracy and individualistic models of learning. Grounded in my experiences with teaching culture, especially in STEM, at elite institutions, I designed the Responsibility for Learning Lens to address the "teach yourself" mentality pervasive in STEM learning. Nevertheless, my lived experiences and positionality inform my belief that individualistic models of learning exacerbate the toxicity and competitive STEM cultures in higher education classrooms. As mentioned in chapter 7, an individualistic model of learning can elevate a white middle- and upper-class cultural capital as the standard for behavior and success, which can racialize teaching, learning, and assessment. An individualistic model of learning helps justify and reify a lack of alignment between instruction and assessment, learning alone without the help needed, and curving down (an approach that promotes competition over collaboration). These are all factors that have characterized weed-out courses and intensified attrition in STEM (Seymour & Hewitt, 1997; Weston et al., 2019). For these reasons, I challenge educators to think beyond and to challenge the pernicious belief that "the only good students are the ones who teach themselves and don't need help." I argue that it is important for instructors to share the responsibility for learning by providing students with low-stakes assessment opportunities to learn independently and collaboratively. At an existential level, I contend that learning is inherently social and distributed, because we learn from each other and from physical and online resources (among others) which are extensions of someone else's mind and social context (Engeström, 1987; Lave & Wenger, 1991; Wenger, 1998; Vygotsky, 1978).

Inspired by van de Sande and Greeno's (2012) perspectival frames (e.g., epistemological, positional, conceptual frames), I contribute 6 lenses that help me understand how instructors think about and frame their teaching and interactions with students. I build on the epistemological frame to delve into understanding instructors' rationales for using qualitative and quantitative data sources to inform teaching. My research extends scholarship on the How People Learn Framework, which has four lenses on learning (Laghari et al., 2017; National Research Council, 1999). These are the community-centered, learner-centered, knowledge-centered, and assessment-centered lenses. I build on the community-centered lens and assessment-centered

lens by contributing the Authority and Community Lens and Data Lens. Through my research, I add equity-focused lenses such as the Equity Perspectives and Definitions Lens, the Learning Goal Lens, and the Responsibility for Learning Lens. These lenses helped me better understand how instructors engage learners and knowledge in the classroom. These lenses can be useful for teacher education and educational development programs.

In my pedagogy, I do not think of knowledge not as an object, but as an activity or process (e.g., learning by doing, reflecting, interacting, and feeling in a sociocultural context) (Engeström, 1987; Gutiérrez et al., 2016; Vygotsky, 1978). This is why it was key for me to create experiential learning opportunities where instructors made sense of data and made pedagogical decisions based on their reflections. I saw that instructors learned through reflection, experience, and emotional processes on how to design, adapt, and iterate equity-oriented practices. From my data, it appeared that this process helped foster the development of equitable teaching competencies among high-growth instructors as they became more aware and responsive to different factors (e.g., equity barriers, peer collaboration) playing a role in student learning.

Among instructors, I also saw elements of distributed and social cognition (Hutchins, 2000), which can refer to how instructors' cognition and social practices were influenced by those of others (e.g., students, colleagues). For example, the high-growth instructors incorporated their students' perspectives and thinking, or cognition, into their pedagogical and social decisions. In this way, the high-growth instructors internalized and responded to the cognitive and social practices that their students brought to the classroom. Throughout this process, cognition was distributed and was not solely a property of the individual instructors' mind as they learned about equitable pedagogies (Hutchins, 2000; Salomon, 1993).

I also see connections between my research and work exploring learning in communities (Engeström & Conant, 2000; Lave & Wenger, 1991). I analyzed how learners in a community of practice can engage in productive learning practices where learners sought and found evidence frequently, made sense of data about students' learning, and leveraged evidence to make decisions. In my context, the productive learning practices could be characterized by the application and development of AEPC. Like Jaber and Hammer (2016), I also explored how epistemic affect and epistemic motivation can play a role in productive disciplinary practices and engagement. For example, I found that high-growth instructors described how their emotions about Computer Science motivated their pedagogical practices to make Computer Science education more inclusive and accessible for all students.

8.3 Limitations and Future Research

Serving as a facilitator is a limitation because I co-created AEP and have applied it in multiple contexts before. This means that the amount of time it took me to apply AEP on instructors in the pedagogy course might be less than the amount of time it would take for someone who is untrained or newly trained in AEP. Regarding adaptation, I spent about 1- 1 hour and 15 minutes per week in the first three weeks.[4] I then spent an hour, on average, or less

[4] Anecdotally, I would like to note that I had more control instructors going to office hours, so the amount of time spent between conditions became comparable. In the treatment, I addressed many of the questions that instructors posed in reflections during class, which I speculate reduced the amount of time I spent with treatment instructors outside of class significantly.

thereafter per week on adapting the pedagogy course. I spent this time reviewing instructors' weekly reflections and noting the main themes to discuss in the conversations at the beginning of class or to make reference to them during class discussions and lectures. These data helped me use my time more productively to frame conversations and pose questions that helped spur reflection that challenged issues like the "teach yourself model" and meritocracy. There can be risks of introducing bias when one both designs and evaluates an intervention. In my research design, I mitigate these risks by running a randomized controlled trial; blinding the grading of assessments by instructor name and condition; bringing in researchers and scorers who were not a part of the pedagogy course teaching team; conducting inter-rater reliability on assessment scores; triangulating data and perspectives with a research team; seeking feedback from other scholars, blindly coding instructors' reflections, and randomly sampling instructors for qualitative analyses after the minoritized instructors were selected. In addition, the effectiveness of the AEP interventions is fairly consistent with prior studies (Phuong et al., 2017; Phuong & Nguyen, 2019; Phuong et al., 2022). I also conducted research where the order of the treatment and control were switched to address any concerns on whether the treatment was more effective because instructors taught it after the control and/or had more time to spend on the treatment within a week. I thought this was important to explore because one could argue that teaching a class the second time will mean it is more effective. I found that the treatment effect was similar under either order (Phuong et al., 2017; Phuong & Nguyen, 2019).

Furthermore, I would suggest creating a train-the-trainer model if the pedagogy course were to be more widely scaled. It might take other facilitators replicating the treatment more or less time to implement it on a weekly basis. Future research should consider how to develop reflection response assignments that can be fed into a technology (e.g., an AI) that can present the most frequent themes. More importantly, I would argue that implementing responsive professional development approaches like AEP are worthwhile because research has shown that there is major return on investment financially and with student success outcomes when contributing resources, time, and energy to educational development programs (Brown & Kurzweil, 2014).

As a reminder, the years of teaching and tutoring experience were not significant predictors of AEPC growth in my dissertation. One limitation is the low variability in teaching experience from my study. Ebert-May et al. (2011) and Ebert-May et al. (2017) found that teaching experience was a significant predictor that was negatively correlated with the adoption of inquiry-based pedagogies among faculty. In a subsequent study, it would be interesting to examine if this relationship would hold if I examined faculty with a larger variation in teaching experience than the instructors in my study. It is possible that this relationship may not hold because inquiry-based learning is a different approach than AEP.

Another limitation is that my qualitative study was structured a priori with reflection questions. In future work, I would be interested in observing how my intervention leads to instructors' behavioral changes in real live classrooms over the course of a semester and beyond. If I interviewed instructors soon after they taught a lesson, I suspect that the Emotional Lens might provide richer qualitative insights on how emotions related to teaching can impact subsequent classroom practices.

With respect to examining organizational change, I would be excited to conduct ethnographic research to examine interactions among faculty, student-instructors, and students; how professors and peers influence assumptions about being a student-instructor; and how

pedagogy course interventions alter pedagogical and social behavior among students, student-instructors and faculty. In an ethnographic study, I suspect that coding for the Authority and Community Lens might produce different results because I can focus on how instructors interact with their peers, other instructors, faculty, and community members.

I also plan to conduct inter-rater reliability on my current qualitative analyses and future analyses. Conducting such research would allow me to refine my coding and better define subcodes related to sharing the responsibility for learning and using multiple data sources.

Moving forward, I plan to go beyond low- and high-growth group analyses by using growth mixture models. This innovative approach allows me to see multiple groups of instructors based on patterns of growth. In this model, instructors can have features about their growth that are shared with the group, and they can have their own unique features. This would allow me to hone in on the unique growth patterns of minoritized instructors across groups, which is key for better understanding the learning trajectories and pathways of college instructors who are underrepresented in the academy.

In terms of inferential qualitative analyses, I can qualitatively code reflections and documents from groups of instructors who share similar profiles, and examine the specific trajectories of individual instructors. I hope to use this research to provide more targeted support, resources and training for instructors, which can promote further growth. This approach has not been done yet when researching graduate and faculty development programs. It is important to pursue this research because I want to delve deeper into how minoritized instructors learn in higher education.

8.4 Implications for Practice, Research, Policy, and Organizational Change

Practice, research, and policy. My work can speak more generally to practice, research, and policy. Based on my research, I would recommend the following:

1) Advance adaptive professional development programs that leverage the AEP framework to support educators' success in higher education. These programs need to identify and respond to instructors' different lenses, such as the ones mentioned in this study. In these programs, I recommend foregrounding AEP concepts like sharing the responsibility for learning and using multiple data sources to inform practice as well as associated topics such as race, racialization of standards for success, intersectionality, and the myth of meritocracy. I also suggest that these programs equip instructors with automated formative assessment tools as described in recommendation # 2 below (e.g., clickers, Poll Everywhere, surveys on equity barriers) and different activities (e.g., like the worksheets mentioned in Appendix A) on how to target learners' needs in a larger classroom. In addition, it would be important to develop a technological tool to flag formative and summative assessment items that exhibit bias against student demographic subgroups (e.g., gender, ethnicity, disability status). Such a tool can draw on differential item functioning, which was described in chapter 5. Further research can be conducted with populations with a larger range of teaching experience and faculty.
2) Automate formative assessment processes for instructors with equity-focused dashboard options that break down assessment data by different learner subgroups. It would also be interesting to pair closed- and open-ended survey data on equity barriers with these assessment data trends. I also see the value in building an Artificial Intelligence (AI) system that can make pedagogical recommendations to instructors based on the AEP

framework. I would suggest conducting research on the extent that this process can help instructors better notice and address equity barriers to learning. Future research should consider how to develop reflection assignments that can be fed into a technology (e.g., an AI) that can present the most frequent themes from instructors' reflections. It is important to train and assess the AI to ensure it does not reproduce racial and other biases.

3) Expand research on different ways to measure equitable pedagogy in teaching portfolios for higher education teaching evaluation. I would suggest exploring how measures such as AEPC can support instructors in documenting inclusive practices in merit review and professional development programs. During such a study, it would be helpful to find a way to automate the scoring of AEPC assessments. This research can support educational institutions that have recommended institutionalizing equity-oriented teaching, mentoring, research, and service in hiring, merit reviews and promotion processes for faculty (UC Berkeley Hispanic Serving Institution Taskforce, 2020; Chicanx and Latinx Standing Committee Recommendations, 2022). These institutionalization efforts would need to be grounded in well validated equity-oriented teaching and mentoring measures that mitigate bias against instructors of color and women commonly found in course evaluations (MacNell et al., 2014; Phuong et al., 2022). Institutionalizing equity-oriented practices in hiring and promotion criteria can incentivize employees to adopt equity-oriented practices because these criteria are being assessed as a part of the job. This recommendation can potentially motivate instructors to apply equity-oriented practices that have been shown to improve student engagement and success. The implementation of these recommendations can promote organizational and cultural transformation that advances equity.
4) Expand, assess, and improve measures for evaluating equitable pedagogies in academic departments. For example, using a data-driven approach (e.g., student performance data from campus), academic departments need to examine and address the opportunity and equity gaps in courses. As one measure of effective teaching, I recommend examining if instructors' pedagogies are reducing or exacerbating these gaps. Doing so may foster a culture that values longer-term student success, and practices such as tracking the impact of courses on student outcomes (e.g., academic achievement, persistence, retention) in subsequent courses. It is important to critically examine if and why equity gaps persist in future courses, and to identify ways this information can inform and strengthen the evaluation of teaching effectiveness. Additionally, I recommend capturing formative assessment data (e.g., surveys) on all students' experiences, including minoritized students, to understand and address barriers and opportunities for student engagement, success, and sense of belonging. Academic departments should use these data to better support faculty and services with the goal of increasing every student's success (defined broadly) and transforming departmental culture. Such data, combined with data on student profiles and outcomes, might provide meaningful insights on diverse trajectories of student success and areas for growth. Given the complexity of the data involved, it may be reasonable to assume that AI or related technology can be leveraged to identify the patterns and suggest courses of action in real time. However it is accomplished (e.g., through AI or other means), connecting members of the university community with each other (e.g., creating mentor-mentee relationships) can promote a greater sense of belonging and community. If AI tools are to be employed, I will continue to emphasize

that researchers and practitioners need to examine and address any bias, racialization, and ableism among other issues embedded within them.

In relation to AEP, it is possible that AI and related technologies could: a) discover and leverage patterns associated with successful learning, b) improve the automation of on-demand and automatically scored formative assessments; and c) identify and intervene, perhaps in culturally responsive ways, in response to early warnings signals present in complex data from the learning ecosystem. Efforts to leverage AI must, however, examine carefully the ways in which underlying algorithms are 'learned' in order to avoid the reinforcement of systemic bias and racism that permeate educational systems (Cheuk, 2021).

Organizational change. I now discuss an organizational change framework that informed how I worked with stakeholders (e.g., senior administrators, faculty) in my study. I present this framework because it has implications for how researchers and educational developers can collaborate with practitioners to advance equity in student outcomes. I mention this framework to be wholly transparent about my process of working with stakeholders.

My equity-oriented organizational change framework entails being sensitive to contextual factors and collaborating with stakeholders to:

- Work in partnership to clarify goals and define intended outcomes with an equity lens
    - Consider a theory of change when thinking about goals
- Diagnose where we are relative to those outcomes
    - Identify strengths, interests, barriers, areas for growth
    - Consider different perspectives and measures of success. Collaboratively define, implement, and refine measures of success.
    - Avoid placing the responsibility for learning and success solely on one group; the responsibility for learning much be shared among stakeholders in a coherent and unified way
- Design and adapt plans based on the diagnosis
- Iterate: Demonstrate a continuous commitment to evaluation, adaptation, and improvement
    - Focus on transforming institutions, structures, and systems to better support student, staff, faculty, and community success.

I have leveraged this framework when collaborating with faculty and administrator stakeholders in this study. I propose drawing and building on this framework to work in partnership with communities to advance inclusive excellence, showcase equity champions to make an inclusive culture visible, and support initiatives that promote data-informed practices to increase equity in student outcomes.

I now move beyond the dissertation to discuss how department action teams can promote equitable organizational change (Ngai et al., 2020; Reinholz et al., 2020). For educational institutions interested in department level change and diffusing innovations (Henderson et al., 2012; Henderson & Dancy, 2011), I recommend the idea of an equity-oriented department action team. In my view, an equity-oriented department action team leverages a collaborative approach where all stakeholders share the responsibility for learning when addressing educational inequities. Creating norms where stakeholders share the responsibility for learning with each other is key because there needs to be a unified and collaborative effort in order to make progress towards equitable and cultural transformation. It is insufficient to expect a diversity officer, faculty member, student, or a staff member to enact organizational change alone.

It is helpful to have a "hub" of stakeholders who champion equity-oriented practices and co-construct a shared vision for advancing inclusive excellence aligned with campus goals. This "hub" of stakeholders can identify, partner with, support, and socialize early adopters with equity-oriented frameworks. These early adopters can form a department action team that performs multiple functions like disseminating and customizing equitable innovations and training for their department or campus unit. The department action team would need to identify key goals, opportunities, challenges, and a theory of change (Ngai et al., 2020). The "hub" of stakeholders can act as a steering committee to provide consultation and support to the department action teams across campus to ensure alignment towards equity goals. With support from the "hub", the department action team can examine synergies and gaps to bring together other "partners" (i.e., stakeholders, committees, initiatives) to identify the role the team and "partners" can play in moving the department and campus towards equity and inclusion goals. For example, the team can bring together faculty to examine departmental climate, create vertical and horizontal alignment in program curriculum, and implement strategies for addressing equity barriers in the department. I suggest that the department action teams collaborate and align educational development efforts with relevant student, faculty, and staff support services. This would include partnering with departments and key stakeholders in creating new initiatives and seeking funding from both internal and external organizations to advance organizational and systemic change.

To make a commitment to equity more visible, the "hub" and department action teams can collaborate with communications teams and stakeholders to disseminate reports and showcase department action teams through media, events, and fundraising efforts that engage communities. In this spirit, I collaboratively drafted recommendations for executive leaders to advance equity, inclusion, belonging, and an anti-racist campus. Drawing on these recommendations, I specifically suggest that educational leaders:

1. Fund and support "equity-oriented department action teams that partner with departments / units on campus to use the equity-oriented framework to address 1-2 opportunities or problems of practice" (Phuong et al., 2022, p. 1).
2. Fund full-time positions for core faculty and coaches to lead "department action teams, programs, and provide individual consultations" (Phuong & Equity-Oriented Advising and Coaching Program, 2022, p. 1). For example, these "core faculty and coaches can partner with department leaders to customize a program where the department applies the equity-oriented framework to collect data and address a problem of practice." (Phuong et al., 2022, p. 1).
    a. Program participants can "discuss shared insights and data on 1-2 problems of practice, apply strategies to address them, and would support each other in transforming their practices based on what is working well and not working well" (Phuong et al., 2022, p. 1).
    b. These programs can train "departments and units to use an equity-oriented framework that has been shown to increase student engagement, success, and community (Phuong et al. 2017; Phuong et al., 2022)" (Phuong et al., 2022, p. 1).

Equity-oriented department action teams can help diffuse innovations and create momentum for transformative and cultural change (Henderson et al., 2012; Henderson & Dancy, 2011).

In conclusion, this dissertation research shows that inclusive teaching competencies can be tracked over time on equity-oriented assessments that do not exhibit gender and URM bias, unlike course evaluations which have been shown to demonstrate such bias (Boring et al., 2016;

MacNell et al., 2015; Phuong et al., 2022; Stark & Freishtat, 2014). This study contributes insights on how to support instructors with learning and adopting inclusive practices, thereby addressing a long-standing barrier to pedagogical reform in STEM (Phuong et al., 2022). This research provides evidence-based approaches to support instructors on how to reflect on and use multiple forms of data effectively—rather than relying on their assumptions—to improve student success (Phuong et al., 2022). If we are to create equitable learning environments, we need to support instructors in making data-driven decisions that address student learning and equity barriers. Moreover, the lenses described in my study can help us better understand instructors' reflective and learning processes. This insight can help us expand instructors' critical consciousness around student learning, so they do not adopt individualistic and deficit models of learning where the responsibility for learning is solely placed on the student. As experimental research on AEP has shown (Phuong et al., 2017; Phuong & Nguyen, 2019), expanding instructors' perspectives and helping them make more targeted instructional decisions based on student data is a key driver to reducing equity barriers and significantly improving student success. The AEP framework, discussions of meritocracy and the racialization of education, and the adaptive pedagogy course model can be beneficial for the implementation of effective faculty development programs and the advancement of equity in student outcomes.

REFERENCES

Adams, R. J., Wu, M. L., & Wilson, M. (2012). A*CER ConQuest 3.01: Generalized Item Response Modeling Software* [Computer software and manual]. Melbourne, Victoria, Australia: Australian Council for Educational Research.

Altman, D. G. (1999). *Practical statistics for medical research*. Chapman & Hall/CRC, London.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). AERA, APA, & NCME. American Educational Research Association.

Amundsen, C., & Wilson, M. (2012). Are we asking the right questions? A conceptual review of the educational development literature in higher education. *Review of Educational Research*, *82*(1), 90-126.

Anderson, C., Turner, A. C., Heath, R. D., & Payne, C. M. (2016). On the meaning of grit… and hope… and fate control… and alienation… and locus of control… and… self-efficacy… and… effort optimism… and…. *The Urban Review*, *48*(2), 198-219.

Baker, P. (1999). Creating learning communities: The unfinished agenda. *The Social Works of Higher Education*, 95-109.

Bernal, D. D. (2002). Critical race theory, Latino critical theory, and critical raced-gendered epistemologies: Recognizing students of color as holders and creators of knowledge. *Qualitative inquiry*, *8*(1), 105-126.

Birt, J. A., Khajeloo, M., Rega-Brodsky, C. C., Siegel, M. A., Hancock, T. S., Cummings, K., & Nguyen, P. D. (2019). Fostering agency to overcome barriers in college science teaching: Going against the grain to enact reform-based ideas. *Science Education*, *103*(4), 770–798.

Black, P., & Wiliam, D. (2001). *Inside the black box: Raising standards through classroom assessment.* Northumberland LEA.

Blum, A. M., Mason, J. M., Kim, J., & Pearson, P. D. (2020). Modeling question-answer relations: the development of the integrative inferential reasoning comic assessment. *Reading and Writing*, 1-30.

Bolman, L. G., & Deal, T. E. (2008). *Reframing organizations: Artistry, choice, and leadership*. Jossey-Bass.

Bonilla-Silva, E. (1999). The essential social fact of race. *American Sociological Review*, *64*(6), 899-906.

Boring, A., Ottoboni, K., & Stark, P. (2016). Student evaluations of teaching (mostly) do not measure teaching effectiveness. *ScienceOpen Research*.

Bourdieu, P. (1973). Cultural reproduction and social reproduction. In R. Brown (Ed.), Knowledge, Education and Social Change (pp. 56-68). London: Tavistock.

Boyer, E. L. (1990). *Scholarship reconsidered: Priorities of the professoriate*. Carnegie Foundation for the Advancement of Teaching.

Brockenbrough, E. (2016). Becoming queerly responsive: Culturally responsive pedagogy for Black and Latino urban queer youth. *Urban Education*, *51*(2), 170-196.

Brown, J., & Kurzweil, M. (2017). Instructional quality, student outcomes, and institutional finances. *American Council on Education (ACE): Leadership and Advocacy*.

Canning, E. A., Muenks, K., Green, D. J., & Murphy, M. C. (2019). STEM faculty who believe ability is fixed have larger racial achievement gaps and inspire less student motivation in their classes. *Science Advances, 5*(2), eaau4734.

CAST (2018). *Universal Design for Learning Guidelines* version 2.2. Retrieved from http://udlguidelines.cast.org

Chang, W. C. C., Ludlow, L. H., Grudnoff, L., Ell, F., Haigh, M., Hill, M., & Cochran-Smith, M. (2019). Measuring the complexity of teaching practice for equity: Development of a scenario-format scale. *Teaching and Teacher Education, 82,* 69-85.

Chang, W.C. (2022). Measuring the complexity of equity-centered teaching practice development and validation of a Rasch/Guttman scenario scale. *Journal of Applied Measurement*, 22(1/2), 35-59.

Chapman, P. D. (1988). *Schools as sorters: Lewis M. Terman, applied psychology, and the intelligence testing movement, 1890-1930*. New York University Press.

Cheuk, T. (2021). Can AI be racist? Color-evasiveness in the application of machine learning to science assessments. *Science Education*, *105*(5), 825-836.

Chicanx and Latinx Standing Committee Recommendations. (2022). *Report for UC Berkeley.*

Clark, C. R., Mercer, S. H., Zeigler-Hill, V., & Dufrene, B. A. (2012). Barriers to the success of ethnic minority students in school psychology graduate programs. *School Psychology Review*, *41*(2), 176-192.

Cochrane, S. V., Chhabra, M., Jones, M. A., & Spragg, D. (Eds.). (2017). *Culturally responsive teaching and reflection in higher education: Promising practices from the cultural literacy curriculum institute*. Taylor & Francis.

Cole, D., & Espinoza, A. (2008). Examining the academic success of Latino students in science technology engineering and mathematics (STEM) majors. *Journal of College Student Development, 49*(4), 285-300.

Condon, W., Iverson, E. R., Manduca, C. A., Rutz, C., & Willett, G. (2016). *Faculty development and student learning: Assessing the connections*. Indiana University Press.

Cox, M. D. (2004). Introduction to faculty learning communities. *New Directions for Teaching and Learning*, *2004*(97), 5-23.

Dalrymple, S. E., Auerbach, A. J., & Schussler, E. E. (2017). Taking a community approach to curriculum change. *International Journal for the Scholarship of Teaching and Learning*, 11(2). https://eric.ed.gov/?id=EJ1150429.

Davis, B. G. (2009). *Tools for teaching*. John Wiley & Sons.

De Boeck, P. & Wilson, M. (Eds.). (2004). *Explanatory item response models: A generalized linear and nonlinear approach.* Springer-Verlag.

Derting, T. L., Ebert-May, D., Henkel, T. P., Maher, J. M., Arnold, B., & Passmore, H. A. (2016). Assessing faculty professional development in STEM higher education: Sustainability of outcomes. *Science Advances*, *2*(3), e1501422, 1-10.

Deslauriers, L., McCarty, L. S., Miller, K., Callaghan, K., & Kestin, G. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *Proceedings of the National Academy of Sciences*, *116*(39), 19251-19257.

Deslauriers, L., Schelew, E., & Wieman, C. (2011). Improved learning in a large-enrollment physics class. *Science*, *332*(6031), 862-864.

DiMaggio, P. J., & Powell, W. W. (1991). *The iron cage revisited.*

Duckor, B., & Holmberg, C. (2019). Exploring how to model formative assessment trajectories of posing-pausing-probing practices: Toward a teacher learning progressions framework for the study of novice teachers. *Journal of Educational Measurement, 56,* 836–890.

Duffy, D. K., & Jones, J. W. (1995). *Teaching within the Rhythms of the Semester. The Jossey-Bass Higher and Adult Education Series*. Jossey-Bass Inc.

Duncan, R. G., & Rivet, A. E. (2013). Science learning progressions. *Science, 339*(6118), 396-397.

Durkheim, E. (2004). *Readings from Emile Durkheim*. Psychology Press.

Ebert-May, D., Derting, T. L., Henkel, T. P., Middlemis Maher, J., Momsen, J. L., Arnold, B., & Passmore, H. A. (2015). Breaking the cycle: Future faculty begin teaching with learner-centered strategies after professional development. CBE—Life Sciences Education, 14(2), ar22.

Enderle, P. J., Southerland, S. A., & Grooms, J. A. (2013). Exploring the context of change: Understanding the kinetics of a studio physics implementation effort. *Physical Review Special Topics - Physics Education Research, 9*(1), 010114.

Engeström, Y. (1987). *Learning by expanding: An activity theoretical approach to developmental research*. Orienta-konsultit.

Esterhazy, R., de Lange, T., Bastiansen, S., & Wittek, A. L. (2021). Moving beyond peer review of teaching: a conceptual framework for collegial faculty development. *Review of Educational Research*, *91*(2), 237-271.

Fairweather, J. (2008). Linking evidence and promising practices in science, technology, engineering, and mathematics (STEM) undergraduate education. *Board of Science Education, National Research Council, The National Academies, Washington, DC*.

Felder, R. M. (2007). Research on teaching and learning in engineering. *Academy of Chemical Engineers Award Lecture, Department of Chemical and Biological Engineering, University of Missouri-Rolla*.

Feldman, J. (2018). *Grading for equity: What it is, why it matters, and how it can transform schools and classrooms*. Corwin Press.

Ferrare, J. J. (2019). A multi-institutional analysis of instructional beliefs and practices in gateway courses to the sciences. *CBE—Life Sciences Education, 18*(2), ar26.

Fink, L. D. (2013). *Creating significant learning experiences: An integrated approach to designing college courses.* John Wiley & Sons.

Fowler, L. (2021). How to implement policy: Coping with ambiguity and uncertainty. *Public Administration, 99*(3), 581-597.

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.

Garcia, G. A., & Okhidoi, O. (2015). Culturally relevant practices that "serve" students at a Hispanic Serving Institution. *Innovative Higher Education, 40*(4), 345-357.

García, S. B., & Ortiz, A. A. (2006). Preventing disproportionate representation: Culturally and linguistically responsive prereferral interventions. *Teaching Exceptional Children, 38*(4), 64-68.

Gay, G. (2010). *Culturally responsive teaching: Theory, research, and practice*. Teachers College Press.

Gehrke, S., & Kezar, A. (2016). STEM reform outcomes through communities of transformation. *Change: The Magazine of Higher Learning*, 48(1), 30–38.

Gehrke, S., & Kezar, A. (2017). The Roles of STEM Faculty Communities of Practice in Institutional and Departmental Reform in Higher Education. *American Educational Research Journal*, 54(5), 803–833.

Gess-Newsome, J., Southerland, S. A., Johnston, A., & Woodbury, S. (2003). Educational reform, personal practical theories, and dissatisfaction: The anatomy of change in college science teaching. *American Educational Research Journal, 40*(3), 731–767.

Glas, C. (2008). Item response theory in educational assessment and evaluation. *Mesure et évaluation en éducation, 31*(2), 19-34.

Goode, J., Ivey, A., Johnson, S. R., Ryoo, J. J., & Ong, C. (2021). Rac (e) ing to computer science for all: how teachers talk and learn about equity in professional development. *Computer Science Education, 31*(3), 374-399.

Goodwin, C. (1994). Professional vision. *American Anthropologist, 96*(3), 606e633.

Grande, S. (2015). *Red pedagogy: Native American social and political thought*. Rowman & Littlefield.

Griful-Freixenet, J., Struyven, K., Verstichele, M., & Andries, C. (2017). Higher education students with disabilities speaking out: Perceived barriers and opportunities of the universal design for learning framework. *Disability & Society, 32*(10), 1627-1649.

Gutiérrez, K. D., Engeström, Y., & Sannino, A. (2016). Expanding educational research and interventionist methodologies. *Cognition and Instruction, 34*(3), 275-284.

Hambleton, R. K., & Van der Linden, W. J. (1982). Advances in item response theory and applications: An introduction. *Applied Psychological Measurement, 6*(4), 373-378.

Hammer, D., Elby, A., Scherr, R. E., & Redish, E. F. (2005). Resources, framing, and transfer. Transfer of learning from a modern multidisciplinary perspective, 89.

Hammond, Z. (2014). *Culturally responsive teaching and the brain: Promoting authentic engagement and rigor among culturally and linguistically divese students.* Corwin Press.

Harrington, C., & Zakrajsek, T. (2017). Dynamic Lecturing: Research-based Strategies to Enhance Lecture Effectiveness. Stylus Publishing, LLC.

Harrison-Bernard, L. M., Augustus-Wallace, A. C., Souza-Smith, F. M., Tsien, F., Casey, G. P., & Gunaldo, T. P. (2020). Knowledge gains in a professional development workshop on diversity, equity, inclusion, and implicit bias in academia. *Advances in Physiology Education, 44*(3), 286-294.

Henderson, C., & Dancy, M. H. (2011). Increasing the impact and diffusion of STEM education innovations. A white paper commissioned for the Characterizing the Impact and Diffusion of Engineering Education Innovations forum, held February 7–8, 2011. *New Orleans, LA. www. nae. edu/File. Aspx*.

Henderson, C., Dancy, M., & Niewiadomska-Bugaj, M. (2012). Use of research-based instructional strategies in introductory physics: Where do faculty leave the innovation-decision process?. *Physical Review Special Topics-Physics Education Research, 8*(2), 020104.

Herman, G. L., Mena, I. B., Greene, J. C., West, M., Tomkin, J., & Mestre, J. (2015). Creating institution-level change in instructional practices through faculty communities of practice. In *Proc. 122nd American Society for Engineering Education Annual Conf. and Exposition (ASEE)*, (vol. 26, pp. 1–26).

Hessler, M., Pöpping, D. M., Hollstein, H., Ohlenburg, H., Arnemann, P. H., Massoth, C., ... & Wenk, M. (2018). Availability of cookies during an academic course session affects evaluation of teaching. *Medical Education, 52*(10), 1064-1072.

Howard, T. O., Winkelmes, M. A., & Shegog, M. (2020). Transparency teaching in the virtual classroom: Assessing the opportunities and challenges of integrating transparency

teaching methods with online learning. *Journal of Political Science Education*, *16*(2), 198-211.
Hudson, K., & Coukos, A. (2005). The dark side of the Protestant ethic: A comparative analysis of welfare reform. *Sociological Theory*, *23*(1), 1-24.
Hutchins, E. (2000). Distributed cognition. *International Encyclopedia of the Social and Behavioral Sciences. Elsevier Science*, *138*.
Jaber, L. Z., & Hammer, D. (2016). Learning to feel like a scientist. *Science Education*, *100*(2), 189-220.
Jackson, K., & Cobb, P. (2010, April). Refining a vision of ambitious mathematics instruction to address issues of equity. In *annual meeting of the American Educational Research Association, Denver, CO*.
Jarvis, P. (1987). *Adult learning in the social context.* Croom Helm.
Kearns, K., Hatcher, M., Bollard, M., DiPietro, M., Donohue-Bergeler, D., Drane, Leslie., Luoma, E., Phuong, A. E., Thain, L., & Wright, M. (2018). "Once a scientist…:" Disciplinary approaches and intellectual dexterity in educational development. *To Improve the Academy: A Journal of Educational Development*. *37*(1), 128-141.
Kendi, I. X. (2019). *How to be an antiracist*. One world.
Kezar, A., & Bernstein-Serra, S. (2020). Scaling Culture Change through Institutional Logics: A Look at the American Association of Universities (AAU) STEM Initiative. *Peabody Journal of Education*, *95*(4), 356-373.
Kinder, D. R., & Sears, D. O. (1981). Prejudice and politics: Symbolic racism versus racial threats to the good life. *Journal of personality and social psychology*, *40*(3), 414.
Kisiel, J. F. (2010). Exploring a school–aquarium collaboration: An intersection of communities of practice. *Science Education*, *94*(1), 95-121.
Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal*, *32*(3), 465-491.
Laghari, A., Ansari, K., & Bughio, F. A. (2017). The Effectiveness of 'How People Learn'(HPL) Framework Via the Star (Software Technology for Action and Reflection) Legacy Cycle of Learning: A Case Study of Undergraduate Students at Institute of English Language and Literature, University of Sindh, Jamshoro, Pakistan. *Grassroots*, *50*(2).
Landis, J. R., & Koch, G. G. (1977). An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics*, 363-374.
Lardier Jr., D. T., Herr, K. G., Barrios, V. R., Garcia-Reid, P., & Reid, R. J. (2019). Merit in meritocracy: Uncovering the myth of exceptionality and self-reliance through the voices of urban youth of color. *Education and Urban Society*, *51*(4), 474-500.
Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge university press.
LeCompte, M. D., & Schensul, J. J. (1999). *Analyzing & interpreting ethnographic data* (Vol. 5). Rowman Altamira.
Lederman, N. G. (2018). Nature of scientific knowledge and scientific inquiry in biology teaching. In *Teaching biology in schools* (pp. 216-235). Routledge.
Levinson-Rose, J., & Menges, R. J. (1981). Improving college teaching: A critical review of research. *Review of Educational Research*, *51*(3), 403-434.
Little, J. W. (2002). Locating learning in teachers' communities of practice: Opening up problems of analysis in records of everyday work. *Teaching and Teacher Education*, *18*(8), 917-946.

Lord, F. M., Novick, M. R., & Birnbaum, A. (1968). *Statistical theories of mental test scores.* Addison-Wesley.
Lund, T. J., & Stains, M. (2015). The importance of context: An exploration of factors influencing the adoption of student-centered teaching among chemistry, biology, and physics faculty. *International Journal of STEM Education, 2*(1), 13.
MacNell, L., Driscoll, A., & Hunt, A. N. (2015). What's in a name: Exposing gender bias in student ratings of teaching. *Innovative Higher Education, 40*(4), 291-303.
Marshall, D. A. (2002). Behavior, belonging, and belief: A theory of ritual practice. *Sociological Theory, 20*(3), 360-380.
Markus, H. R., & Conner, A. (2013). *Clash!: 8 cultural conflicts that make us who we are* (p. 320). New York, NY: Hudson Street Press.
Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47*(2), 149-174.
Masters, G. N. (2010). The Partial Credit Model. *Handbook of Polytomous Item Response Theory Models*, 109-122.
Mazur, E. (2009). Farewell, lecture?. *Science*, *323*(5910), 50-51.
McCallum, R. (2013). *Diagnostic Assessment and Strategic Teaching*. Personal Collection of R. McCallum, University of California Berkeley, Berkeley, CA.
McCoy, H., & Bocala, C. (2022). Building Virtual Communities of Practice for Equity in Education. In *Teaching and Learning for Social Justice and Equity in Higher Education* (pp. 187-210). Palgrave Macmillan, Cham.
Metevier, A. J., Hunter, L., Goza, B. K., Raschke, L. M., & Seagroves, S. (2010). Improvements in professional development program participants' understandings about inclusive teaching. In *Learning from Inquiry in Practice* (Vol. 436, p. 515).
Moll, L. C., Amanti, C., Neff, D., & Gonzalez, N. (1992). Funds of knowledge for teaching: Using a qualitative approach to connect homes and classrooms. *Theory into practice*, *31*(2), 132-141.
Montuori, A., & Purser, R. E. (1995). Deconstructing the lone genius myth: Toward a contextual view of creativity. *Journal of Humanistic Psychology*, *35*(3), 69-112.
Mulnix, A. B. (2016). STEM faculty as learners in pedagogical reform and the role of research articles as professional development opportunities. *CBE—Life Sciences Education*, *15*(4), es8, 1-9.
National Research Council. (1999). *How people learn: Bridging research and practice*. National Academies Press.
Ngai, C., Pilgrim, M. E., Reinholz, D. L., Corbo, J. C., & Quan, G. M. (2020). Developing the DELTA: Capturing cultural changes in undergraduate departments. *CBE—Life Sciences Education, 19*(2), ar15.
Nguyen, J., Phuong, A. E., Hunn, C. T., & Mejia, F. D. (2021). Zone of Proximal Self: A Sociocultural Framework for Examining the Development of Possible Selves and Social-Emotional Competencies. In de Vries, E., Hod, Y., & Ahn, J. (Eds.), *Proceedings of the 15th International Conference of the Learning Sciences* - ICLS 2021. (pp. 1173-1174). Bochum, Germany: International Society of the Learning Sciences.
Nguyen, J., Phuong, A. E., Salehi, S., Mejia, F. D., & Hagstrom, S. (2022). The Zone of Proximal Self Model to Improve Equity in Higher Ed. In *Proceedings of the 16th International Conference of the Learning Sciences* - ICLS 2022. (pp. 767-774). Hiroshima, Japan (Online Conference): International Society of the Learning Sciences.

North, J., & Scholl, S. C. (1979). POD: The founding of a national network. *POD Quarterly: The Journal of the Professional and Organizational Development Network in Higher Education,* 6, 10-17. https://digitalcommons.unl.edu/podqtrly/6

O'Leary, E. S., Shapiro, C., Toma, S., Sayson, H. W., Levis-Fitzgerald, M., Johnson, T., & Sork, V. L. (2020). Creating inclusive classrooms by engaging STEM faculty in culturally responsive teaching workshops. *International Journal of STEM Education*, 7(1), 1-15.

Pace, D., & Schwartz, D. (2008). Accessibility in Post Secondary Education: Application of UDL to College Curriculum. *Online Submission*, *5*(12), 20-26.

Paek, I., & Wilson, M. (2011). Formulating the Rasch differential item functioning model under the marginal maximum likelihood estimation context and its comparison with Mantel–Haenszel procedure in short test and small sample conditions. *Educational and Psychological Measurement, 71*(6), 1023-1046.

Pelletreau, K. N., Knight, J. K., Lemons, P. P., McCourt, J. S., Merrill, J. E., Nehm, R. H., … Smith, M. K. (2018). A faculty professional development model that improves student learning, encourages active-learning instructional practices, and works for faculty at multiple institutions. *CBE—Life Sciences Education*, 17(2), es5.

Parker, C., Morrell, C., Morrell, C., & Chang, L. (2016). Shifting understandings of community college faculty members: Results of an equity-focused professional development experience. *The Journal of Faculty Development*, 30(3), 41-48.

Phuong, A. E. (2021a). Investigating the Cognitive, Emotional, and Socio-Organizational Mechanisms that Drive College STEM Instructional Improvement. *Proceedings of the First Annual Meeting of the International Society of the Learning Sciences Conference,* 2021. (pp. 135-136). Bochum, Germany: International Society of the Learning Sciences.

Phuong, A. E. (2021b). Adaptive Equity-Oriented Pedagogy: Equipping Educators to Overcome Barriers to Learning in STEM. Teaching Effectiveness Award Essay. UC Berkeley GSI Teaching and Resource Center.

Phuong, A. E., & Nguyen, J. (2019). Evaluating an Adaptive Equity-Oriented Pedagogy on Student Collaboration Outcomes Through Randomized Controlled Trials. In Lund, K., Niccolai, G. P., Lavoué, E., Hmelo-Silver, C., Gweon, G., & Baker, M. (Eds.), *A Wide Lens: Combining Embodied, Enactive, Extended, and Embedded Learning in Collaborative Settings, 13th International Conference on Computer Supported Collaborative Learning (CSCL) 2019, Volume 1* (pp. 496-503). International Society of the Learning Sciences.

Phuong, A. E., Nguyen, J., Hunn, C. T., & Mejia, F. D. (2021a). A Learning Sciences and Organizational Behavior Framework for Analyzing How College Instructors Learn Inclusive Pedagogies. In de Vries, E., Hod, Y., & Ahn, J. (Eds.), *Proceedings of the 15th International Conference of the Learning Sciences* - ICLS 2021. (pp. 1171-1172). International Society of the Learning Sciences.

Phuong, A. E., Nguyen, J. & Marie, D. (2017a). Evaluating an adaptive equity-oriented pedagogy: A study of its impacts in higher education. *The Journal of Effective Teaching, 17*(2), 5-44.

Phuong, A. E., Nguyen, J. & Marie, D. (2017b). Conceptualizing an adaptive and data-driven equity-oriented pedagogy. *Transformative Dialogues: Teaching & Learning Journal*, *10*(2).

Phuong, A. E., Nguyen, J., Mejia, F., Hunn, T., & Marie, D. (2021b). Equitable teaching that creates pathways to success for all students. *Times Higher Education.*

Phuong, A. E., Nguyen, J., Hunn, C. T., Vo, J., Mejia, F. D., & Huang, F. (2022). Examining Instructors' Adaptive Equity-Oriented Pedagogical Competency: A Validated Measure that Promotes College Students' Success. In *Proceedings of the 16th International Conference of the Learning Sciences* - ICLS 2022. (pp. 1645-1648). Hiroshima, Japan (Online Conference): International Society of the Learning Sciences.

Phuong, A. E., Nguyen, J., Vo, J., Hunn, C. T., & Mejia, F. D. (2022). Increasing College STEM Instructors' Equity-Oriented Teaching Competencies. In *Proceedings of the 16th International Conference of the Learning Sciences* - ICLS 2022. (pp. 1665-1668). Hiroshima, Japan (Online Conference): International Society of the Learning Sciences.

Phuong, A. E., Peña-Rager, I., Mejia, F., Naranjo, L., Vivar, D., Nguyen, N., Yaden, Y., Chiarelli, J., Velazques, M., Gerker-Junsay, J., Hunn, C., Kuah, Y., Pantig-Naval, K., Jackson, L., Calhoun, S., & Equity-Oriented Advising and Coaching Program (2022). Recommendations to Executive Leaders and Administrators at UC Berkeley: Equity-Oriented Advising and Coaching Opportunities and Recommendations. Internal Document at UC Berkeley.

POD Network. (2016). What is educational development?. The Professional and Organizational Development (POD) Network in Higher Education. Retrieved February 1, 2022.

Podolsky, A., Darling-Hammond, L., Doss, C., & Reardon, S. (2019). *California's positive outliers: Districts beating the odds*. Learning Policy Institute.

Rämö, J., Reinholz, D., Häsä, J., & Lahdenperä, J. (2019). Extreme apprenticeship: Instructional change as a gateway to systemic improvement. *Innovative Higher Education*.

Reinholz, D. L., & Apkarian, N. (2018). Four frames for systemic change in STEM departments. *International Journal of STEM Education*, *5*(1), 3.

Reinholz, D. L., Matz, R. L., Cole, R., & Apkarian, N. (2019). STEM is not a monolith: A preliminary analysis of variations in STEM disciplinary cultures and implications for change. *CBE—Life Sciences Education*, *18*(4), mr4.

Reinholz, D. L., Ngai, C., Quan, G., Pilgrim, M. E., Corbo, J. C., & Finkelstein, N. (2019). Fostering sustainable improvements in science educations: An analysis through four frames. *Science Education*, *103*(5), 1125–1150.

Reinholz, D. L., Pawlak, A., Ngai, C., & Pilgrim, M. (2020). Departmental action teams: Empowering students as change agents in academic departments. *International Journal for Students as Partners*, *4*(1), 128-137.

Reinholz, D. L., White, I., & Andrews, T. (2021). Change theory in STEM higher education: a systematic review. *International Journal of STEM Education*, *8*(1), 1-22.

Robert, J., & Carlsen, W. S. (2017). Teaching and research at a large university: Case studies of science professors. *Journal of Research in Science Teaching*, *54*(7), 937-960.

Rogoff, B. (1995). Observing sociocultural activity on three planes: Participatory appropriation, guided participation, and apprenticeship. In J. V. Wertsch, P. del Rio, & A. Alvarez (Eds.), Sociocultural studies of mind (pp. 139–164). Cambridge University Press.

Rooney, S., Enszer, J., Maresca, J., Shah, S. I., Hewlett, S., & Buckley, J. (2020, May). Faculty Development Mini-Modules on Evidence-Based Inclusive Teaching and Mentoring Practices in Engineering. In *2020 ASEE Virtual Annual Conference*.

Rose, D. H., Harbour, W. S., Johnston, C. S., Daley, S. G., & Abarbanell, L. (2006). Universal design for learning in postsecondary education: Reflections on principles and their application. *Journal of Postsecondary Education and Disability*, *19*(2), 135-151.

Saldana, J. (2011). *Fundamentals of qualitative research*. Oxford university press.

Sax, L., Blaney, J., Lehman, K., Rodriguez, S., George, K., & Zavala, C. (2018). Sense of belonging in computing: the role of introductory courses for women and underrepresented minority students. *Social Sciences, 7*(8), 122.

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction*, *16*(4), 475-522.

Schön, D. A. (1987). *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions*. Jossey-Bass.

Scott, R. (2006). *Organizations: Rational, natural, and open systems*. Prentice Hall.

Sears, D. O., & Henry, P. J. (2003). The origins of symbolic racism. *Journal of Personality and Social Psychology*, *85*(2), 259.

Seldin, P., Miller, J. E., & Seldin, C. A. (2010). *The teaching portfolio: A practical guide to improved performance and promotion/tenure decisions*. John Wiley & Sons.

Seymour, E., & Hewitt, N. (1997). *Talking about leaving: why undergraduates leave the sciences*. Westview Press.

Sherin, B., & Star, J. R. (2011). Reflections on the study of teacher noticing. *Mathematics Teacher Noticing: Seeing through Teachers' Eyes*, *66*.

Shermis, M. D. (2015). Contrasting state-of-the-art in the machine scoring of short-form constructed responses. *Educational Assessment*, *20*(1), 46-65.

Sim, J., & Wright, C. C. (2005). The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Physical Therapy*, *85*(3), 257-268.

Smith, G. A. (2020). Motivated Reasoning and Persuading Faculty Change in Teaching. *To Improve the Academy: A Journal of Educational Development*, *39*(1).

Stains, M., Pilarz, M., & Chakraverty, D. (2015). Short and long-term impacts of the Cottrell Scholars Collaborative New Faculty Workshop. *Journal of Chemical Education, 92*(9), 1466–1476.

Stark, P., & Freishtat, R. (2014). An evaluation of course evaluations. ScienceOpen. Center for Teaching and Learning, University of California, Berkeley. Retrieved from https://www.scienceopen. com/document.

Steen, L. A. (1988). *Calculus for a New Century: A Pump, Not a Filter. Papers Presented at a Colloquium (Washington, DC, October 28-29, 1987). MAA Notes Number 8*. Mathematical Association of America.

Theobald, E. J., Hill, M. J., Tran, E., Agrawal, S., Arroyo, E. N., Behling, S., ... & Freeman, S. (2020). Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math. *Proceedings of the National Academy of Sciences*, *117*(12), 6476-6483.

Thornton, P. H., Ocasio, W. C., & Lounsbury, M. (2012). *The institutional logics perspective: A new approach to culture, structure and process*. Oxford University Press. Ch 1-2.

Tinnell, T. L., Ralston, P. A. S., Tretter, T. R., & Mills, M. E. (2019). Sustaining pedagogical change via faculty learning community. *International Journal of STEM Education*, *6*(1), 26.

Titus, A. L., Knoll, K., Sertich, J. J., Yamamura, D., Suarez, C. A., Glasspool, I. J., ... & Roberts, E. M. (2021). Geology and taphonomy of a unique tyrannosaurid bonebed from the upper Campanian Kaiparowits Formation of southern Utah: implications for tyrannosaurid gregariousness. PeerJ, 9, e11013.

UC Berkeley Hispanic Serving Institution Taskforce. (2020). *Report for UC Berkeley.*

Valenzuela, A. (2010). *Subtractive schooling: US-Mexican youth and the politics of caring*. Suny Press.
van de Sande, C. C., & Greeno, J. G. (2012). Achieving alignment of perspectival framings in problem-solving discourse. *Journal of the Learning Sciences*, 21(1), 1-44.
van Es, E. A., Hand, V., Agarwal, P., & Sandoval, C. (2022). Multidimensional noticing for equity: Theorizing mathematics teachers' systems of noticing to disrupt inequities. *Journal for Research in Mathematics Education, 53*(2), 114-132.
van Es, E. A.., Hand, V., & Mercado, J. (2017). Making visible the relationship between teachers' noticing for equity and equitable teaching practice. In *Teacher noticing: Bridging and broadening perspectives, contexts, and frameworks* (pp. 251-270). Springer, Cham.
VERBI Software. (2021). *MAXQDA 2022* [computer software]. Berlin, Germany: VERBI Software. Available from maxqda.com.
von Hoene, L., Nilson, L. B., & Siering, G. (2017). Intended teaching effectiveness outcomes for instructors and faculty development.*Institutional commitment to teaching excellence: Assessing the impacts and outcomes of faculty development*. American Council on Education., 17.
Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes (M. Cole, V. John-Steiner, S. Scribner & E. Souberman., Eds.) (A. R. Luria, M. Lopez-Morillas & M. Cole [with J. V. Wertsch], Trans.) Harvard University Press (Original manuscripts [ca. 1930-1934]), Chapter 6, 79-91. Weinstein, R. S. (2009). Reaching higher. Harvard University Press.
Wenger, E. (1998). Communities of practice: Learning as a social system. *Systems Thinker*, *9*(5), 2-3.
Weston, T. J. (2019). Patterns of switching and relocation. In *Talking about leaving revisited* (pp. 55-85). Springer, Cham.
Wheeler, L. B., & Bach, D. (2021). Understanding the impact of educational development interventions on classroom instruction and student success. *International Journal for Academic Development*, *26*(1), 24-40.
Wieman, C. (2017). *Improving how universities teach science*. Harvard University Press.
Wilson, M. (2005). *Constructing measures: an item response modeling approach*. Mahwah, NJ: Lawrence Erlbaum Assoc.
Wilson, M. (2009). Measuring progressions: Assessment structures underlying a learning progression. Journal of Research in Science Teaching: *The Official Journal of the National Association for Research in Science Teaching, 46*(6), 716-730.
Wilson, M. & Sloane, K. (2000). From principles to practice: An embedded assessment system. *Applied Measurement in Education, 13*(2), 181-208.
Winkelmes, M. A. (2014). Transparent assignment design enhances students' success.
Yoon Yoon, S., Evans, M. G., & Strobel, J. (2014). Validation of the teaching engineering self-efficacy scale for K-12 teachers: A structural equation modeling approach. *Journal of Engineering Education*, *103*(3), 463-485.

APPENDICES

APPENDIX A: AEP Practices Tip Sheet for Instructors

Adaptive Equity-Oriented Pedagogical Strategies (Phuong et al., 2017; Phuong et al., 2022)

This tip sheet provides strategies for applying AEP, a framework for adapting teaching to address equity barriers to learning based on student data (e.g., formative assessment, observations, surveys) (Phuong et al., 2017). The tips below are organized by the six competency elements, which build on each other in a mastery learning model. Some of the strategies are more relevant for instructors who design and teach courses; however, many of these strategies can broadly be applied by instructors who do not have full autonomy over their course and assessments. It is not expected that instructors apply every strategy listed below in their classroom. Instructors are encouraged to identify, adapt, and iterate the strategies that make sense for their context.

Table A

Examples of Adaptive Equity-Oriented Pedagogy (AEP) Practices

Timing	AEP Competency Element	Strategies
Before and during teaching	**(1) Clarify** learning outcomes, prerequisite knowledge, and equitable course policies	● Articulate learning outcomes in behavioral terms; be explicit about what students should know and be able to do by the end of instruction. Identify within the syllabus where the final assignment or assessment questions are taught. ● Articulate learning outcomes focused on equity goals (e.g., promoting a sense of belonging, sense of self-efficacy, identity). ● In course learning outcomes, encourage students to synthesize concepts to build a more equitable future through assignments and project-based learning. ● Generate concept maps to clarify the issues to be covered and their relationships; use the map to assess prerequisite knowledge and identify breakdowns in student learning. ● Develop community norms by asking students to reflect on and identify teaching practices that hindered and promoted their learning. ● Align students', instructors', & graders' expectations through rubrics & norming. ● Provide annotated examples of assignments of exemplary assignments with detailed comments on how they meet a rubric's expectations.

		● Provide class discussion questions or lesson plan outline before class - explain how your lessons can engage students' interests or thought-provoking topics! ● Provide strategies for effective use of time outside class and indicators as to when students should ask for help, e.g., you can include the following in an assignment: "If you're spending more than three hours on this assignment, come to office hours or send us an email. We're committed to your success!" ● Build equitable courses policies grounded in the grading for equity literature (e.g., offer flexible deadlines and extensions, allow assessment re-do's, and adjust policies based on students' needs and contextual circumstances). These policies and practices can foster a growth mindset within a brave space for learning where students can take chances, take risks, make mistakes, and grow in an environment that is focused on mastery learning (Phuong et al., 2017; Phuong et al., 2022). ● When building curricula and syllabi, center non-dominant voices and present perspectives and stories from "hidden figures". ● Diversify knowledge presented in the curriculum: expose students to diverse perspectives and outlets (e.g., journal articles, oral histories, new media, news articles, blogs, podcasts, guest speakers, etc.). ● Build course policies and practices that center universal design for learning (e.g., incorporate the ALLY tool that converts text into accessible formats and different languages, encourage students and instructors to make slide decks and course content visible, audible, and accessible for everyone in the classroom).
Before and during teaching	**(2) Align** formative assessments and teaching activities with outcomes/ summative assessments	● Implement formative assessments that are aligned with the rigor of summative assessments; use these formative assessments to guide your instruction and improvements for your course(s). I suggest using ungraded questions and/or assignments that scaffold students towards meeting learning outcomes. ● Offer assessment prompts and activities with language clarifying what instructors expect to see in students' responses. ● Model key skills and expert thinking in problem solving as well as how students can engage course content and the disciplinary field; offer cues and

		strategies (in writing) to help students analyze different question prompts, which can help students develop various skills and thought processes to apply when completing assessments and tasks. ● Provide collaborative learning and deliberate practice opportunities. ● Audit curriculum and formative/ summative assessments for bias; reflect on whether assessments questions are asking about contexts (e.g., obscure sports like cricket) that many students may not understand; examine whether students from certain subgroups are disproportionately answering an assessment incorrectly compared to other subgroups. ● Make assessments more accessible (e.g., ensure questions are screen readable, provide a bulleted form of the key points for long word problems (Toyama, 2021), offer alternative-text for images, etc.). ● Break larger assignments and projects into shorter installments that students can submit for ungraded feedback (McCallum, 2013); grade the final assignment where students can incorporate feedback and focus on mastery learning. ● Apply culturally relevant pedagogy, where students can bring themselves and their interests into the curriculum. ● Incorporate project-based, service-based, problem-based, and/or experiential learning to build more authentic assessments that engage students' interests, career goals, and lived experiences. Proactively seek ways to make these assessments culturally responsive and sustaining.
During teaching	**(3) Identify** students' competencies, interests, and needs	● Survey students' career and professional interests; make connections between course concepts and these interests. ● Administer a Student Interest Survey and Venn Diagram activity (McCallum, 2013). ● Foster a brave space where students are encouraged to seek support and express confusion as well as their learning needs. ● Use assessments to identify specific resources, activities, & page numbers from readings matching students' needs. ● Encourage groups of students to identify their strengths and interests; groups of students can create a task chart to assign roles during group projects based on

		discussion of strengths and interests; use this chart to establish point people, tasks, and deadlines to create a system of accountability; to foster motivation for early tasks, encourage students to work on tasks that engage their strengths and interests. ● Democratize the syllabus and classroom: empower students to voice how curricula can address their interests and needs through surveys and dialogue (Phuong et al., 2017).
During teaching	**(4) Understand** students' equity barriers and contextual challenges to meeting outcomes	● Survey students about their learning history and experiences. Based on capacity, include close-ended and open-ended survey questions. Use a survey tool that can transform close-ended responses into data visualizations. ● Survey students about potential equity barriers asking, "What factors outside the classroom impact your success inside the classroom? Only share what you're comfortable with." ● Initiate conversations with students about barriers they are comfortable expressing, such as imposter phenomenon. ● Use assessment/ survey questions to understand if students are having issues with basic needs access, stereotype threat, and imposter phenomenon. ● Provide opportunities for peer-to-peer and instructor-student validation to enhance psychosocial outcomes. ● Offer incentives such as game-based learning approaches to foster students' academic and professional identity.
During teaching	**(5) Adapt** teaching activities based on students' needs and barriers	● Use data to adjust methods of making expectations and thought processes explicit, such as modeling how students are expected to approach a problem or task with step-by-step instructions (Phuong et al., 2017). ● Apply Dynamic Lecturing - storytelling, reference students' interests, human polling, poll everywhere, ask students to predict answers to exciting questions that relate to the course content and that academics have trouble figuring out (Harrington & Zakrajsek, 2017). ● Adjust the difficulty of activities to provide students with appropriate levels of challenge and practice. ● Provide problems with varying levels of difficulty, so students, possibly working in self-organized groups, can find appropriate levels of challenge. Avoid

		stigmatizing groups. For example, provide level 1 (i.e., easy), level 2 (i.e., medium), and level 3 (i.e., hard) problems that students can self-organize and work on together and present back to the class — while creatively labeling the levels without calling them easy, medium, and hard. ● Apply design-based research to projects; make projects relevant to students' career or personal goals.
During and after teaching	**(6) Iterate**: Reflect upon pedagogy to support continuous learning, adaptation, and growth	● Provide feedback for students to incorporate in future class meetings and assignments that are aligned with final's rigor. ● Make student data and instructional decisions transparent. ● Provide and model strategies for how students can use their time effectively outside of class.

If you have questions, please email andrew.e.phuong@gmail.com.

I developed this tip sheet in collaboration with Judy Nguyen, Daniel Garcia, and Richard McCallum.

List of useful readings and resources:

Harrington, C., & Zakrajsek, T. (2017). Dynamic Lecturing: Research-based Strategies to Enhance Lecture Effectiveness. Stylus Publishing, LLC.

Harris, M. S., Major, C. H., & Zakrajsek, T. (2015). Teaching for learning: 101 intentionally designed educational activities to put students on the path to success. Routledge.

Kearns, K., Hatcher, M., Bollard, M., DiPietro, M., Donohue-Bergeler, D., Drane, Leslie., Luoma, E., Phuong, A. E., Thain, L., & Wright, M. (2018). "Once a scientist…:" Disciplinary approaches and intellectual dexterity in educational development. To Improve the Academy: A Journal of Educational Development. 37(1), 128-141.

Phuong, A. E., Nguyen, J., Mejia, F., Hunn, T., & Marie, D. (2021). Equitable teaching that creates pathways to success for all students. *Times Higher Education.*

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.

Wieman, C., & Gilbert, S. (2014). The teaching practices inventory: a new tool for characterizing college and university teaching in mathematics and science. *CBE-Life Sciences Education*, *13*(3), 558-559.

http://gsi.berkeley.edu/

https://teaching.berkeley.edu

https://diversity.berkeley.edu/programs-services/faculty

https://campusclimate.berkeley.edu/faculty/resources

APPENDIX B: AEPC Construct Map

I offer a brief summary of the AEPC learning progression that advances reflective and equitable teaching with supporting literature. This learning progression serves as the AEPC construct map under Wilson's (2005) four building blocks for constructing measurement instruments. The construct map is the foundation for the AEPC assessment. See the table below for the construct map. I provide the learning progression level, number of AEP elements used, a description of each level, and a summary response. More detailed responses can be found in the scoring guide.

Table B

AEPC Construct Map.

Learning Progression Level	AEP Elements Used	Description	Response
Level 0: No AEPC Mastery	0	0 AEP elements used, the instructor does not consider learner needs (McCallum, 2013). This level of competency reflects a lack of consideration of the needs and perspectives of the learner.	Provides an irrelevant answer, does not apply any elements
Level 1: Low AEPC Mastery	1-2	The instructor uses AEP elements 1-2: 1. Clarify learning outcomes and pre-requisite knowledge 2. Align formative assessments and activities (e.g., clickers, low-stakes assignments and quizzes, discussion prompts) with outcomes Fink (2013), Davis (2009), and Weinstein (2009) highlight the importance of clarifying and supporting students in meeting rigorous learning outcomes. Clarifying expectations and aligning *both* instruction and assessment with outcomes is foundational to helping students equitably navigate course curricula. The lack of alignment engenders equity issues in STEM, since it	Describes applying AEP element 1 and/or 2 in response

		often privileges students who have more background knowledge or those who have more access to financial resources to support their success. Designing for equity requires one to build and clarify learning outcomes and course policies (e.g., grading, assessment, accommodations) that are informed by the equity literature (Feldman, 2018; Phuong et al., 2017). For example, instructors can provide flexibility with deadlines (e.g., extensions), offer opportunities for re-grading, avoid curving down, and adjust policies based on students' contextual circumstances. These policies and practices can foster a growth mindset within a brave space for learning where students can take chances, take risks, make mistakes, and grow in an environment that is focused on mastery learning (Phuong et al., 2017; Phuong et al., 2022).	
Level 2: Moderate AEPC Mastery	1-4	The instructor uses AEP elements 1-4. Elements 1 and 2 mentioned above 3. Identify students' competencies, interests, and needs 4. Understand equity barriers/ contextual challenges to meeting outcomes Reaching this level requires that level 1 be satisfied as a prerequisite. This level represents moderate mastery because research shows it is beneficial to identify students' existing competencies, equity	Diagnoses learning competency and interest **Or** Diagnoses learning competency and learning barriers

		barriers (e.g., imposter syndrome, stereotype threat, no internet connection), and learning needs via formative assessments that are aligned with learning outcomes (McCallum, 2013; Phuong et al., 2017).	
Level 3: High AEPC Mastery	1-5	The instructor uses AEP elements 1-5. Elements 1-4 mentioned above 5. Adapt teaching practices based on these needs and barriers Drawing on diagnostic data aligned with course learning outcomes, the instructor adjusts instruction, which includes how they engage students in productive struggle, model key skills, provide deliberate practice opportunities, and offer feedback (Phuong et al., 2017). Level 3 represents a higher degree of AEPC because adapting instruction in a manner consistent with AEP requires the prior levels. Adapting instruction without this data-driven strategy can lead to misguided instruction and assessment practices that do not address key learning barriers (Phuong et al., 2021)	Adapts based on full diagnosis, not only based on diagnosis of learning competency
Level 4: Very High AEPC Mastery	1-6	The instructor uses AEP elements 1-6. Elements 1-5 mentioned above 6. Iterate: Reflect upon pedagogy to support continuous learning, adaptation, and growth The instructor demonstrates a commitment to adaptation,	Demonstrates applying all 6 AEP elements with supporting instructor/ student evidence and explains how their practice impacts the student **Or** Demonstrates applying all 6 AEP elements with instructor/ student evidence and explains rationale for how practice impacts student

	iteration, and growth with students, improving the application of previous elements recursively. This level suggests that the instructor can recognize breakdowns in their own teaching practices and are flexible enough to employ and adapt culturally responsive and universal design approaches (e.g., video, multimedia) to further their students' success (Hammond, 2014; CAST, 2018). In their reflections, the instructors describe rationales for improving teaching in ways that are grounded in multiple forms of student learning data. This level is the highest because in order to address breakdowns in teaching and learning, instructors need to have informed experiences associated with levels 1-3 (i.e., with AEP elements 1-5). Demonstrating a high-level of mastery suggests the instructor is willing to continuously adapt practices and policies. It is noteworthy that elevating ongoing, critical reflection (Kendi, 2019), perspective taking, and adaptation to reduce equity barriers advances anti-racist pedagogical practice.	

APPENDIX C: Reflection Questions for Qualitative Coding

I qualitatively coded questions from the post-semester AEPC assessment because these items ask instructors to describe what they have done, the evidence of demonstrating AEP competency, and their reflections on how they have adjusted their perspectives and pedagogical practices to advance student success.

Questions from Post-Semester AEPC that were qualitatively coded

Question 1. Please provide your teaching philosophy

Your teaching philosophy should include

- Your beliefs and attitudes about how students learn best and the instructors' role in supporting that learning.
- What do you think the learner's role is in the learning process?
- A description of how you define equity and why equity is important to you
- A description of your philosophy on addressing equity barriers students face
- A description of how you are an ally for all groups of students
- A description of how your teaching philosophy connects to the AEP framework

Question 2. Please provide a list of teaching practices (e.g., behaviors and language) that support your philosophy. Explain how your practices align with all the core elements of Adaptive Equity-Oriented Pedagogy (AEP).

In this response, please address **all** the following questions:

- **What are the key barriers to learning that your students have experienced? What factors do you think influence these barriers?**
- **How are you adjusting your teaching practices to help mitigate these learning barriers that result from equity issues that your students may face inside and outside the classroom?**
- **How have you been drawing on your students' interests to engage students in the course content?**

Question 3. Which of your concrete teaching strategies did you find most impactful for student success? Why? Please cite any evidence.

Question 4. Which of your concrete teaching strategies did you find least impactful for student success? Why? Please cite any evidence.

Question 6. Based on a comparison of your initial and current teaching philosophy, please provide a reflection where you respond to the statement: "I used to think, but I now think …". Before responding to this statement, please review your midterm and your teaching materials. *(Comment: Instructors were also prompted to consider: I used to do, but now do).*

Question 11. How has your understanding and use of diagnostic and ongoing assessment changed since the beginning of the semester?

Question 15. Please provide any qualitative evidence (e.g., quotations, images, videos, written descriptions from observation notes and surveys) that helped you understand your students' success.

Question 16. Please provide qualitative evidence of the impact that your teaching practices have on *one* psychosocial outcome. Psychosocial outcomes include: motivation, sense of self-efficacy, sense of belonging, stereotype threat, growth mindset, and attitudes toward learning.

The following two questions ask you to reflect on the course you're teaching or tutoring for:

Question 17. What features of the course design (e.g., syllabus schedule, teaching strategies, formative assessment, alignment of instruction and graded assessments) do you think best supported student success? Why? Please cite any evidence.

Question 18. What features of the course design (e.g., syllabus schedule, teaching strategies, formative assessment, alignment of instruction and graded assessments) do you think hindered student success? Why? Please cite any evidence.

Pre-Semester Reflection

I draw on the following items. These were not scored as part of AEPC, but provide great insight.

Why does a student not learn a concept? Whose fault is it if the student does not learn a concept?

APPENDIX D: AEPC Question Prompts

The AEPC assessment items used for analyzing instructor growth over time (data set 2) had the same item stem with unique sub-question elaborations at each time point. Please see the tables below.

Table D1
Pre-semester AEPC assessment items

Item Stems	Sub-question Elaborations	Points
Item 1: Please provide your teaching philosophy	Describe how, if at all, your teaching philosophy and practices align with the core elements of Adaptive Equity-Oriented Pedagogy	7
Item 2: Describe your clarification of learning outcomes	What are your learning goals/ outcomes? Have you set yourself goals as a teacher? What are you striving towards?	1
Item 3: What did you do? Did it work?	How do you teach or tutor? Describe a highlight. What went well? What didn't? How did your students react? How did you feel afterwards about this experience?	7
Item 4: What would you do next?	Thinking back, is there anything you would do differently to further your students' learning?	7

The same rubric was used for scoring these items that was presented in chapter 5.

Table D2
Mid-Semester AEPC assessment items

Item Stems	Sub-question Elaborations	Points
Item 1: Please provide your teaching philosophy	Write your teaching philosophy below. Your teaching philosophy should include your beliefs and attitudes about how students learn best and the instructors' role in supporting that learning. What do you think the learner's role is in the learning process? Be sure to address barriers to equity that students may experience. How is your teaching philosophy actualized in your teaching practices? Explain how your practices align with the core elements of Adaptive Equity-Oriented Pedagogy (AEP)?	7

Item 2: Describe your clarification of learning outcomes	What is the core learning outcome? In other words, what do you want students to know about this concept and/or be able to do at the end of the lesson?	1
Item 3: What did you do? Did it work?	What is a list of the teaching practices and activities you will implement to reach your learning outcome? How do these practices and activities incorporate the core elements of Adaptive Equity-Oriented Pedagogy? In particular, please describe how your teaching practices respond to your students' existing competencies, interests, and content area knowledge. How did the lesson go and what patterns of student behavior and learning did you notice? Based on data collected from your students, did your teaching practice from your lesson work? If so why? If not, why not? In your response, describe all the specific AEP teaching practice elements you applied and how students responded to these practices.	7
Item 4: What would you do next?	Based on the data you collected from your students, what would you do next to improve your students' success? What did you learn today that you did not know about your students? What else would you like to know about your students to help you further support their learning? What new or refined practices would you employ to achieve this goal and why?	7

The same rubric was used for scoring these items that was presented in chapter 5.

Table D3

Post-semester AEPC assessment items

Item Stems	Sub-question Elaborations	Points
	Your teaching philosophy should include • Your beliefs and attitudes about how students learn best and the instructors' role in supporting that learning. • What do you think the learner's role is in the learning process? • A description of how you define equity and why equity is important to you • A description of your philosophy on addressing equity barriers students face • A description of how you are an ally for all groups of students • A description of how your teaching philosophy connects to the AEP framework	
Item 1: Please provide your teaching philosophy	Please provide a list of teaching practices (e.g., behaviors and language) that support your philosophy. Explain how your practices align with all the core elements of Adaptive Equity-Oriented Pedagogy (AEP). What are the key barriers to learning that your students have experienced? What factors do you think influence these barriers? How are you adjusting your teaching practices to help mitigate these learning barriers that result from equity issues that your students may face inside and outside the classroom? How have you been drawing on your students' interests to engage students in the course content?	7
Item 2: Describe clarification of learning outcomes	How have you communicated your learning outcomes to your students?	1

	Which of your concrete teaching strategies did you find most impactful for student success? Why? Please cite any evidence. Which of your concrete teaching strategies did you find least impactful for student success? Why? Please cite any evidence.	
Item 3: What did you do? Did it work?	Please provide any qualitative evidence (e.g., quotations, images, videos, written descriptions from observation notes and surveys) that helped you understand your students' success. (Note: this question is asked after instructors provide pre- and post-assessment scores) Please provide qualitative evidence of the impact that your teaching practices have on one psychosocial outcome. Psychosocial outcomes include: motivation, sense of self-efficacy, sense of belonging, stereotype threat, growth mindset, and attitudes toward learning.	7
	Please provide your initial teaching philosophy below that you submitted at the beginning of the semester. Based on a comparison of your initial and current teaching philosophy, please provide a reflection where you respond to the statement: "I used to think, but I now think …".	
Item 4: What would you do next?	Before responding to this statement, please review your midterm and your teaching materials. If you were to become a TA, instructor, or play a course staff role in the future, how would you draw on your insights to improve teaching and assessment practices in CS, DS, or the field you teach in?	7

The same rubric was used for scoring these items that was presented in chapter 5.

APPENDIX E: Scoring Guide

This appendix presents the rubric and scoring guides for the AEPC assessment. For data set 1, the construct-based rubric with 0-7 item scores (presented in chapter 5 and reproduced below) was used to score items i1a, i8a, i10a, i11a, and i12a. For other items, I provide the rubric under each item.

As mentioned in chapter 5, I defined the outcome space for the AEPC construct by creating scoring rubrics for the items. In particular, for the items scored from 0-7, the rubric in the following table was used:

Table E1
Scoring Rubric for 0-7 Item Scores

Construct Map Level	Item Score	Description
4c: Continuous commitment to adaptation and improvement (Very High)	7	Educator describes all AEP elements with an example and explains how their practice impacts the students' success (Citing student and instructor evidence) or Describes all AEP elements with an example and explains rationale for how practice impacts students' success (Citing student and instructor evidence)
4b: Continuous commitment to adaptation and improvement (Very High)	6	Describes application of all AEP elements with an example (Citing instructor and student evidence)
4a: Continuous commitment to adaptation and improvement (Very High)	5	Describes application of all AEP elements
3: Adaptation based on diagnosis (High)	4	Adapts based on full diagnosis, not only based on diagnosis of learning competency
2b: Full Diagnosis (Moderate)	3	Diagnoses learning competency and interest (Full level 2) **Or**

		Diagnoses learning competency (gets the concept right) and learning barriers (equity barriers, lack of interest, or misconceptions) (Alternative version of Full Level 2)
2a: Partial diagnosis (Moderate)	2	Diagnoses student learning competency **Or** Diagnoses interest
1: Clarifies learning outcomes; Aligns formative assessment and instruction with learning outcomes (Low)	1	Demonstrates application of element 1 and/or 2 Element 1: Communication and clarification of learning outcomes Element 2: Alignment of formative assessment with learning outcomes
0	0	Provides an irrelevant answer, does not apply any elements

For item itc which is scored from 0-1, scorers focused on assessing whether or not instructors met APEC level 1. The following collapsed rubric below was applied:

Table E2

Scoring Rubric for 0-1 Item Scores

Construct Map Level	**Item Score**	**Description**
1: Clarifies learning outcomes; Aligns formative assessment and instruction with learning outcomes	1	Demonstrates application of element 1 and/or 2 both questions for item itc. Element 1: Communication and clarification of learning outcomes Element 2: Alignment of formative assessment with learning outcomes. For example, instructors would need to answer the question as to whether the pre- or post-assessments align with the learning outcomes reflected in graded assessments (i.e., homework, midterm, final).
0	0	Does not fully demonstrate elements 1 and 2 for both questions for item itc. For example, the participant only answers one of the questions

	for item itc correctly and does not answer both itc questions correctly. Provides an irrelevant answer, does not apply any elements

Each item has an associated rubric that was presented in chapter 5. To support consistent scoring using these rubrics, a detailed scoring guide was created for each and provided to the scorers along with sample answers that were not a part of the dataset.

The scoring guide for each item is presented below. Repetitive examples and explanations were removed for brevity.

ITEM 1/ ITEM i1a (0-7 points).

The rubric for item i1a was presented in chapter 5. Here is ITEM 1/ ITEM i1a:

Question 1a. Write your teaching philosophy below. Your teaching philosophy should include your beliefs and attitudes about how students learn best and the instructors' role in supporting that learning. What do you think the learner's role is in the learning process? Be sure to address barriers to equity that students may experience.

Question 1b. How is your teaching philosophy actualized in your teaching practices? Explain how your practices align with the core elements of Adaptive Equity-Oriented Pedagogy (AEP)?

Questions 1a-1b can be scored from level 0-7 because instructors are being asked about how their teaching practices align with AEP. Scorers would search for elements of AEP practice.

Here is a sample response at level 6. This was scored at 6 because there's little explanation of how the practices would specifically impact a student's learning.

SAMPLE RESPONSE

I believe that every student is capable of excelling (i.e., understands that students have strengths, interests, potential, and areas for growth). To achieve this goal, I need to demonstrate a growth mindset and commitment to continuously improve diverse students' success. Part of this includes interpreting ongoing feedback and assessment data as an asset for diagnosing challenges in the classroom, innovating instruction, addressing student needs, and developing a growth mindset as an instructor.

I do not believe that only the good students are the ones who teach themselves (i.e., students should learn solely on their own) (McCallum, 2013). I do not favor students with more background knowledge or those who confirm instructors' biases

I believe that students learn best when they have opportunities to become emotionally

engaged with course content that is relevant to them. They also learn effectively when course learning outcomes and expectations are made clear.

To actualize my philosophy, I communicate and clarify course learning outcomes. This entails careful consideration and effective communication of the learning outcomes and expectations for academic success in the course. An example would be to provide a syllabus with course learning outcomes and expectations.

To provide clear expectations, I align my formative assessments and course activities with course learning outcomes. This means including formative assessment and course activities that align with the rigor of homework, midterm, and final assessment questions. In other words, the formative assessment is equally as difficult as your homework, midterm, and final assessment questions.

I also diagnose my students' learning needs (i.e., their strengths, interests, and areas for growth) relative to course learning outcomes. This means diagnosing students' existing competencies, interests, and content area knowledge relative to the course learning outcomes and expectations. Diagnosis is accomplished through ongoing formative assessment, observation, and student surveys.

I then adapt my instruction based on this diagnosis based on formative assessment, observation, and student surveys. In doing so, I leverage student learning data to adjust my teaching practices, including how key skills are modeled, what deliberate practice opportunities are afforded, and the feedback provided.

Furthermore, I strive to demonstrate a continuous commitment to adaptation. I use data rather than my assumption to recognize and proactively address areas for improvement in student learning and my instructional practice continuously. Adapting continuously refers to the state of mastery where I go beyond a single iteration of adaptation based on data. In doing so, I seek to improve how I incorporate the first 5 AEP practice elements continuously. This ongoing commitment to adaptation is reflected through changes in my course design, implementation, revision to the course, the ways the outcomes are communicated, and assessment re-designs.

Applying all 6 elements of AEP practice, I use data to recognize breakdowns in my own teaching practices and I strive to be continuously adaptive enough to use a variety of responsive instructional approaches (e.g., video, multimedia, mapping). These efforts inform my plan, teach, and reflect cycle. I am passionate about highlighting patterns in student behavior and learning to respond to them.

ITEM 2 (0-1 points)

The rubric for item 2 is out of 1 point given if the instructor states learning outcomes, indicating what the student can do. Here is ITEM 2:

Question 2. What is the core learning outcome? In other words, what do you want students

to know about this concept and be able to do at the end of the lesson?

Here are sample responses for each score level.

> **SAMPLE RESPONSE**
>
> Sample response for level 0:
>
> No response, provides an irrelevant response, does not articulate course learning outcomes.
>
> Sample response for level 1:
> Here is a sample response at level 1. This was scored at 1 because they clarify learning outcomes
>
> I have already covered the definition and assumptions of a *t*-test. Therefore, I want to focus my lesson on helping my students learn how to interpret *t*-test output. By the end of my lesson, I would like my students to be able to do the following:
>
> 1. Interpret the output of a two-sample *t*-test, comparing control and treatment conditions
> 2. Compare means, standard deviations, and statistical significance of the two conditions
> 3. Explain which sample/ condition performed better
> 4. Explain how the findings may be useful for future practice

ITEM itc (0-1 points).

The rubric for item itc was presented in chapter 5. Here is ITEM itc:

Question itc. What is the pre-assessment activity that you used to determine whether or not the learning outcome has been achieved? Explain how this assessment aligns with a homework, midterm, and/or final assessment. Consider how you scored this assessment to compare it to the post-assessment. Ideally, this assessment should not take more than 10 minutes.

Question itc. What is the post-assessment activity that you used to determine whether or not the learning outcome has been achieved? Explain how this assessment aligns with a homework, midterm, and/or final assessment. Consider how you scored this assessment to compare it to the pre-assessment. Ideally, this assessment should not take more than 10 minutes.

Here is a sample response of a score of 1. For both questions below, instructors describes how they align formative assessment with learning outcomes. For example, instructors would need to

answer the question as to whether the pre- or post-assessments align with the learning outcomes reflected in graded assessments (i.e., homework, midterm, final).

SAMPLE RESPONSE

The assessment questions below are similar to a midterm's assessment questions for the course I am supporting. I will score this assessment out of 10 points and it should take about 10 minutes to complete.

Assessment

You are a researcher collaborating with faculty members trying to evaluate a new curriculum. You are analyzing statistical outputs that compare an old curriculum (i.e., the control condition) with the new curriculum (i.e., the treatment intervention).

Your tasks are listed in the prompts below and are centered around making recommendations on whether the treatment is worth implementing at a larger scale. Please use complete sentences in writing all your answers to keep your reputation as the top research collaborator in your institution.

Here is a description of the variables in the two-sample *t*-test output that you will be interpreting:

Outcome variable:

Final Exam
The final exam is a continuous variable on a scale of 0-100 points and sought to identify the extent to which students met the course objectives.

Predictor variables:

Treatment Group
The treatment group variable (*treatment*) is a binary variable indicating whether students were randomly assigned to the treatment or control condition. The value of the variable is 1 for students in the treatment course with the new curriculum and 0 for students in the control course with the old curriculum.

In light of this information, please interpret the output below responding to the prompts that are listed under the output.

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
Control	65	83.66154	1.004623	8.099531	81.65457	85.6685
Treatmen	62	97.82258	.2223905	1.751105	97.37788	98.26728
combined	127	90.5748	.819592	9.236332	88.95285	92.19675
diff		-14.16104	1.051499		-16.24209	-12.07999

diff = mean(Control) - mean(Treatmen) t = -13.4675
Ho: diff = 0 degrees of freedom = 125

Ha: diff < 0 Ha: diff != 0 Ha: diff > 0
Pr(T < t) = 0.0000 Pr(|T| > |t|) = 0.0000 Pr(T > t) = 1.0000

In your interpretation, please make sure to provide

1) the estimated mean for the treatment group and control group on the final,
2) standard error for the treatment coefficient,
3) the t-statistic,
4) degrees of freedom,
5) 95% confidence interval, and
6) *p*-value in your interpretation.

Please state which group had a higher mean. Is there a significant difference on the final between treatment and control groups at the 5% level?

Based on your findings, please explain why the treatment effect may or may not be useful for practice.

What recommendations would you have for instructors?

What recommendations would you have for researchers?

- **What is the post-assessment activity that you used to determine whether or not the learning outcome has been achieved? Explain how this assessment aligns with a homework, midterm, and/or final assessment. Consider how you scored this assessment to compare it to the pre-assessment. Ideally, this assessment should not take more than 10 minutes. This post-assessment should be given near the end of your lesson to see if you met your lesson plan learning outcomes.**

Same as above
This assessment is similar to a midterm assessment.

ITEM i8a (0-7 points).

The rubric for item i8a was presented in chapter 5. Here is ITEM i8a:

Question 8. What is a list of the teaching practices and activities you will implement to reach your learning outcome?

Question 8. How do these practices and activities incorporate the core elements of Adaptive Equity-Oriented Pedagogy? In particular, please describe how your teaching practices respond to your students' existing competencies, interests, and content area knowledge.

Here is a sample response of 6 because it does not explain how it will impact their specific student. A 7 could be satisfied with a more detailed description of how the practices impact students in the response.

SAMPLE RESPONSE

I provide my teaching practices below with timing that runs for about an hour.

1. I will **explain and clarify** my learning outcomes (2 minutes)
2. I will use the **pre-assessment** to diagnose what my students understand and do not understand. My goal is to understand how the student is making sense of the course material. (10 minutes)
3. I will draw **on student-interest survey data** to explain how this skill is related to my students' professional goals (2 minutes). For example, I will explain how the competencies from this lesson are in job descriptions for becoming a data scientist.
4. Using **another practice question** of similar rigor to the pre-assessment, I will model and explain how to read the assessment questions (5 minutes)
5. I will then **show an example with steps o**f how to interpret the output and respond to the assessment question (5 minutes)
6. I will then **ask my students if they have any questions** and address misunderstanding/ misconceptions. (5 minutes)
7. I will provide an opportunity for stu**dents to practice a different problem** that is similar to the rigor of the pre-assessment, using the steps and example I provided. (10 minutes)
8. I will provide **some feedback and** re-teach any parts they didn't understand in a different way. (5 minutes)
9. I will administer the **post-assessment** and see if the student understands the material (10 minutes).
10. I will **debrief** for 2-3 minutes if time permits. Otherwise, I will take time to examine their response and debrief in our next lesson.

Based on the practices listed above, I applied all 6 AEP practice elements:

I satisfied by communicating and clarifying expectations.

I aligned formative assessment questions with the midterm questions.

I diagnosed students' learning relative to my course learning outco6me/ midterm questions. I addressed students' existing competencies, content area knowledge, and how their interests (i.e., job interests) tie to the content.

I adapted based on diagnosis by building on what they did know, indicating how the lesson is relevant to their interests, and showing students how to solve the problem.

I showed and will continue to demonstrate a continuous commitment to adaptation by teaching materials in new ways.

ITEM i10a (0-7 points).

The rubric for item i10a was presented in chapter 5. Here is ITEM i10a:

Question 10. How did the lesson go and what patterns of student behavior and learning did you notice? Based on data collected from your students, did your teaching practice from your lesson work? If so why? If not, why not? In your response, describe all the specific AEP teaching practice elements you applied and how students responded to these practices.

Here is a sample response that would be scored as 7. A good 7 has an explanation and a rationale for how the instructor would iterate teaching and directly impact the student success.

SAMPLE RESPONSE

Example response: My post-test was better than my pre test and I noticed that students were more engaged in the collaborative activities. This behavior was consistent whenever I used collaborative activities. These tests were on a scale of 0-10. The average pre-test score was about 3 points with a mode of 3 and the average post-test score was about 9 points with a mode of 9. The range for the pre-test was from 0-5. The range for the post-test was 8-10.5

I calculated the average gain score by computing the following: average gain score = average post-test score - pre-test score.

The average gain score was about 6 points, since 9-3 = 6. The range was also smaller on the post-test. Therefore, I think the lesson worked because there were improvements in student learning.

Since these tests were **aligned** with the rigor of the midterm, I was able to work with my student to have a better understanding of the expectations so that they could focus on meeting course objectives. I **communicated the lesson plan's learning outcomes** to the student by contextualizing the pre- and post-test questions with the course. This helped my student learn more effectively because my student was able to reflect on their learning and ask more meaningful questions, which helped me address misconceptions.

Consistent with my philosophy on adaptive equity-oriented instruction, this approach probably worked because we used formative assessment and student interest surveys to **diagnose** students' existing competencies and interests.

I used these data to adapt my instruction to increase student engagement with developing academic skills. These data were helpful because I was able to foreground why the course concept was relevant to the students' professional goals as I modeled key skills, provided practice opportunities, and gave feedback.

I think these practices worked because I was able to address what students did not know based on the pre-test and was able to provide teaching practices and examples that engaged their interests and areas for growth. Consistent with the AEP framework, the cycles of ongoing formative feedback and practice opportunities enabled me to address misconceptions and reinforce concepts through multiple modalities (e.g., visual, auditory, kinesthetic). **I demonstrated a commitment to continuous adaptation and improvement because I re-taught the concept in a visual way after noticing that my student was still confused with the other way I taught the concept.**

ITEM i11a (0-7 points).

The rubric for item i11a was presented in chapter 5. Here is ITEM i11a:

Question 11. Based on the data you collected from your students, what would you do next to improve your students' success?

Here are sample responses for each score based on the rubric.

SAMPLE RESPONSE

Score 0

No response or irrelevant response

Score 1
N/A

Score 2
I diagnosed that my student doesn't understand recursion but not do anything different

Score 3
I diagnosed that my student doesn't understand recursion and that they feel imposter syndrome but I'm not going to do anything different.

Score 4 if the instructor said they did some sort of adaptation, but doesn't want to

continuously adapt

My student did really well based on my teaching adaptations. I am not going to change anything else from what I have done, especially if I have applied the other elements.

(Comment: This response does not show a continuous commitment to adaptation and is therefore not level 4 (i.e., score 5 or higher)).

Score 5.
I want to demonstrate a continuous commitment to adaptation and improvement and/or the other AEP elements.

Score of 6
I want to demonstrate a continuous commitment with an example of how I will adapt

Example response for *t*-test lesson: They did well, I am going to provide more challenging deliberate practice activities to help further my students' learning.

Score of 7
I want to demonstrate a continuous commitment with an example of how I will adapt and I explain why this is going to improve student success

Example response 2 for Computer Science: Based on the post-assessment score, the student still did not have a clear understanding of recursion. Therefore, I will try to teach recursion in a more visual way by mapping out the thinking processes and writing this process out for my student. For example, I will use environment diagrams to visually represent recursion. I am doing this because I observed that the student learned better and was more engaged when I presented information visually.

ITEM i12a (0-7 points).

The rubric for item i12a was presented in chapter 5. Here is ITEM i12a:

Question 12. What did you learn today that you did not know about your students? What else would you like to know about your students to help you further support their learning? What new or refined practices would you employ to achieve this goal and why?

Here are sample responses for each score based on the rubric.

SAMPLE RESPONSE

Score 0

No response or irrelevant response

Score 1

N/A

Score 2
I learned something from my student, but I have no desire to learn anything or do anything new

Score 3

I learned something from my students' interests, equity barriers, and learning competencies, and I'm not sure what I would do (OR ACTION PLAN is not concrete (e.g., I'll research something), but there is a full diagnosis)

Score 4 if the instructor said they did some sort of adaptation, but doesn't want to continuously adapt

I learned a lot about my student based on my teaching adaptations. I am not going to change anything else from what I have done, especially if I have applied the other elements.

(Comment: This response does not show a continuous commitment to adaptation and is therefore not level 4 (i.e., score 5 or higher)).

Score 5.
I want to demonstrate a continuous commitment to adaptation and improvement and the other AEP elements to learn more about my student.

Score of 6

Example response for *t*-test lesson: I want to learn more about how my student responds to more challenging deliberate practice exercises. I am going to provide more challenging deliberate practice activities to help further my students' learning.

Score 7

> Sample response: I learned my student was more engaged when I mapped concepts. I want to learn more about what teaching practices my student found to be effective and less effective for supporting their learning. I will implement an anonymous course feedback form to ask about how specific teaching practices impacted their learning. I also want to ask them ways I could engage them more in this form.
>
> I noticed that some students did not write much on the pre-assessment. I would also like to ask them more about their thought processes next time in my course-feedback form. This would help me learn more about my student's thought

process so that I can further their success and identify key misconceptions to address in the future.

As described in chapter 6, these rubrics and rubrics were adapted for dataset 2.

APPENDIX F: Codebook

Table F1
Authority and Community Lens super codes and subcodes

Super code: Authority & Community Lens

Definition: Instructors mimic and value practices from previous instructors, their department, or academic field. "I am following or am heavily influenced by someone's pedagogical model."

I hypothesize that higher-growth instructors will mimic equity-oriented instructors and norms.

Name of subcode	Label of subcode	Description
I mimic instructors whose style was effective for me	*AnC-EFF*	Instructors are motivated to adopt teaching practices that worked for them
I mimic esteemed equity-oriented instructors	*AnC-EOI*	Instructors are motivated to adopt practices used by other esteemed instructors who are equity-oriented (e.g., breaks assignments into installments with formative feedback, uses data to adjust teaching) Inspired by instructors who "grade for equity"-- e.g., believe in clobber policy/ assignment re-dos
Inspired by instructors who demand that students work hard to meet high expectations	*AnC-EXPEC*	Demands that students work hard and provides very difficult exams that require students to learn more, curving down and limiting "A's" is good Give a 3-hour exam with 2 hours to take it.
They value this teaching practice in my discipline	*AnC-DISP*	Instructors are motivated to adopt practices that are aligned with disciplinary norms/expectations.
They value this teaching practice in my department	*AnC-DEPT*	Instructors are motivated to adopt practices that are aligned with departmental norms/expectations.
I independently found a source from an authoritative	*AnC-AUTH*	Instructor finds a source outside of the university to inspire their teaching

figure that I use for my teaching		
Something in the department constrains my teaching effectiveness	*AnC-CONS*	The course curriculum is overpacked and moves too fast. Therefore, it's hard to adapt.

Table F2
Data Lens super code and subcodes

Super code: Data Lens

Describes how valued forms of data and analytical methods affect pedagogical practice

The Data Lens is defined as the kinds of knowledge and data that are important and relevant in an academic and socio-organizational culture. Here is a question that one may consider when reflecting on this lens: Do I value certain kinds of practices because they are tested using data sources and methods that are important to me or my field?

An instructor can use the Data Lens to understand student perspectives and learning processes. This lens can be useful for diagnosing student needs and adapting teaching.

The subcodes below impact motivation to learn and adopt AEP.

Name of subcode	Label of subcode	Description
I value quantitative data	*DATA-QUANT*	Instructors rely on quantitative data and/or its validity to adjust teaching. I value reliable quantitative data
I value qualitative data	*DATA-QUAL*	Instructor relies on qualitative data and/or its validity to adjust teaching I value reliable qualitative data
It depends	*DATA-DEP*	Instructor prioritize certain types of data based on the instructional context *Conditional (If/ then) or "it depends on the context")
Formative assessment is important	*DATA-FAI*	I trust and rely on formative assessment to adjust teaching. I value diagnostic data.

I believe it is important to use data to measure progress towards learning outcomes	*DATA-MLO*	I assess content knowledge and level of academic preparation. I adjust teaching based on that assessment.
Data about socio historical context and student lived experience is important	*DATA-SLE*	When adjusting teaching, I identify equity barriers and/or the students' socio historical experience, interests, and context
Multiple forms of data are used to make pedagogical decisions that further student success	*DATA-MULT*	The instructor explicitly uses multiple forms of data to inform pedagogical decisions that further student success

Table F3

Content Lens super code and subcodes

Super code: Learning Goal Lens

The Learning Goal Lens is defined as the concepts or ideas that one foregrounds in their mind. Here are questions that one may consider when reflecting on this lens: Is the AEP practice or knowledge relevant to my teaching, goals, and priorities that I am foregrounding in my mind? Are these AEP practices relevant to the course content and concepts that my organization and I prioritize?

The subcodes below impact motivation to learn and adopt AEP.

Name of subcode	Label of subcode	Description
Academic Lens	*LG-ACA*	Goal is to teach academic content better. I focus on academic learning outcomes.
Equity Lens	*LG-EQU*	Goal is to promote an equity goal for learning: sense of belonging, community, making students feel comfortable or engaged, social interaction, student voice and agency, bringing students into curricula, psychological well-being, etc. I value student input to adjust what concepts I prioritize teaching and the way I teach

Table F4

Emotional Lens super code and subcodes

Super code: Emotional Lens

The Emotional Lens is defined as the emotions that arise from engaging in a practice, experience, or reflection. I contend that the norms of a socio-organizational culture impact how individuals think they should feel and how they express their feelings.

The Emotional Lens can explain motivations for adapting teaching.

The subcodes below impact motivation to learn and adopt AEP.

Name of subcode	Label of subcode	Description
Positive	*EMO-P*	Instructor feels positive emotions related to teaching and learning
Negative	*EMO-N*	Instructor feels negative emotions related to teaching and learning
Emotions are tied to how students DID in my class (measured via their performance)	*EMO-DID*	Student performance outcomes motivate or demotivate instructors to adapt
Emotions are tied to how students FELT in my class (measured via student feedback)	*EMO-FELT*	Positive feedback creates positive emotions that address burnout and promote continuous pedagogical adaptation Negative feedback creates negative emotions that hinder adaptation. My students don't appreciate me so I'm not going to try anymore.
Emotions tied to personal learning experiences	*EMO-PLO*	Instructors' empowering and/or disempowering learning experiences motivate them to adapt
Emotions related to what I did and how I felt during class	*EMO-I*	Instructor feels a certain way based on what they did in class.
Emotions related to teaching online	*EMO-ONLINE*	Instructors may experience positive or negative emotions related to teaching online. These emotions can motivate or demotivate one to learn AEP.

Table F5
Responsibility for Learning Lens super code and subcodes

Super code: Responsibility of Learning Lens

Their model of how students should learn can motivate instructors to adopt or not adopt AEP.

Name of subcode	Label of subcode	Description
It's up to me	*RESP-ME*	Instructor responses that suggest instructors themselves should take responsibility to mitigate the impact of these barriers on student success; The instructor believes in their responsibility to proactively understand their students' questions, barriers, needs, interests, and experiences. They see an important part of their role to be proactively addressing student learning needs and equity barriers.
It's on my student	*RESP-STUD*	Instructor responses that suggest the student should play a proactive role in advancing their own learning. Example: It's on my student, and I'm not willing to adapt because they need to own the learning and work harder.
It's something else	*RESP-ELSE*	Instructor responses that indicate another contextual factor (e.g., equity barrier, stress on the community or family) plays a role in impacting student learning.

Table F6
Equity Definitions and Perspectives Lens super code and subcodes

Supercode: Equity Definitions and Perspectives Lens

Definitions of equity can motivate instructors to adopt or not adopt AEP.

Name of subcode	Label of subcode	Description
Equity is giving everyone the same thing (i.e., equality)	*EQUITY-SAME*	Instructor defines equity as giving everyone the same thing
Equity is adapting to	*EQUITY-ADAPT*	Instructor defines equity as adapting to address

address student needs and their socio historical context -- I expect that students have different pathways to learning and success		student needs and their socio historical context. The goal is to create greater opportunities and outcomes for success by reducing equity barriers. Instructors understand that students have different learning needs and are responsive and supportive of different pathways to learning
I expect students to behave, learn, and succeed in the same ways or in the way I learned	*EQUITY-SBEHAVE*	Instructors provide a one-size-fits-all model of learning It possible the instructor is basing this on how they were socialized to learn-- i.e., the good students are the ones who teach themselves
I want all my students to succeed/ I address equity issues	*EQUITY-WSS*	I adapt my teaching to support all students
I don't adapt to support all students-- e.g., I teach only to one group of students (high, middle, or low performing students)	*EQUITY-NOADAPT*	I teach only to one group of students (high, middle, or low performing students)

Below, I provide two examples of two instructors' teaching philosophy and practice statements from the AEPC post-assessment as coded using the book above. All quoted written reflection excerpts below are entire responses to a reflection unless it's noted. All subcodes are identified in the written excerpts below with brackets and a label, which are described below. One statement can be coded for multiple super codes (e.g., Authority and Community Lens). The subcodes (e.g., *AnC-EFF, AnC-EOI)* can have some overlap and can contain one another.

Example Coding I: Teaching Philosophy and Practice Statement Part 1

> I believe students learn best when they are interested in the material and feel excited about engaging with it. Instructors' should strive to create a safe and comfortable environment where students feel heard and respected. We must be honest, intentional, and adaptable in our lessons. [I think it's the teacher's responsibility to get to know their students, including their unique background, learning needs, and passions. (*DATA-SLE; RESP-ME*)] Teachers should do their best to ensure that the material they're teaching is tailored to the level of the students, without being too easy or unnecessarily challenging. [Teachers should plan lessons carefully, considering their students' zone of proximal development and how best they might scaffold them towards desired student learning outcomes. I believe it's also important to tie course

material to other topics students are interested in outside the classroom, helping to spark passion and increase engagement with the material. (LG-*ACA*)].

[The learner's role in this learning process is equally important. Students should be open-minded and be ready to put in the time and effort they might need to practice and master the material. They should keep a positive mindset, knowing that if they keep trying, they will succeed. (*RESP-STUD)].* It is also important that learners seek out support when they need it. Teachers should provide an environment and resources that benefit the students' learning, but the students must reach out and take advantage of these resources.

It's important that students feel comfortable to speak up, asking questions and discussing ideas with others. It is in the best interest of the students for them to be their own advocate – ensuring that they have everything they need to succeed.

Equity in the context of learning and teaching is this idea that all students have access to whatever resources they need for success. Some students may need more support, while others may need less. Every student is different and has their own set of unique skills, knowledge, and challenges – facing many forms of equity barriers. It's important that each student is looked at as an individual and can get support for their individual needs. [My philosophy on addressing equity barriers is that I should be as supportive and approachable as possible. (LG-*EQU)]* [I want to be able to know my students on a personal level so that I can be aware of any inequities they face and be able to support them academically through them all. (*EQUITY-ADAPT)*] To me, without equity we risk the chance of missing out on so much greatness. If students aren't given the support they each need to succeed, we as educators are failing them. We aren't allowing them to reach their full potential. Everyone has their own special way that they can positively impact the world and each person has their own path to getting there. As teachers, to do the best we can for the world, we must be an ally to all students, ensuring that none of their paths to success are blocked, by actively doing all we can to support our students through their learning process.

This whole philosophy relates back to the Adaptive Equity-Oriented Pedagogy (AEP) framework, which focuses on five key practices that teachers should engage in to help provide equitable instruction. Focusing on the students' individual needs and continually adapting to those needs is critical. [Teachers must prioritize feedback and assessments so that they are effectively able to identify any problem areas and continue to grow and improve for their students. As teachers, our most important responsibility is to provide all our students with the best opportunity for learning that we can. (*DATA-QUANT)]*

Example Coding II: Teaching Philosophy and Practice Statement Part 2

My teaching practices:

· **Positive language and excitement about the material**

o I feel that as a teacher, one of my primary goals is to empower my students to want to learn and to push themselves. I want to inspire students to care about the material and to embrace any challenges, instead of fearing them. I think being authentic and honest with my students is an important aspect of my teaching. I want my students to feel comfortable asking questions and being honest with me about their learning. I never want them to be afraid to share their thoughts. Every student has a unique background and a unique way of learning and working with material, and I think it's crucial to recognize and adapt my teaching style to this. I don't think that with teaching there is a simple 'one- size-fits-all' approach. I think it often takes working closely with your students and finding ways to assess their progress and adapt to their needs.

· **Clear communication**

o [I think it's very important to clearly communicate what the learning outcomes of your lessons are. I try to start each of my lessons like this as I feel that this is the first road map students get when introducing new material. This gives students something to look forward to as well as a bit of context for what they're about to learn. (LG-*ACA)]*

· **Perform assessments**

o I feel that it's crucial as an instructor to perform formative assessments which align with those learning outcomes. Not only are assessments important to understand where your students current learning is at, but it's also vital to provide assessments that align with the goals of the course and your lesson. [[Questions that are too far removed from the material at hand can be confusing and discouraging for students, so it's important that all assessments are relevant and are something the students can really engage with. Formative assessments are a great tool to help diagnose and evaluate student learning. (*DATA-QUANT; DATA-FAI; DATA-MLO)]*. [I like to couple them with anonymous feedback surveys which not only assess how the students is feeling about their progress in the course, but also give me a chance to learn more about each individual students' interests and anything else that might be going on in their lives. (*DATA-QUAL; DATA-SLE; DATA-MULT)]*

· **Adapt, adapt, adapt**

o The assessments and feedback surveys I do help me make more informed decisions about my teaching style and lesson plans and can help notify me if a student might need extra support. With this information, I am able to adapt my lesson to fit the specific needs of the students I'm working with. As I'm able to work with students more, this only becomes better as I'm able to continuously adapt to my students as I learn more about them. I constantly strive to take in feedback so that I can keep improving my lessons.

Some of the barriers to learning that my students have experienced are not being prepared enough for the content of the course. Most of my students did not have very much, if any, coding experience before CS 61A, and so when entering CS 61B (the course I teach), they are woefully underprepared, especially when compared to many of their fellow classmates. I think this generally is due to a lack of CS education at many high schools. There is also some pressure to take CS 61A as a first class, even

though it is not a good introduction to computer science course. The pace and content of CS 61A and CS 61B often leave bright students feeling discouraged simply because they haven't had enough preparation for these classes.

[A majority of the students I've taught this semester have also been women, and some have felt uncomfortable and even looked down upon for being a woman in a course and major dominated by men. This adds an extra layer of self-doubt and insecurity that only negatively impacts their learning. (EMO-N)] I think being intentional about treating students equitably and with equal respect regardless of their gender, race, religion, sexual orientation, etc., is crucial to overcoming these types of learning barriers. (*RESP-ELSE)]*

To help mitigate these learning barriers, I try to adapt my teaching practices as much as possible to the individual students I work with. One of the main adaptations I make is with the pace of my lessons. I try to always ensure that the student I'm teaching is understanding the material by checking in with them frequently during the lesson, ensuring that students have an opportunity to speak up about any confusion and clear up misunderstandings. I think this is especially important for students who may have less preparation coming into the course and may feel overwhelmed or discouraged by the course material.

I try to always start my sections off by asking my students how they're doing, in life and in the course. I really do care about each and every one of them and I try to be mindful of any extra difficulties they are having. [I always try to encourage my students to stay positive and to have a growth mindset (*EMO-P*).] If I can, I like to tell them about the struggles that I faced when I was in their position and the things that helped me overcome those struggles.

I try to always relate material back to the "big picture" of the course, and I often talk about how students might use the material in the 'real world' or in later courses. [When I can, I try to relate examples or concepts back to my students' majors or outside areas of interests . I'm a
very visual learner and I know many of my students are as well, so I also try to present the material in multiple modes whenever possible, often seeking out cool graphics or videos for course concepts that are more interesting ways of visualizing the material. *(EQUITY-ADAPT)*]