

# Predictive Validity of the Algebra Concept Inventory (ACI): The Importance of Algebraic Conceptual Understanding to College Math and STEM Outcomes

Claire Wladis 

*Mathematics Department, Borough of Manhattan Community College  
at the City University of New York*

*Urban Education Department, The Graduate Center at the City University of New York*

Benjamin Sencindiver 

*Mathematics Department, University of Texas at San Antonio*

Alyse C. Hachey 

*Department of Teacher Education, University of Texas at El Paso*

Nils Myszkowski 

*Department of Psychology, Pace University*

Kathleen Offenholley

*Mathematics Department, Borough of Manhattan Community College at the City University of New York*

Jason Samuels 

*Mathematics Department, Borough of Manhattan Community College  
at the City University of New York*

Geillan Aly

*Urban Education Department, Graduate Center at the City University of New York*

The Algebra Concept Inventory (ACI) is the first large-scale instrument validated to measure the foundational algebraic conceptual understanding of college students. This study uses ACI scores to conduct the first quantitative analysis on the relationship between algebraic conceptual understanding and college outcomes, thus exploring the predictive validity of the ACI. Specifically, we investigate whether ACI scores predict: (1) math course grades; (2) subsequent completion of math courses required for STEM majors; (3) completion of STEM versus non-STEM degrees; and (4) the extent that differences in these outcomes by race/ethnicity or gender are explained by ACI scores. Results indicate that ACI scores significantly predict math course outcomes and STEM versus non-STEM degree completion, as well as significant proportions of differences in these outcomes by race/ethnicity and gender. This illustrates the importance of providing every student instruction that supports development of the kinds of foundational algebraic conceptual understanding measured by the ACI.

Keywords: *higher education, colleges, assessment, equity, mathematics education, hierarchical linear modeling, regression analyses, research in undergraduate mathematics education, concept inventories, algebra, conceptual understanding, predictive validity, STEM majors*



HIGHER education researchers have identified algebra as a barrier to college and STEM degrees (e.g., Adelman, 2006; Bailey et al., 2010; Cohen & Kelly, 2019), and researchers in undergraduate mathematics education have documented how algebra knowledge can be critical to higher-level college courses, such as calculus (e.g., Frank & Thompson, 2021; Stewart & Reeder, 2017; Stewart et al., 2018), which are often a required part of STEM degree sequences. Yet, despite the acknowledged importance of algebra to college outcomes, studies investigating college math and STEM outcomes have not generally measured algebra knowledge directly, often relying on imperfect proxies that assess a much wider range of skills or knowledge. Most existing research has used successful completion of various secondary or postsecondary mathematics courses as a proxy for algebra skills and knowledge (e.g., Gipson, 2016; Maltese, 2008; Nicholls et al., 2013). Some studies have measured mathematical skills and knowledge more directly by using course grades or SAT/ACT test scores (e.g., Alkhasawneh & Hargraves, 2014; LeBeau et al., 2012; Sahin et al., 2012; Wolniak, 2016). However, grades “assess a multidimensional construct containing both cognitive and non-cognitive factors” (Brookhart et al., 2016, p. 803), and tests like the SAT/ACT combine a wide array of broader skills and knowledge. Thus, these measures were designed for a different purpose, and are not designed to assess cognitive measures of specific types of algebraic thinking.

Without a better understanding of which types of algebra knowledge are critical to college math and STEM outcomes, it is difficult to design college curricula that better support students in developing this knowledge. One type of mathematics knowledge that has been stressed as important (yet is often overlooked in instruction) is conceptual understanding (e.g., Aly, 2022; Boyce & O’Halloran, 2020; National Research Council, 2001; Richland et al., 2012; Weibel et al., 2017). Yet, algebra courses in college tend to focus on procedural practice and memorization disconnected from conceptual understanding (Crooks & Alibali, 2014; Hammerman & Goldberg, 2003; Hodara, 2011; Rittle-Johnson & Schneider, 2014) and research has found that most K–12 students graduate without flexible, conceptual mathematics knowledge (Richland et al., 2012). This may leave students unprepared to apply algebraic knowledge flexibly in novel contexts related to their STEM field, which is critical to success in the major and in the STEM workforce (Quarles & Davis, 2017).

Access to the kinds of algebra instruction that support development of conceptual understanding is also an equity issue. Racial/ethnic minorities, women, and students with disabilities are more likely to be placed into non-credit algebra classes in college (e.g., Chen & Simone, 2016; Hodara, 2019; Sanabria et al., 2020) that tend to rely more on rote procedural instruction (Crooks & Alibali, 2014; Hammerman & Goldberg, 2003; Hodara, 2011; Rittle-Johnson &

Schneider, 2014). This may exacerbate existing K–12 inequities, where more marginalized students have been shown to have less access to the kinds of rich instruction that provide opportunities for sensemaking and other higher-level skills needed in college (Aly, 2022; Schoenfeld, 2022; Stepter, 2023; Yeh et al., 2020). This may be one reason why students from groups that have been traditionally underrepresented and underserved in STEM are more likely to change their mind about majoring in STEM in college (e.g., Black et al., 2021; Hatfield et al., 2022; Riegle-Crumb et al., 2019; Wright et al., 2023; Zhang, 2021).

In this study, we explore the extent to which algebraic conceptual understanding predicts college math course and STEM degree outcomes in college. The Algebra Concept Inventory (ACI) is the first assessment of its kind to undergo large-scale validation for measuring the foundational algebraic conceptual understanding of college students (Wladis et al., 2024a, 2024b; Wladis, Murray, & Aly, 2025b; Wladis, Murray, Hachey, et al., 2025c; Wladis et al., manuscript under review). Here, we analyze whether ACI scores predict students’ subsequent college math course grades, completion of core math requirements for STEM majors, and completion of STEM versus non-STEM degrees. We also investigate whether ACI scores explain differences in these outcomes by race/ethnicity and gender.

Performing these analyses allowed us to explore how supporting and measuring the development of algebraic conceptual understanding may be important to college math and STEM outcomes, which goes beyond just enrollment in or completion of particular math courses in high school or college. This is an area that has often been hypothesized as important (e.g., Al-Mutawah et al., 2019; Crooks & Alibali, 2014; National Research Council, 2001; Richland et al., 2012), but rarely explored quantitatively in empirical studies. If ACI scores predict math and STEM outcomes in college, then this suggests that it may be particularly important for mathematics educators and researchers to attend to the ways that algebraic conceptual understanding is taught throughout the K–16 curriculum, and to consider whether all students have access to instruction that supports the development of this understanding. This may be one malleable factor that could help to improve enrollment in and completion of STEM degrees, particularly for groups that have been traditionally underrepresented and underserved in STEM fields.

## Literature Review

### *The Leaky STEM Pipeline*

Studies have long explored various factors related to enrollment, persistence, and STEM degree completion in college. Data show that students who initially indicate an interest in STEM often leave college with non-STEM degrees (e.g., van den Hurk et al., 2018; Zhang, 2021). Further, students from traditionally underrepresented groups

are less likely to both enroll in and complete STEM degrees (e.g., Black et al., 2021; Hatfield et al., 2022; Wright et al., 2023; Zhang, 2021); this persistent discrepancy sets STEM fields apart from non-STEM fields (Riegler-Crumb et al., 2019). The loss of students from STEM at various points in their educational trajectory is commonly referred to as the “Leaky STEM Pipeline,” and it represents a significant potential loss of talent for STEM professions (e.g., van den Hurk et al., 2018). Much of the research on the Leaky STEM Pipeline has focused on equal access concerns, as certain groups are more highly represented in STEM degrees and the STEM workforce than others. For example, Black and Hispanic students, low-income students, women, and students with disabilities are all more likely to switch out of STEM degrees, leading to underrepresentation (e.g., Wang, 2013; Zhang, 2021) and a lack of diversity that is critical to innovation (e.g., Freeman & Huang, 2014).

Research has described structural barriers in K–12 education that impact pathways into STEM for low income, rural, and students of color, such as lack of access to equitable school funding, science resources, high-quality teachers, and technology access (Charleston et al., 2014; Reardon et al., 2021; Scott & Martin, 2014). Studies consistently suggest that these underrepresented students have less access to pre-college experiences that prepare them for college STEM majors, including advanced math and science coursework (e.g., Office for Civil Rights, US Department of Education, 2023; Schoenfeld, 2022). Both prior high school preparation and socioeconomic status have been found to be a factor in STEM major choice and persistence (Anderson & Kim, 2006; Chen & Weko, 2009; Ellington, 2006; Herrera & Hurtado, 2011; Tyson et al., 2007; Wang, 2013; Zhang, 2021). Mathematics placement is also a contributing factor, as research indicates that students of color and lower socioeconomic status students are disproportionately overrepresented in developmental algebra courses in college (e.g., Hodara, 2019), where passing rates are often below 50% (e.g., Coltharp, 2020); this serves to delay/derail STEM degree attainment.

#### *Connection Between Algebra Learning and STEM Degree Attainment*

Algebra has been studied as a predictive factor in determining whether students enroll in or complete STEM degrees. Substantial research has explored whether the timing or amount of particular math courses taken during K–12 schooling predicts STEM degree enrollment or completion in college (e.g., Anderson & Kim, 2006; Chen & Weko, 2009; Ellington, 2006; Redmond-Sanogo et al., 2016; Tyson et al., 2007; Wang, 2013). For example, Tyson et al. (2007) report that Black and Hispanic students on average complete lower levels of math courses in high school than other groups, but that Black and Hispanic students who take higher

level math courses are as likely as White students to pursue STEM degrees. This may be related to the fact that Black and Hispanic students have less access to higher-level math and science courses in school compared to White and Asian students (Civil Rights Data Collection, 2023).

Using a nationally representative dataset, Chen and Weko (2009) found that taking higher level math courses in high school was associated with STEM pathway entrance in college. Wang (2013), also using a national representative dataset, found that the effect of high school exposure to math and science on STEM pursuit in college was statistically significant and positive across all racial groups, with math achievement in 12th grade positively associated with intent to pursue STEM fields. Further, (Minaya, 2021) reports findings that dual credit enrollment in high school algebra (where students receive both high school and college credit for the course) increases the likelihood of selecting a STEM major in college; this study found a particularly strong relationship with beginning college as a STEM major and persisting in the major for Black and Hispanic students. Cohen and Kelly (2019) report that students who take developmental algebra in college (which is often conceptualized as a proxy for non-taking or non-success in algebra in high school<sup>1</sup>) were more likely to change to non-STEM majors than those who were labeled by the institution as prepared for “college-level” math at the start of postsecondary education. This is in line with Crisp et al. (2009), who found that students who enrolled for credit in “non-developmental” algebra (or higher) during the first semester of college were more likely to persist toward a STEM degree.

However, all these studies focus on students’ course-taking, rather than on measures of mathematical knowledge that students may have acquired as a result of these courses. This can be problematic because which courses students take is the result of many complex confounding structural factors (e.g., availability of course offerings in school, academic advisement, college placement procedures, etc.), which may be unrelated to students’ mathematical knowledge. Some research has attempted to address this by exploring the relationship between grades or test scores and college outcomes. Studies have demonstrated a correlation between high school math course grades or overall GPA and college STEM outcomes (e.g., Gipson, 2016; Maltese, 2008; Nicholls et al., 2013). Other research has analyzed the correlation between scores on certain standardized assessments such as the SAT or ACT and subsequent STEM enrollment and degree completion in college: students with higher scores on the mathematics portions of these tests are more likely to enroll in and persist in STEM majors (Alkhasawneh & Hargraves, 2014; LeBeau et al., 2012; Sahin et al., 2012; Wolniak, 2016). Additionally, research has noted a correlation between advance placement (AP) exam scores (on the Calculus and other science AP exams) and STEM enrollment/retention (Ackerman et al., 2013; see review in Patel, 2024). These

assessments all measure different types of mathematical knowledge that are either more general than or distinct from algebraic conceptual understanding. While these assessments may include some items that measure algebraic conceptual understanding, they do not measure it as a separate construct; for example, many items on these assessments focus on computation or other mathematical topics beyond algebra. Thus, no existing large-scale research has explored the relationship specifically between algebraic conceptual understanding and college math/STEM outcomes.

### *The Importance of Algebraic Conceptual Understanding*

Conceptual understanding is considered to be a critical component of mathematical knowledge (e.g., Al-Mutawah et al., 2019; National Research Council, 2001; Richland et al., 2012), and algebra is considered foundational to much of math and science (Juraev & Bozorov, 2024). However, research suggests that existing algebra instruction has not been particularly successful in helping students to develop conceptual understanding. Algebra courses in college tend to focus on procedural practice and memorization disconnected from conceptual understanding (Crooks & Alibali, 2014; Goldrick-Rab, 2006; Hammerman & Goldberg, 2003; Hodara, 2011; Rittle-Johnson & Schneider, 2014). Further, access to instruction that supports development of conceptual understanding has been found to be an equity issue, with more marginalized students often relegated to mathematics courses that provide fewer opportunities for developing conceptual skills such as sense-making (Schoenfeld, 2022; Stepter, 2023; Yeh et al., 2020). Thus, differences in algebraic conceptual understanding may contribute to differential college math and STEM outcomes.

What research exists on college students' algebraic thinking shows that understanding of core algebraic concepts is a critical factor in student success at all levels of mathematics from developmental algebra (Givvin et al., 2011; Stigler et al., 2010) through calculus (e.g., Frank & Thompson, 2021; Stewart & Reeder, 2017; Stewart et al., 2018). Without effective instruction in conceptual understanding, students may interpret algebra as a sequence of arbitrary algorithms that they apply without understanding (e.g., Hiebert & Grouws, 2007; Richland et al., 2012), and as a result, they may be unable to transfer knowledge from one context to another (Blanton Otto, 2018; Rebello et al., 2017; Richland et al., 2012) and may struggle to use procedures flexibly and in the appropriate contexts (e.g., Al-Mutawah et al., 2019; Givvin et al., 2011; Stigler et al., 2010). Conceptual understanding may be particularly important for STEM majors, since STEM students need to be able to apply their algebraic knowledge flexibly in novel contexts related to their particular STEM field (Quarles & Davis, 2017). However, the lack of existing research directly linking algebraic conceptual understanding

to students' college math and STEM outcomes makes it difficult to determine the precise role that algebraic conceptual understanding may play in the leaky STEM pipeline. This study aims to expand existing models of STEM retention to include algebraic conceptual understanding specifically as a potentially important factor that has been overlooked in prior frameworks.

### **Conceptual Framework: Models of STEM Retention**

Most existing models of STEM major retention that include academic factors as predictors rely on broad and more diffuse measures of knowledge, such as highest math course completed in high school, GPA, or scores on general standardized assessments (e.g., Alkhasawneh & Hargraves, 2014; Gipson, 2016; see review in Snyder & Cudney, 2017). These are useful for certain higher-level policy decisions (e.g., supporting students to enroll in more math courses in high school), but they do not provide sufficient information about *which types* of knowledge are particularly important for college mathematics and STEM courses, and thus are not designed to inform instructional change.

Standardized tests (e.g. ACT, SAT, state math exams) and classroom assignments (used to calculate high school grades) contain many different problem types on different topics, and often focus heavily on computation rather than measuring understanding (Alkhasawneh & Hargraves, 2014; Gipson, 2016; LeBeau et al., 2012; Maltese, 2008; Nicholls et al., 2013; Sahin et al., 2012; Wolniak, 2016). GPA and standardized assessments like these are relatively diffuse measures of more general (mathematics) knowledge, and, as such, it is difficult to tease out which knowledge types are important to STEM enrollment, persistence, and completion, or to measure whether students are being successfully supported to develop those knowledge types in different instructional contexts.

Measures of high school math course completion are particularly problematic as a proxy for knowledge, because this influences college mathematics course placement, which subsequently exerts its own effects on a student's STEM trajectory in multiple ways that can be unrelated to mathematical knowledge upon college entry. For example, the math course in which a student is placed at college entry can impact the time investment needed to complete a degree. Being able to start college by enrolling in higher-level math courses allows students to complete STEM degrees in fewer terms, which could influence decisions to enroll in or persist in the major, particularly for more marginalized students who have on average less financial or time capital to invest in college (Alkhasawneh & Hargraves, 2014; Crisp et al., 2009; Wladis et al., 2024a, 2024b). Placement may also influence STEM degree enrollment or completion through the quality of instruction and culture in lower- versus higher-level

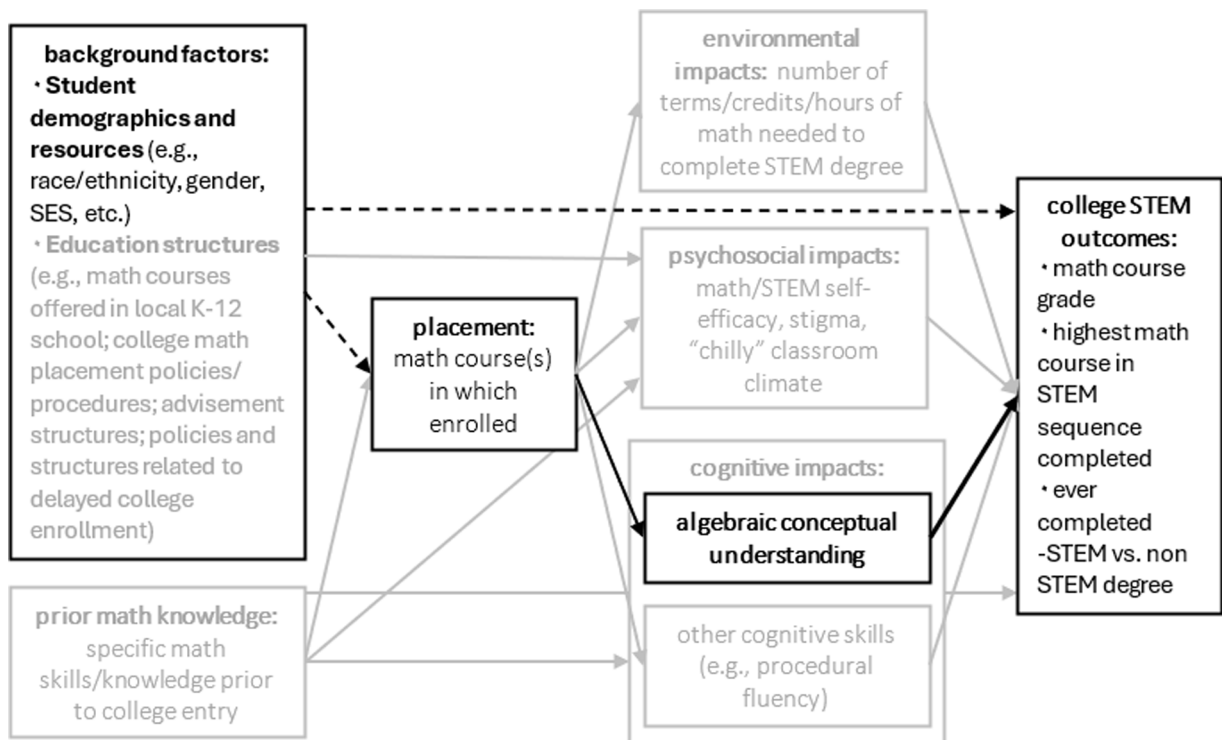


FIGURE 1. *Model of STEM retention accounting for algebraic conceptual understanding.*

Note. Grey arrows represent relationships tested in prior research. The thick black arrow at the far right represents main relationship tested in this study. The thin black arrow represents fixed effects by course (controlled for in all analyses). Dotted black arrows represent control variables included in “full” models that include controls.

math courses in college. Negative instructional experiences (e.g., Ellis et al., 2016; Hatfield et al., 2022) and stigma in lower- versus higher-level math courses (e.g., Larnell, 2016; Roberts, 2020), may influence student decisions to enroll or persist in STEM degrees. These experiences may lower students’ mathematics self-efficacy or emotional well-being, or a “chilly” climate in certain courses may convince them that they are not a good “fit” for a particular STEM major at the college even when they remain confident of their own skills (Charleston et al., 2014; Jensen & Deemer, 2019; Lin et al., 2018; Palid et al., 2023; Scott & Martin, 2014). A third possibility of how math placement could impact STEM attainment is that the types of learning opportunities afforded in higher- versus lower-level math courses to develop different types of knowledge may be important to STEM outcomes (Berkowitz & Stern, 2018; Quarles & Davis, 2017). For example, more advanced math courses often value conceptual skills such as abstraction, generalization, and justification, so students who have access to courses that stress these knowledge types may be better positioned to enroll and persist in STEM degrees (Quarles & Davis, 2017), regardless of the level of students’ mathematics knowledge upon college entry.

To better tease out the unique contribution of algebraic conceptual understanding from other potential impacts of

course placement or more diffuse measures of general mathematics knowledge, we theorize a more comprehensive model of STEM major retention that includes algebraic conceptual understanding as one component that has been unexplored in prior large-scale research (Figure 1). This model depicts both previously tested relationships and hypothesized relationships analyzed in this study.

Prior research has explored several pathways in the model depicted in Figure 1. For example, research that has explored changing college math placement policies (Ngo et al., 2018; Scott-Clayton, 2012; Stancher, 2019) has often considered the entire path from placement to college outcomes without analyzing the individual components between those two boxes separately. This research has also tended to focus more on non-STEM majors (an exception is Park et al. [2018], who found that placement mismatch impacted STEM college outcomes).

Other research has explored specific categories of factors that predict outcomes, with or without including math course placement as a predictive factor. For example, much of the research on co-requisite courses has focused on the pathway from placement to environmental impact and then to outcomes; this research has considered how the specific course sequence offered to students may impact their ability to complete degrees by influencing the number of terms needed

to finish a degree, and has tended to focus primarily on non-STEM majors (see reviews in Emblom-Callahan et al., 2019; Ryu et al., 2022). Some studies have considered the impact of the classroom or department climate on students' STEM outcomes, or the pathway from psychosocial impacts to outcomes depicted in Figure 1 (Charleston et al., 2014; Jensen & Deemer, 2019; Lin et al., 2018; Palid et al., 2023; Scott & Martin, 2014). These studies have tended to focus less on placement. Additional research has explored how various measures of general cognitive skills (math-specific or even broader) may predict STEM outcomes, considering the pathway as a whole from the cognitive impacts box (Figure 1) to outcomes without considering how individual components in the cognitive impacts box might predict STEM outcomes separately (Berkowitz & Stern, 2018; Fagan et al., 2019; Wai et al., 2009).<sup>2</sup>

However, we could find no studies that explicitly explore the black path in Figure 1, where measures of algebraic conceptual understanding are used to predict STEM outcomes. In this study, we aim to fill this gap by analyzing the extent to which algebraic conceptual understanding predicts STEM outcomes in college (the rightmost black arrow in Figure 1). We do this while including fixed effects by course (thin black line), thus controlling for the specific course in which a student is enrolled; also included are models with controls (dotted black lines), to control for background variables. Also investigated in this study are various mediation models, to explore the extent to which any observed differences in STEM outcomes by race/ethnicity or gender can be explained by differences in algebraic conceptual understanding. These models are described in more detail in the Method section.

Here, we conceptualize STEM outcomes as successes or failures measured at the institutional, rather than student level. The extent to which different students are successfully completing particular college math classes and/or electing to complete a STEM versus non-STEM degree is the result of many institutional and societal-level factors that are often beyond the control of individual students. Our goal in exploring the relationship between algebraic conceptual understanding and STEM-related college outcomes is *not* to label or categorize students based on their current knowledge states, or to make claims about which students should pursue STEM degrees—any such usage of this research for these purposes is expressly counter to the goals of this study and the validated use of the ACI (Wladis et al., 2024a, 2024b; Wladis et al., manuscript under review).

Rather, we frame a student's current level of algebraic conceptual understanding as the result of the cumulative effect of their prior formal and informal instructional experiences (inside and outside school, in both the K–12 and college context). By exploring the relationship between student scores on the ACI and their subsequent STEM outcomes in college, it allows for consideration of the extent to which

unequal access to opportunities to develop algebraic conceptual understanding may be a critical component of unequal access to STEM degrees. If algebraic conceptual understanding is a strong predictor of STEM outcomes in college, separate from the level of the specific math course in which a student has currently been placed, then this suggests that improving instruction in algebraic conceptual understanding (for instance, in high-school and early college) is an essential component of ensuring that every student has the opportunity to pursue a STEM degree.

### The Algebra Concept Inventory (ACI)

The Algebra Concept Inventory (ACI) has been validated in recent large-scale studies as a measure of the foundational algebraic conceptual understanding of students in a wide range of math courses in college starting at elementary algebra,<sup>3</sup> and it is the first instrument validated for this purpose (Wladis et al., 2024a, 2024b; Wladis, Murray, & Aly, 2025b; Wladis, Murray, Hachey, et al., 2025; Wladis et al., manuscript under review). “Classical”<sup>4</sup> algebra (in contrast to linear/abstract or “modern” algebra [Cooke, 2008]) denotes algebra focused on the transformation, representation, structure, and properties of algebraic expressions and equations, typically in progressively more complex, generalized and abstract ways as students move through the curriculum. The term “foundational” is used to denote concepts from classical algebra that are important to the college mathematics curriculum but also accessible to students as soon as they start elementary algebra (e.g., solutions of equations; structure of symbolic representations). While classical algebra may grow more advanced by including concepts beyond foundational algebra in subsequent mathematics courses (e.g., logarithmic or trigonometric functions), many core concepts from foundational algebra remain critical throughout all levels of mathematics courses in college (e.g., Frank & Thompson, 2021; Stewart & Reeder, 2017). For brevity, throughout the rest of this paper, we use the term “algebraic conceptual understanding” to refer to conceptual understanding of foundational classical algebra.

On the ACI, conceptual understanding was defined as “understanding related directly to the meaning of [an algebra] concept (Wladis et al., 2018), rather than ability to apply procedures or produce a formal proof” (Melhuish & Hicks, 2019, p. 123). In particular, ACI items (as described in Wladis et al., manuscript under review; Wladis et al., 2024a, 2024b) were designed to measure algebraic conceptual understanding separately, rather than as one component of a larger construct such as algebraic proficiency, as has been done in many national and international assessments (e.g., NAEP [NCES, 2024]; TIMSS [Mullis et al., 2023]). Thus, ACI items were designed to minimize construct-irrelevant variance (e.g., variation in performance due to procedural fluency skills or knowledge in domains outside algebra).

**Item 1.** Can the expressions  $(5x^{-4}3y)^{-2}$  and  $a^{-n}$  represent the same operations on the same things in the same order if we choose the right things for  $a$  and  $n$ ?

- a. No
- b. Yes, if  $a = y$
- c. Yes, if  $a = 3y$
- d. Yes, if  $a = 5x^{-4}3y$

**Item 2.** When solving an equation that contains an  $x$ , a student gets the answer  $8 = 8$ .

This means that:

- a. This could only happen if the student made a mistake somewhere.
- b. Every possible value of  $x$  is a solution to the original equation.
- c. There is no solution to the original equation.
- d. The solution to the original equation is  $x = 8$ .

FIGURE 2. Algebra Concept Inventory example items.<sup>5</sup>

ACI items were created both (1) to require mathematical sensemaking to answer correctly (i.e., a correct answer cannot be obtained solely through executing rote procedures or restating memorized facts); and (2) to *not* require calculation or transformation to answer correctly. Procedural fluency is important and can be related to conceptual understanding, but items measuring this skill were deemed inappropriate on this assessment intended to measure conceptual understanding specifically (Wladis et al., 2024a, 2024b; Wladis et al., manuscript under review).

As an illustration, we present two ACI items focused on different areas in algebra: the first explores conceptions of algebraic syntax structure, and the second conceptions of solutions of equations (See Figure 2).

For Item 1, students conceptualizing the standard intended structure of the syntax of the expression would typically choose D. Students who choose A, B, or C typically extract the structure of the syntax from non-mathematically salient features drawn from instructional experiences and may conceptualize the meaning and structure of algebraic syntax differently than those who link it to standard operational precedence (e.g., Wladis, Murray, & Aly, 2025a; Wladis, Sencindiver, et al., 2023). For Item 2, students who understand both that valid transformations preserve the solution of an equation and that all values of  $x$  make the equation  $8 = 8$  true typically choose B. Students who choose A, C, and D often have alternative conceptions about what it means to solve an equation or what it means for an equation to have a solution; for example, some students who choose D conceptualize solving an equation as a process of performing “allowed” procedures to get a numerical value as a final result in a way that is de-linked from the standard definition of an equation solution (e.g., Wladis et al., 2022). Both of these items measure conceptual understanding as defined for the purposes of the ACI because both require mathematical sensemaking (neither can be answered solely by restating

memorized facts or performing a rote procedure without understanding its justification), and neither require computation or transformation to answer correctly.

The ACI is currently the only large-scale validated college concept inventory that focuses explicitly on college students’ conceptual understanding of foundational algebra and therefore does not include any algebra concepts that would be inaccessible to students in an elementary algebra course (Wladis et al., 2024a, 2024b; Wladis et al., manuscript under review). Measuring foundational classical algebra across all math courses in college is important because students in elementary and intermediate algebra classes in college have often had little access to rich instruction that supports the development of conceptual understanding in addition to procedural skills (Crooks & Alibali, 2014; Goldrick-Rab, 2006; Hammerman & Goldberg, 2003; Hodara, 2011), and because foundational classical algebra concepts have been found to impact student work in mathematics courses throughout the college curriculum (Frank & Thompson, 2021; Stewart & Reeder, 2017; Stewart et al., 2018; A. Weinberg et al., 2016). A detailed interpretation and use statement for the ACI can be found at [www.algebraconceptinventory.org](http://www.algebraconceptinventory.org).

Analysis has shown that the ACI has excellent reliability and validity as a measure of the foundational algebraic conceptual understanding of college students across a wide range of mathematics courses (Wladis et al., 2024a, 2024b; Wladis et al., manuscript under review). Because the ACI requires only limited prior exposure to algebra (i.e., current/prior enrollment in Algebra I or elementary algebra), covers a wide range of algebra concepts, and has undergone large-scale mixed methods validation with college students, it is an excellent candidate for exploring hypotheses about whether algebraic conceptual understanding can predict college students’ math course outcomes and STEM degree progress or attainment. In addition, because the ACI has been generated based on specific constructs that have been

explicitly operationalized (i.e., each item on the inventory is designed to measure specific concepts and conceptions [Wladis et al., manuscript under review; Wladis et al., 2024a, 2024b]), if ACI scores do predict math and STEM outcomes in college, the framework used to generate the ACI provides an actionable starting point for revising curriculum and instruction to better prepare students for college math and STEM courses.

## Methods

### *Data Collection*

Data for this study were obtained from the City University of New York (CUNY) Institutional Research (IR) Offices. This included information about all students enrolled in any mathematics course at the largest community college at CUNY between Spring 2019 and Summer 2023. Data included the course in which the student was enrolled, the grade in that course, gender, race/ethnicity, age, GPA at the beginning of the term, number of credits earned by the beginning of the term, major, any degrees earned, and home zipcode (so that this could be merged with household income by zipcode data from the American Community Survey and used as a partial proxy for socioeconomic status [SES]). For more details, see the Measures subsection below. For some robustness tests (see Table A1 in the Appendix), a measure of English-language-learner (ELL) status was also used; this was based on placement assessments administered by the college. Results in Table A1 demonstrate that models have identical or near identical coefficients, standard errors, and  $p$  values, regardless of whether ELL status is included in the model.

This study merged institutional records with an existing dataset of scores on the ACI originally generated during ACI validation. Math classes included elementary algebra, intermediate algebra, mathematics for Elementary Educators I and II, quantitative reasoning, various levels of statistics, mathematics for liberal arts majors, mathematics for health science majors, discrete mathematics, precalculus, Calculus I, II and III, advanced calculus, linear algebra, abstract algebra, and differential equations. A total of 402 ACI items were taken by students during validation. CUNY provided institutional research data for a total of 6,582 students with ACI scores.<sup>6</sup>

### *Analytical Methods*

*Measures.* The primary independent variable of interest was a student's score on the ACI. We explored a variety of dependent variables as potential outcomes that could be predicted by ACI score. First, course grade was generated by translating letter grades into a corresponding GPA scale (0–4); official withdrawals, incomplete grades, and audit grades were not assigned any grade value. Because some courses did not have grades beyond pass/fail designations

(e.g., developmental courses), we also computed “successful course completion,” defined as completing the course with a C or better, the typical criterion necessary for transfer or credit in the major. A variable to indicate whether a student obtained a STEM versus non-STEM degree during the study period was generated by considering the CIP code of the associated degree (Manly et al., 2018). In addition to dependent variable data on student course grades, majors, and degree completion, control variables in this study were also taken from Institutional Research datasets and included: gender, race/ethnicity, age, GPA at the beginning of the term, number of credits earned at the beginning of the term, and home zipcode as a partial proxy for SES (Wladis, Hachey, et al., 2024). Home zipcode was used to generate median household income of the zipcode using the U.S. Census Bureau's pre-pandemic American Community Survey data (because pre-pandemic data was not subject to the issues with data quality that the Census Bureau cited with the 2020 data due to the pandemic; see, for example, Villa Ross et al., 2021).

There were virtually no missing data in IR variables; the only variables with missing data were race/ethnicity, gender and age (missing under 0.1%) and U.S. zipcode (missing 1.6%). Zipcode missingness included some data entry errors (e.g., zipcode entered into wrong field) and some foreign students with no U.S. address on file (96.3% of students with a foreign address on file had a U.S. address on file, so this represented a small percentage of foreign students). Because of the small number of missing IR variables, and the fact that we cannot assume that zipcode is missing at random,<sup>7</sup> we used listwise deletion. This may mean that full model results (that use control variables) may be less applicable to a small group of foreign students who do not have U.S. addresses on file.

For mediation analyses only, race/ethnicity was dichotomized into a single variable to provide sufficient statistical power to detect effects: underrepresented minority status (URM) was assigned to students who identified as Black, Hispanic, or American Indian/Native Alaskan, since these groups have historically been underrepresented and underserved in STEM majors (e.g., Black et al., 2021; Hatfield et al., 2022; Wright et al., 2023; Zhang, 2021). While there are limitations to dichotomizing such data, categories were collapsed into a binary variable to increase power so that important differences could be better detected by statistical models, and because these categories reflect to some extent politically and socially constructed categories that are at the root of much structural marginalization and discrimination (Balestra & Fleischer, 2018). This variable was *not* dichotomized in other regression models.

For each math course in which a student was enrolled during the study period, we classified courses two ways: (1) by commonly-used categories typical to U.S. STEM degree math course sequences (elementary algebra, intermediate/college algebra, precalculus, Calculus I, and Calculus II); and (2) by course prerequisite sequence (see Table 1).

TABLE 1  
Course Sequence Classification Based on Prerequisites

Prerequisite	Sequence Number
No Algebra Prerequisite	1
Elementary Algebra Pre-Requisite	2
Intermediate/College Algebra Prerequisite	3
Precalculus Prerequisite, or Equivalent	4
Calculus I Prerequisite	5
Calculus II Prerequisite	6
Calculus III or Linear Algebra Prerequisite	7

“Elementary” and “intermediate” algebra are often labeled as “below-college-level” and are often designed to review procedures from first- and second-year K–12 algebra courses, respectively. “College” algebra may be practically indistinguishable from intermediate algebra, or may be more advanced, and typically carries college credit. We refer to these designations not to endorse them (issues with these definitions are discussed at length elsewhere; Wladis, Bjorkman, et al., 2023; Wladis, Makowski, et al., 2023) but simply to describe existing course designations and prerequisite sequences.

Analyses controlled for specific course in which a student was enrolled or the sequence number of the enrolled course, typically by including the level in the course sequence as a fixed effect (unless analysis was already limited to a single course or sequence level), because scores on the inventory are correlated with the sequence level of the course in which a student is enrolled (for more details see Wladis et al., 2024a, 2024b, manuscript under review; Wladis, Murray, & Aly, 2025a; Wladis, Murray, Hachey, et al., 2025b). Random effects were also used to control for clustering by instructor, as students enrolled in math courses with the same instructor are more likely to have similar outcomes. More details on how analysis was carried out are provided in the next subsections.

*ACI scores.* ACI scores were originally generated for each student during validation studies of the ACI (Wladis et al., 2024a, 2024b; Wladis et al., manuscript under review), using two-parameter logistic (2PL) item-response-theory (IRT) models (Birnbaum, 1968). These were estimated using marginal maximum likelihood (MML) estimation on each wave, using the R package “mirt” (Chalmers, 2012). 2PL models are commonly used to analyze dichotomous items and consist of logistic models where items are allowed to differ on both difficulty level and discrimination (which represents how well they distinguish between students with higher versus lower levels of “algebraic conceptual understanding”). The equation for 2PL models is:

$$P(Y_{ij} = 1|\theta_j) = \frac{\exp(a_i(\theta_j - d_i))}{1 + \exp(a_i(\theta_j - d_i))}$$

where  $a_i$  is the discrimination and  $d_i$  the difficulty of item  $i$ , and  $\theta_j$  is the theta score of person  $j$ . These 2PL models were used to generate an ACI score for each student, representing the student’s  $\theta$  value, or the number of standard deviations above/below the mean their score on the ACI is in comparison to other students in the population. The ACI was administered throughout the term, so there was not one set date of administration; the median date for taking the ACI was about 6 weeks into the 15-week term, with the 10th percentile completion date 1 month into the term and the 90th percentile completion date 1 month prior to final exams. All students took the ACI prior to final exam week. Some students took the ACI more than once in a single term (e.g., if they were enrolled in multiple math courses that term). For those students, we took the mean of ACI scores for that term for use in analysis.

*Regression models.* Multilevel fixed effects linear and logistic regression were used to assess the validity of the ACI as a predictor of college math and STEM outcomes. While we give a few descriptive results without fixed effects, most analyses reported here control for fixed effects by course level (where outcomes are only compared among students at the same level in the course sequence) and random effects by course instructor (to control for the fact that students taught by the same instructor may have more similar outcomes). The linear multilevel model equations used were:

$$\text{First level equation: } Y_j = \beta_{0j} + \beta_{1j}x_1 + \dots + \beta_{nj}x_n + e$$

$$\text{Second level equations: } \begin{aligned} \beta_{0j} &= \alpha_{00} + \alpha_{10}z + u_{0j} \\ \beta_{ij} &= \alpha_{0i} + u_{ij} \end{aligned}$$

for  $i = 1, \dots, n$  and  $j = 1, \dots, m$ , and where  $Y_j$  represents the dependent variable (e.g., math course grade, whether the student ever successfully completed intermediate algebra if they enrolled, whether the student received a STEM vs. non-STEM degree during the study period). Here  $j$  denotes the specific group (i.e., students are grouped based on the mathematics instructor a student has in their current course) and  $m$  denotes the number of groups (i.e., instructors) in the sample. The  $x_1, \dots, x_n$  represent independent variables;  $x_1$  represents the ACI score;  $x_2$  represents the course sequence number (see Table 1) to control for fixed effects of course level;  $x_3, \dots, x_n$  represent control variables (e.g. race/ethnicity, GPA, median household income). In these equations,  $\alpha_{00}$  is the average value for the whole population on the dependent variable (e.g., mean math course grade), and  $u_{0j}$  is an error term specific to each group,

denoting the amount by which the mean value of the dependent variable (e.g., mean math course grade) for each group is above or below the whole sample mean. The variable  $e$  denotes the amount by which each individual deviates from the group mean, just as the variables  $x_1, \dots, x_n$  take on different values for each subject.

For all models with a binary dependent variable, we ran both logistic and linear probability fixed-effects models. Both produced similar results, so we report linear probability models because readers are much less likely to misinterpret the coefficients of linear probability models than odds ratios (e.g., Norton & Dowd, 2018) and because we cannot compare coefficients across logistic regression models due to rescaling (Buis, 2010; Erikson et al., 2005; Long, 1997; Winship & Mare, 1984; Wooldridge, 2002).

For logit models (e.g., successful completion of an intermediate algebra course; STEM degree attainment), the same equations were used as in linear regression, but with a logit link employed to model the probability distribution:

$$\text{First level equation: } \lambda(Y_j) = \beta_{0j} + \beta_{1j}x_1 + \dots + \beta_{nj}x_n$$

$$\text{Logit link: } \lambda(Y_j) = \frac{e^{Y_j}}{1 + e^{Y_j}}$$

$$\text{Second level equations: } \begin{aligned} \beta_{0j} &= \alpha_{00} + \alpha_{10}z + u_{0j} \\ \beta_{ij} &= \alpha_{0i} + u_{ij} \end{aligned}$$

where  $\lambda(Y_j)$  represents the probability value of the dependent variable (e.g., the probability that the student successfully completed intermediate algebra, or received a STEM degree), and other variables are the same as in the linear regression models.

*Mediation models.* This paper also considers models that aim to understand to what extent disparities in dependent variables such as math course grade or STEM degree attainment by race/ethnicity (or gender) can be explained by disparities in ACI score by race/ethnicity (or gender). Following van VanderWeele and Robinson (2014), we conceptualize ACI score as a mediator and employ statistical models from the mediation literature to decompose disparities into two parts.<sup>8</sup> The first part is an “indirect” disparity that is a function of the combined effects of: (a) the relationship between ACI score and race/ethnicity (or gender) and (b) the relationship between ACI score and the dependent variable (e.g., math course grade; STEM degree attainment). The second part is a “direct” disparity, which quantifies the proportion of the total disparity that would remain even if the distribution of ACI score were to be equalized across racial/ethnic (or gender) groups. Direct and indirect disparities, as we call them here, are the same coefficients that are often referred to in the mediation literature as direct and indirect “effects”; however, we avoid the use of the term “effect” because this

is an observational rather than causal study (however, to improve readability, we include the term “effect” in quotes inside a parenthetical reference in tables for those readers more familiar with the terminology used in the mediation literature; we couple this with reminders to the reader that these results do not support causal inference).

For mediation analysis, we used the KHB decomposition method, available in Stata (Buis, 2010). This is a general decomposition method based on an SEM framework that can be used with either linear or logistic regression. In linear regression mediation is modeled as follows:

$$Y = \alpha_D + \beta_D X + \Gamma_D M + \epsilon_D \quad (1)$$

$$Y = \alpha_T + \beta_T X + \epsilon_T \quad (2)$$

where  $Y$  is the dependent variable,  $X$  is the independent variable of interest and  $M$  is the potential mediator.  $\beta_D$  represents the direct disparity of  $X$  on  $Y$ ,  $\beta_T$  represents the total disparity of  $X$  on  $Y$ , and  $\beta_I = \beta_T - \beta_D$  represents the indirect disparity of  $X$  on  $Y$ . Then residuals of the linear regression of  $M$  on  $X$  are calculated:

$$R = M - (a + bX) \quad (3)$$

and then  $R$  is used in place of  $M$  in Equation (1) that estimates the direct disparity of  $X$  on  $Y$ :

$$Y = \tilde{\alpha}_T + \tilde{\beta}_T X + \tilde{\Gamma}_T R + \tilde{\epsilon}_T \quad (4)$$

Adding  $R$  to the model adds only the component of  $M$  that is uncorrelated with  $X$ , so the residuals have the same standard deviation in (1) and (4), but the coefficient of  $X$  in (1) is the direct disparity, whereas the coefficient of  $X$  in Equation (4) is the total disparity.

### *Descriptive Statistics and Weighting*

Students who consent to have their scores on a math assessment included in research may vary from those who do not; for example, they might be more likely to have higher math self-efficacy or achievement. Thus, it was important to check for potential differences between students for whom ACI scores are available vs. those who are not. Descriptive statistics comparing these two groups are provided in Table 2.

Students for whom ACI scores were available were more likely to be Asian and less likely to be Hispanic; more likely to be women; were slightly older on average; had a slightly lower number of credits earned by the start of the term; had slightly higher GPAs on average, and were enrolled in a slightly higher math course in the sequence. However, there was no difference in terms of other race/ethnicity categories or median household income of zipcode. We note that despite the observed

TABLE 2

*Summary Statistics Before Weighting, Comparing Those for Whom ACI Score Was Available to Total Math Course Population*

	All Students Enrolled in Math Courses			Only Those Students With ACI Score		
	Mean/Prop.	SE	CI	Mean/Prop.	SE	CI
<b>Ethnicity</b>						
Black	33.0%	0.1%	[32.8%, 33.3%]	33.2%	0.5%	[32.2%, 34.1%]
Hispanic	36.8%	0.1%	[36.5%, 37.1%]	31.9%	0.5%	[30.9%, 32.8%]
White	12.8%	0.2%	[12.4%, 13.2%]	12.9%	0.6%	[11.7%, 14.2%]
Asian/PI	16.9%	0.1%	[16.7%, 17.1%]	21.6%	0.4%	[20.8%, 22.4%]
AI/NA	0.4%	0.0%	[0.4%, 0.5%]	0.4%	0.1%	[0.3%, 0.6%]
<b>Gender</b>						
Male	46.9%	0.3%	[46.4%, 47.4%]	41.0%	0.9%	[39.3%, 42.8%]
Female	53.0%	0.2%	[52.7%, 53.3%]	59.0%	0.5%	[58.0%, 59.9%]
Diverse Gen. Cat.	0.1%	0.0%	[0.1%, 0.1%]	0.0%	0.0%	[0.0%, 0.0%]
Age	22.3	0.0	[22.3, 22.4]	23.2	0.1	[23.1, 23.3]
<b>GPA</b>						
No GPA	36.1%	0.2%	[35.7%, 36.5%]	37.4%	0.7%	[36.0%, 38.7%]
Under 2.0	13.1%	0.1%	[12.9%, 13.3%]	9.0%	0.3%	[8.5%, 9.6%]
2.0–2.49	10.9%	0.1%	[10.8%, 11.1%]	8.4%	0.3%	[7.9%, 9.0%]
2.5–3.99	12.0%	0.1%	[11.8%, 12.1%]	10.4%	0.3%	[9.8%, 11.0%]
3.0–3.49	12.7%	0.1%	[12.5%, 12.9%]	13.9%	0.4%	[13.2%, 14.6%]
3.5–4.0	15.2%	0.1%	[15.0%, 15.4%]	20.8%	0.4%	[20.0%, 21.6%]
Creds Start of Term	17.5	0.1	[17.4, 17.7]	17.0	0.2	[16.6, 17.4]
Course Seq. Number	2.36	0.00	[2.35, 2.37]	2.50	0.01	[2.48, 2.53]
Median Household Income of Zipcode	\$67,663	\$115	[\$67,437, \$67,889]	\$67,494	\$371	[\$66,766, \$68,222]

*Note.* ACI = Algebra Concept Inventory; SE = Standard Error; CI = 95% Confidence Interval; PI = Pacific Islander; AI = American Indian; NA = Native American.

differences, students with all GPA levels, course levels, and racial/ethnic, and gender groups were well-represented in the ACI data, and, thus, weighting can be used to address variation on observables between students for whom ACI scores are available and the full sample. To address this variation, we performed entropy balancing (Hainmueller, 2012). Entropy balancing is a weighting method, like other propensity score weighting methods, where the goal is to re-weight data points to improve covariate balance so that the sample is more reflective of the population, or so that a treatment variable (in this case, whether a student took the ACI) becomes independent of measured background characteristics (Powell et al., 2019).

Entropy balancing differs from other weighting methods because it allows the researcher to impose balance constraints, requiring that covariate distributions of the sample and the population match on all prespecified moments (Hainmueller, 2012). In contrast, other propensity score methods often require the researcher to manually go through multiple rounds of weighting and checking the data for covariate balance, and improving the balance on some covariates may occur at the cost of worsening balance on others (Thomas et al., 2023). Weighting using entropy balancing resulted in a dataset of students for whom ACI scores were available for which the mean/proportion on all control

variables of the new weighted dataset matched the characteristics of the population of all students enrolled in math courses with standardized mean differences reduced to zero across all variables (see Table 3).

Table 3 illustrates that the resulting weighted dataset used for this study has covariate distributions that reflect the full population of students enrolled in math courses in the study population.

## Results

We explored whether and to what extent scores on the ACI predict mathematics course grades, completion of key courses in the STEM mathematics sequence, and completion of a STEM versus non-STEM degree.<sup>9</sup> We also considered whether ACI scores explain any differential outcomes by race/ethnicity or gender.

### *ACI Score as Predictor of Mathematics and STEM Outcomes in College*

*ACI score as a predictor of math course grades.* Table 4 reports results of overall regression models on the full sample of all courses, using course sequence as a fixed effect; it also reports the results of separate regression models for each

TABLE 3

Summary Statistics After Weighting, Comparing Those for Whom ACI Scores Were Available to Total Math Course Population

	All Students Enrolled in Math Courses		Only Those Students With ACI Score		SMD
	Mean	SD	Mean	SD	
Ethnicity					
Black	33.0%	0.1%	33.0%	0.5%	0.0
Hispanic	36.8%	0.1%	36.8%	0.5%	0.0
White	12.8%	0.2%	12.8%	0.6%	0.0
Asian/PI	16.9%	0.1%	16.9%	0.4%	0.0
AI/NA	0.4%	0.0%	0.4%	0.1%	0.0
Gender					
Male	46.9%	0.3%	46.9%	0.9%	0.0
Female	53.0%	0.2%	53.0%	0.5%	0.0
Diverse Gen. Cat.	0.1%	0.0%	0.1%	0.0%	0.0
Age					
	22.3	0.0	22.3	0.1	0.0
GPA					
No GPA	36.1%	0.2%	36.1%	0.7%	0.0
Under 2.0	13.1%	0.1%	13.1%	0.3%	0.0
2.0–2.49	10.9%	0.1%	10.9%	0.3%	0.0
2.5–3.99	12.0%	0.1%	12.0%	0.3%	0.0
3.0–3.49	12.7%	0.1%	12.7%	0.3%	0.0
3.5–4.0	15.2%	0.1%	15.2%	0.4%	0.0
Credits Start of Term					
	17.5	0.1	17.5	0.2	0.0
Course Seq. Number					
	2.4	0.0	2.4	0.0	0.0
Median Household Income of Zipcode					
	\$67,663	\$115	\$67,664	\$376	0.0

Note. ACI = Algebra Concept Inventory; SD = Standard Deviation; SMD = Standardized Mean Difference; PI = Pacific Islander; AI = American Indian; NA = Native American.

course type. Table 4 includes the results of base models (without controls) and full models (with controls). Base models have the benefit that they illustrate existing relationships between variables based on characteristics of the students who are actually enrolled in current courses; full models have the benefit of showing us what the relationship would theoretically be if all students in each course were compared only to other students with the same characteristics on the control variables in that course. Thus, each type of model provides different information about existing patterns.

Table 4 indicates that on average, a one standard deviation (*SD*) increase in ACI score correlates with a roughly 0.4 grade point increase (on a scale of 1–4) for the math course in which a student is enrolled ( $p < .001$ ) when controlling for fixed effects by course sequence number (a measure of course level based on course prerequisite requirements). This relationship is reduced somewhat but still highly significant ( $p < .001$ ) when including all covariates in the model, suggesting that some of the variation in ACI score is likely due to correlation between various student characteristics (GPA, number of prior credits earned, race/ethnicity, gender, age, and median household income of zipcode) course grade, and ACI score.

When looking at individual courses, ACI scores significantly predicted grades in intermediate/college algebra,<sup>10</sup> precalculus, and Calculus I in models both with and without controls, and in differential equations once all controls are added. ( $p < .001$  for all models except model for differential equations, where  $p = .006$ ). This result makes sense since all these courses rely heavily on algebra. Interestingly, ACI scores also significantly predicted grades in courses for future elementary school teachers (listed as mathematics education in Table 4), statistics courses, quantitative reasoning courses, and math courses for health science majors (e.g., nursing, paramedics, respiratory therapy), even though these courses are not typically thought of as part of the algebra trajectory. This suggests that algebraic conceptual understanding, as measured by the ACI, may be important for these courses as well. Scores on the ACI were not predictive, in either base or full models, of grades in Calculus III, and were only predictive of grades in Calculus II in base models, and not full models with controls, in this sample.

*ACI as predicting completion of math courses in the STEM major sequence.* Next, we investigated the extent to which students' scores on the ACI are predictive of whether they

TABLE 4  
*ACI Score as a Predictor of Math Course Grade (Grade Scale: 1–4)*

Base Models					Full Models <sup>a</sup>				
Fixed Effects by Course Sequence Number, Regression Results									
	Coeff.	SE	<i>p</i>	sig.	Coeff.	SE	<i>p</i>	sig.	<i>n</i>
ACI Score	0.40	0.03	< .001	***	0.26	0.02	< .001	***	7,908
Sequence	0.08	0.03	.007	**	0.05	0.03	0.066	.	
Separate Models for Each Course Type, Regression Results									
	Coeff.	SE	<i>p</i>	sig.	Coeff.	SE	<i>p</i>	sig.	<i>n</i>
Intermediate/College Algebra	0.55	0.05	< .001	***	0.40	0.04	< .001	***	1,960
Precalculus	0.51	0.04	< .001	***	0.36	0.04	< .001	***	2,167
Calculus I	0.46	0.09	< .001	***	0.24	0.05	< .001	***	1,020
Calculus II	0.28	0.10	.006	**	0.06	0.11	.600		495
Calculus III	0.05	0.11	.667		0.01	0.08	.951		201
Differential Equations	0.62	0.38	.105		0.90	0.33	.006	**	68
Mathematics Education	0.38	0.08	< .001	***	0.23	0.06	< .001	***	337
Statistics	0.34	0.05	< .001	***	0.20	0.05	< .001	***	2,166
Quantitative Reasoning	0.35	0.11	.001	**	0.18	0.09	.048	*	625
Math For Health Majors	0.35	0.07	< .001	***	0.18	0.07	.009	**	606

Note. Multi-Level Weighted Regression, fixed effects by course (first row) or separate models by courses (listed in first column), random effects by instructor, math course grade as dependent variable, ACI score as independent variable for which coefficients are reported.

<sup>a</sup>Full models include control variables: GPA at start of term, number of prior credits earned by start of term, race/ethnicity, gender, age, and median household income of zipcode.

\*\*\*  $\alpha = 0.001$ , \*\*  $\alpha = 0.01$ , \*  $\alpha = 0.05$ , .  $\alpha = 0.10$ .

TABLE 5  
*ACI Score as a Predictor of Successful Completion of Core Math Courses in STEM Major Requirements*

Base Models					Full Models <sup>a</sup>				
Ever Completed:	Coefficient for ACI Score	SE	<i>p</i>	sig.	Coefficient for ACI Score	SE	<i>p</i>	sig.	<i>n</i>
Elementary Algebra	0.16	0.04	< .001	***	0.15	0.05	.001	**	499
Intermediate/College Algebra	0.14	0.02	< .001	***	0.11	0.02	< .001	***	2,597
Precalculus	0.13	0.02	< .001	***	0.09	0.02	< .001	***	2,418
Calculus I	0.15	0.03	< .001	***	0.09	0.02	< .001	***	1,248
Calculus II	0.11	0.02	< .001	***	0.08	0.03	.008	**	611

Note. Linear Probability Models Using Multi-Level Weighted Regression: fixed-effects by course sequence number of current course, random effects by instructor, ever completed (if attempted) particular math course (listed in first column) as dependent variable, ACI score as independent variable for which coefficients are reported.

<sup>a</sup>Full models include control variables: GPA at start of term, number of prior credits earned by start of term, race/ethnicity, gender, age, and median household income of zipcode.

\*\*\*  $\alpha = 0.001$ , \*\*  $\alpha = 0.01$ , \*  $\alpha = 0.05$ , .  $\alpha = 0.10$ .

successfully completed math courses in the standard STEM math sequence during the study period (Table 5). We considered both base and full fixed effects regression models predicting whether students ever successfully completed elementary algebra; intermediate or college algebra; precalculus, Calculus I, or Calculus II, if enrolled. We only considered those students who had not yet successfully completed

these courses *prior* to taking the ACI, so that scores are only used to predict *future* (not past) course completion.<sup>11</sup>

Table 5 illustrates that ACI scores significantly predicted whether a student would ever successfully complete each of the five core STEM major math requirements if they attempted them, both in base and full models; this was significant at the  $\alpha = 0.001$  level for all models, except the full model

TABLE 6  
*ACI Score as Predictor of STEM Versus Non-STEM Degree Completion*

Base Models				
Ever Attained STEM Versus Non-STEM Degree	Coefficient	SE	p	sig.
ACI Score	0.058	0.009	< .001	***
ACI Score	0.059	0.010	< .001	***
Course Grade, Standardized	0.033	0.011	.003	**
Full <sup>a</sup> Model				
Ever Attained STEM Versus Non-STEM Degree	Coefficient	SE	p	sig.
ACI Score	0.053	0.010	< .001	***
ACI Score	0.051	0.010	< .001	***
Course Grade, Standardized	0.040	0.012	.001	**

*Note.* ACI = Algebra Concept Inventory; STEM = Science, Technology, Engineering, and Mathematics. Linear Probability Models Using Weighted Regression: fixed-effects by course sequence number of current course, ever attained STEM (vs. non-STEM) degree as dependent variable, ACI score as independent variable.

<sup>a</sup>Full models include control variables: GPA at start of term, number of prior credits earned by start of term, race/ethnicity, gender, age, and median household income of zipcode.

\*\*\*  $\alpha = 0.001$ , \*\*  $\alpha = 0.01$ , \*  $\alpha = 0.05$ , .  $\alpha = 0.10$ .

predicting Calculus II completion where  $p = 0.008$ . We note that because all models included fixed effects for course sequence level, these differences were not due to correlations between ACI score and current course level—only students currently enrolled in the same level of the math course sequence are being compared in these models.

Thus, ACI score was strongly predictive of whether a student would successfully complete courses in the standard STEM major math course sequence, with the strongest relationship present for elementary algebra. This aligns with the design of the ACI to focus on concepts first introduced in elementary algebra. Students were 16 percentage points more likely to successfully complete elementary algebra for every 1 *SD* increase in ACI score (15 percentage points after controls). Additionally, students were also significantly more likely to successfully complete intermediate/college algebra, precalculus, Calculus I or Calculus II if they ever enrolled in these courses, with every 1 *SD* increase in ACI score correlating with 11–15 percentage point increases in successful course completion of these classes (8–11 percentage points with controls). This illustrates the importance of algebraic conceptual understanding to completing the core mathematics courses required for a STEM major.

*ACI as predicting STEM versus non-STEM degree completion.* Next, we explored whether ACI scores predicted student degree attainment in STEM versus non-STEM fields. This analysis only considered those students who completed a 2- or 4-year degree during the study period ( $n = 3,556$ ). We analyzed whether ACI score predicted STEM versus non-STEM major completion while

controlling for course fixed effects by sequence level. Regression results can be seen in Table 6.

Table 6 shows that for every 1 *SD* increase in ACI score, students were 5.8 percentage points more likely to complete a STEM versus non-STEM degree when including fixed effects for course sequence number. Thus, for students at the same level of the math course sequence, those with higher ACI scores were significantly more likely to graduate with a STEM degree. Adding in controls did not change this relationship, altering its magnitude only slightly (primarily reducing the coefficient only for course grade as a predictor).

Even more notable is that ACI score separately predicted STEM degree completion, above and beyond math course grade alone. ACI score was just as good (or better) than math course grade at predicting STEM versus non-STEM degree completion: in base models, an increase in 1 *SD* in course grade correlated with a 3.3 percentage point increase in STEM versus non-STEM degree completion and an increase in 1 *SD* in ACI score correlated with a 5.9 percentage point increase (this shifted to 5.1 and 4.0 percentage points, respectively, in full models). Adding in course grade as a predictor in addition to ACI score only slightly altered the coefficient for the predictive relationship between ACI score and STEM degree completion, with 1 *SD* increase in ACI score still predicting over a five percentage point greater likelihood of graduating with a STEM versus non-STEM degree (when comparing only those students with the same grade in the same math course to one another). Thus, the ACI is measuring something important to STEM versus non-STEM degree completion that is not currently captured by course grades alone.

TABLE 7

*ACI Score as Mediator Between Underrepresented Minority Status (URM)<sup>a</sup> or Gender and Math Course Grade*

Base Models					Full Models <sup>b</sup>			
	Coef.	SE	P	Sig.	Coef.	SE	P	Sig.
Underrepresented Minority Status								
Total Discrepancy (Total “Effect”)	-0.65	0.05	< 0.001	***	-0.42	0.06	< 0.001	***
Direct Discrepancy (Direct “Effect”)	-0.54	0.05	< 0.001	***	-0.36	0.06	< 0.001	***
Indirect Discrepancy (Indirect “Effect”)	-0.11	0.01	< 0.001	***	-0.06	0.01	< 0.001	***
Female Gender								
Total Discrepancy (Total “Effect”)	0.24	0.04	< 0.001	***	0.23	0.05	< 0.001	***
Direct Discrepancy (Direct “Effect”)	0.29	0.04	< 0.001	***	0.26	0.05	< 0.001	***
Indirect Discrepancy (Indirect “Effect”)	-0.04	0.01	< 0.001	***	-0.03	0.01	0.001	**

*Note.* Mediation using the KHB method: Mediation of the relationship between underrepresented minority status or gender and math course grade by ACI score, fixed-effects by course sequence number, clustering by instructor.

<sup>a</sup>Underrepresented minority status indicates students of color who have been traditionally underrepresented and underserved in STEM fields; this includes Black, Hispanic, and American Indian/Native Alaskan students, but not White or Asian/Pacific Islander students. <sup>b</sup>Full models included control variables: GPA at start of term, number of prior credits earned by start of term, race/ethnicity (for gender models), gender (for URM models), age, and median household income of zipcode, instructor.

\*\*\*  $\alpha = 0.001$ , \*\*  $\alpha = 0.01$ , \*  $\alpha = 0.05$ , .  $\alpha = 0.10$ .

In particular, we performed an incremental validity analysis as a robustness check of these results (see Table A3 in the Appendix). Comparisons of Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) for all full models in Tables 5, 6, and 7 illustrates that there is “very strong” (Raftery, 1995, p. 140) evidence that ACI scores add significant information in predicting course grade, whether students ever completed a math course in the STEM sequence, or whether they graduated with a STEM versus non-STEM degree. In particular, for all models except the one full model from Table 5 predicting Calculus III course grade there is positive evidence that ACI scores add predictive information, and for all models except the one full model from Table 5 predicting Calculus II course grade there is “very strong” (Raftery, 1995, p. 139) evidence that ACI scores add predictive information above GPA, math course grade, and other SES and demographic variables. In particular, the “conservative” (1995, p. 139) cutoff suggested by Raftery to demonstrate “very strong” evidence of added information from an additional predictor is a difference in BIC scores greater than 10; all the BIC differences in Table A3 other than those for the full models in Table 5 for predicting Calculus II and III course grades are well above 10, ranging from 21 to 2,568. This provides further evidence that ACI scores measure something beyond course grade.

#### *ACI Score as a Predictor of Differential Outcomes by URM Status and Gender*

Next, we considered whether ACI score could explain some of the observed gaps in math or STEM outcomes by URM status.

*Math course grades.* We begin by considering whether ACI scores mediate the relationship between URM status or gender and math grades. The results of these mediation models are reported in Table 7.

Table 7 reveals several trends. First, for URM status, the direct and indirect discrepancies are both highly significant in base and full models, suggesting that both ACI score and other factors correlated with URM status explain the lower average math course grades of URM students. ACI score accounts for 17% of the difference in math course grades by URM status in base models (proportion of total discrepancy equal to the indirect discrepancy, or  $-0.11 / -0.65$ ) and 13% in full models ( $-0.06 / -0.42$ ).

Mediation models for gender show different trends: while both the direct and indirect discrepancies are highly significant in both base and full models, the direction of the relationship is in the opposite direction. Women on average earn higher grades in math courses despite having lower ACI scores on average than men, so considering a mediation model that includes ACI scores reveals a suppressor “effect” where women on average earn even higher grades than men with equivalent ACI scores. Thus, ACI scores do not explain women’s higher math course grades, but rather women’s higher average math course grades mask that they have lower mean scores on the ACI.

*Successful completion of core STEM major math courses.* Next, we considered whether ACI scores mediate the relationship between URM status or gender and successful completion of commonly required math courses for STEM majors, including elementary algebra, intermediate/college

TABLE 8

*ACI Score as Mediator Between Underrepresented Minority Status (URM)<sup>a</sup> or Gender and Successful Completion of Various Core Math Courses in STEM Degree Requirements*

Base Models					Full Models <sup>b</sup>				
Mediation for Differential Outcomes by Underrepresented Minority Status									
Elementary Algebra	Coeff.	SE	p	sig.	Coeff.	SE	p	sig.	n
Total Discrepancy (Total "Effect")	-0.19	0.05	< .001	***	-0.18	0.08	.020	*	499
Direct Discrepancy (Direct "Effect")	-0.14	0.05	.011	*	-0.16	0.08	.039	*	
Indirect Discrepancy (Indirect "Effect")	-0.05	0.02	.002	**	-0.02	0.02	.329		
Intermediate/College Algebra	Coeff.	SE	p	sig.	Coeff.	SE	p	sig.	n
Total Discrepancy (Total "Effect")	-0.21	0.02	< .001	***	-0.11	0.03	< .001	***	2,597
Direct Discrepancy (Direct "Effect")	-0.17	0.02	< .001	***	-0.10	0.03	< .001	***	
Indirect Discrepancy (Indirect "Effect")	-0.04	0.01	< .001	***	-0.01	0.01	.043	*	
Precalculus	Coeff.	SE	p	sig.	Coeff.	SE	p	sig.	n
Total Discrepancy (Total "Effect")	-0.22	0.02	< .001	***	-0.14	0.03	< .001	***	2,418
Direct Discrepancy (Direct "Effect")	-0.18	0.02	< .001	***	-0.13	0.03	< .001	***	
Indirect Discrepancy (Indirect "Effect")	-0.04	0.01	< .001	***	-0.01	0.00	.025	*	
Calculus I	Coeff.	SE	p	sig.	Coeff.	SE	p	sig.	n
Total Discrepancy (Total "Effect")	-0.15	0.02	< .001	***	-0.08	0.03	.001	**	1,248
Direct Discrepancy (Direct "Effect")	-0.12	0.03	< .001	***	-0.07	0.03	.005	**	
Indirect Discrepancy (Indirect "Effect")	-0.02	0.01	< .001	***	-0.01	0.00	.026	*	
Calculus II	Coeff.	SE	p	sig.	Coeff.	SE	p	sig.	n
Total Discrepancy (Total "Effect")	-0.20	0.03	< .001	***	-0.18	0.04	< .001	***	611
Direct Discrepancy (Direct "Effect")	-0.19	0.04	< .001	***	-0.17	0.04	< .001	***	
Indirect Discrepancy (Indirect "Effect")	-0.02	0.01	.056	.	-0.01	0.01	.257		
Mediation For Differential Outcomes by Female Gender Status									
Elementary Algebra	Coeff.	SE	p	sig.	Coeff.	SE	p	sig.	n
Total Discrepancy (Total "Effect")	0.08	0.04	.093	.	0.05	0.06	.439		499
Direct Discrepancy (Direct "Effect")	0.07	0.04	.140		0.07	0.06	.271		
Indirect Discrepancy (Indirect "Effect")	0.01	0.01	.440		-0.02	0.01	.150		
Intermediate/College Algebra	Coeff.	SE	p	sig.	Coeff.	SE	p	sig.	n
Total Discrepancy (Total "Effect")	0.07	0.02	< .001	***	0.06	0.03	.026	*	2,597
Direct Discrepancy (Direct "Effect")	0.09	0.02	< .001	***	0.07	0.03	.006	**	
Indirect Discrepancy (Indirect "Effect")	-0.02	0.00	.002	**	-0.01	0.00	.004	**	
Precalculus	Coeff.	SE	p	sig.	Coeff.	SE	p	sig.	n
Total Discrepancy (Total "Effect")	0.05	0.02	.006	**	0.04	0.03	.119		2,418
Direct Discrepancy (Direct "Effect")	0.07	0.02	< .000	***	0.05	0.03	.044	*	
Indirect Discrepancy (Indirect "Effect")	-0.02	0.00	< .001	***	-0.01	0.00	.007	**	
Calculus I	Coeff.	SE	p	sig.	Coeff.	SE	p	sig.	n
Total Discrepancy (Total "Effect")	0.05	0.03	.055	.	0.03	0.02	.285		1,248
Direct Discrepancy (Direct "Effect")	0.08	0.03	.002	**	0.05	0.03	.068	.	
Indirect Discrepancy (Indirect "Effect")	-0.03	0.01	< .001	***	-0.02	0.01	< .001	***	
Calculus II	Coeff.	SE	p	sig.	Coeff.	SE	p	sig.	n
Total Discrepancy (Total "Effect")	0.04	0.04	.242	.	0.04	0.04	.271		611
Direct Discrepancy (Direct "Effect")	0.06	0.04	.097	.	0.04	0.04	.213		
Indirect Discrepancy (Indirect "Effect")	-0.02	0.01	.022	*	-0.01	0.00	.282		

*Note.* Linear probability model mediation using the KHB Method: Mediation of the relationship between underrepresented minority status or gender and successful completion of core math courses in the standard STEM major course sequence (for those students who enrolled) by ACI score, fixed-effects by course sequence number, clustering by instructor. <sup>a</sup>Underrepresented minority status indicates students of color who have been traditionally underrepresented and underserved in STEM fields, this includes Black, Hispanic, and American Indian/Native Alaskan students, but not White or Asian/Pacific Islander students. <sup>b</sup>Full models included control variables: GPA at start of term, number of prior credits earned by start of term, race/ethnicity (for gender models), gender (for URM models), age, and median household income of zipcode, instructor.

\*\*\*  $\alpha = 0.001$ , \*\*  $\alpha = 0.01$ , \*  $\alpha = 0.05$ , .  $\alpha = 0.10$ .

TABLE 9

*ACI Score as Mediator Between Underrepresented Minority Status (URM)<sup>a</sup> or Gender and STEM Versus Non-STEM Degree Completion*

Base Models					Full Models <sup>b</sup>			
	Coef.	SE	<i>p</i>	sig.	Coef.	SE	<i>p</i>	sig.
Underrepresented Minority Status								
Total Discrepancy (Total “Effect”)	−0.02	0.02	.442		−0.00	0.03	.928	
Direct Discrepancy (Direct “Effect”)	0.01	0.02	.812		0.01	0.03	.806	
Indirect Discrepancy (Indirect “Effect”)	−0.02	0.01	< .001	***	−0.01	0.00	.010	*
Female Gender								
Total Discrepancy (Total “Effect”)	−0.11	0.03	< .001	***	−0.08	0.03	.002	**
Direct Discrepancy (Direct “Effect”)	−0.10	0.03	< .001	***	−0.07	0.03	.005	**
Indirect Discrepancy (Indirect “Effect”)	−0.01	0.00	.008	**	−0.01	0.00	.008	**

*Note.* ACI = Algebra Concept Inventory; STEM = Science, Technology, Engineering, and Mathematics. linear probability model mediation using the KHB method: Mediation of the relationship between underrepresented minority status or gender and STEM versus non-STEM degree completion by ACI score, fixed-effects by course sequence number, clustering by instructor.

<sup>a</sup>Underrepresented minority status indicates students of color who have been traditionally underrepresented and underserved in STEM fields, this includes Black, Hispanic, and American Indian/Native Alaskan students, but not White or Asian/Pacific Islander students. <sup>b</sup>Full models included control variables: GPA at start of term, number of prior credits earned by start of term, race/ethnicity (for gender models), gender (for URM models), age, and median household income of zipcode, instructor.

\*\*\*  $\alpha = 0.001$ , \*\*  $\alpha = 0.01$ , \*  $\alpha = 0.05$ , .  $\alpha = 0.10$ .

algebra, precalculus, Calculus I, and Calculus II. The results of these mediation models are reported in Table 8.

For URM status, both direct and indirect discrepancies are highly significant in base and full models for all courses (except full models for elementary algebra). This suggests that both ACI score and other factors associated with URM status explain the lower probability of successfully completing courses above elementary algebra for URM students. ACI scores explain about 16%–18% of the difference for intermediate/college algebra through Calculus I in base models and 7%–11% in full models (i.e., the proportion of total discrepancy equal to the indirect discrepancy in each model).

For gender, these models again suggest a suppressor “effect” (significant for all courses except elementary algebra and Calculus II full models), where women are, on average, more likely to successfully complete core math course requirements for STEM majors, despite scoring lower on average on the ACI than men. For some models in Table 8 (e.g., for some precalculus, Calculus I and II models), the total discrepancy is not statistically significant, but the direct and indirect discrepancies are, yet in opposite directions. Thus, these relationships stay hidden when looking only at average effects, when we do not consider the mediating role of ACI score.

*STEM versus non-STEM degree completion.* Finally, we considered whether ACI scores mediate the relationship between race/ethnicity or gender and STEM versus non-STEM degree completion. Results are reported in Table 9.

For URM status, neither the total nor the direct discrepancy (Table 9) is significant in base or full models, while the indirect discrepancy is. In fact, the direction of the direct

discrepancy has reversed sign—thus, after accounting for the indirect relationship of ACI score as an explanatory variable for differences in STEM versus non-STEM major attainment by URM status, URM students are *more* likely to complete a STEM versus non-STEM degree (although not statistically significantly so). Thus, ACI score explains 100% of the total discrepancy in STEM versus non-STEM degree attainment by URM status (although absolute differences are small, at about 2 percentage points).

Exploring mediation models for gender shows a different trend: the total, direct, and indirect discrepancies are all statistically significant, with ACI score explaining about 10% of the total discrepancy in both base and full models. Thus, about 10% of differences in STEM versus non-STEM degree attainment between men and women can be explained by women’s lower average scores on the ACI, but the remaining 90% must be accounted for by other factors correlated with gender.

### Limitations

While a wide range of students were represented in the data, students who completed the ACI did differ somewhat on average from the total population of students who were enrolled in math courses. However, data in this study still included good representation of all racial/ethnic and gender groups, as well as students across all GPA bands and course levels, and, for all analysis, weighting was used to adjust for all observed differences between those with an ACI score and the general population of students enrolled in math courses, with excellent balance achieved (Table 3). Further, control variables were used across all research questions to

allow for comparison of students with the same values on these characteristics.

This dataset also has some limitations in how gender and race/ethnicity were recorded. CUNY institutional data had at the time only a binary category for gender and used limited federal race/ethnicity categories; therefore, these variables may not accurately represent how all students self-identify. Further research that includes more nuanced race/ethnicity and gender categories is necessary. However, in the meantime, this study may provide a starting point for understanding how certain politically and socially constructed categories can relate to structural marginalization and discrimination (Balestra & Fleischer, 2018) in higher education.

The population at CUNY, where this study was conducted, is more diverse than the average U.S. college. It includes a higher proportion of ethnic/racial minorities, foreign-born students, first-generation students, lower-SES students, and students requiring developmental coursework. We note that we did not directly analyze or control for all of these groups in this particular study (e.g., we do not include models here that consider how foreign-born or first-generation status relate to ACI scores)—considering the relationship between ACI scores and college outcomes for other groups of students (beyond gender and race/ethnicity) is an important area for future research. Further, because of CUNY's diversity, we cannot claim that these data are nationally representative. However, these results are likely generalizable to a wider population nationally, and will be of particular interest to colleges with diverse populations. Further, this makes CUNY an excellent context for investigating the relationship between algebraic conceptual understanding and college math/STEM outcomes for groups traditionally underrepresented and underserved in higher education.

## Discussion

ACI scores significantly predicted math course grade, both overall, and for almost every individual course that was analyzed in this study. For courses that depend heavily on algebra (e.g., elementary algebra, intermediate algebra, precalculus) this result was not surprising. However, the strong predictive power of ACI scores for grades in courses that do not explicitly focus on algebra, such as statistics, Calculus I, and courses for future primary teachers and medical professionals, illustrates how important algebraic concepts may be to a range of mathematical domains across different STEM fields.

We note that the content of the ACI was generated to be accessible to any student who had enrolled in Algebra I in high school or elementary algebra in college and so it did not include algebraic concepts or objects (e.g., function notation, exponential growth, trigonometric functions) that might be more relevant to higher-level mathematics courses. Thus, the observed pattern in which ACI scores were more

predictive for lower-level courses (e.g., intermediate/college algebra and precalculus) than some higher-level courses (e.g., Calculus II and III) provides validity evidence for the intended interpretation of the ACI. An expanded instrument that includes concepts beyond those relevant to elementary algebra may be more predictive of grades in upper-level math courses. However, the results found here show that algebraic conceptions tested by the ACI significantly predict course grades in a wide variety of mathematics courses that are important to both STEM and non-STEM majors. This suggests that a more explicit focus in instruction at all levels on conceptual understanding of foundational classical algebra may be a critical component of preparing students to succeed in college math courses and the degrees that require these courses.

ACI scores also significantly predicted the likelihood of ever subsequently completing each of the courses on the standard STEM major math course sequence from elementary algebra to Calculus II, independent of the course in which a student was currently enrolled. This suggests that algebraic conceptual understanding as measured by the ACI is important to student completion of many different required math courses for STEM majors, even though the ACI does not include algebra concepts introduced in courses beyond elementary algebra. Findings point to the importance of algebra courses including an explicit focus on conceptual understanding if they are to prepare students to be successful in the college math courses needed for STEM degrees.

ACI scores additionally predicted the likelihood of a student completing a STEM versus non-STEM major during the study period, independent of the math course in which they were currently enrolled. It was particularly noteworthy that ACI scores predicted STEM versus non-STEM degree completion separately but equally as well as math course grade. This suggests that ACI scores capture critical information about students' likelihood of completing STEM degrees that is not captured by existing math course grades. Perhaps existing course grades depend heavily on accurate procedural fluency and less on conceptual understanding, and/or perhaps they measure conceptual understanding on concepts outside foundational classical algebra. Regardless of the reason for the difference, this study suggests that it is critical to better understand how to provide every student interested in STEM ample opportunities to cultivate the kinds of foundational classical algebraic conceptual understanding measured by the ACI. This is likely an important task throughout the K–16 mathematics curriculum, if we want every student to have access to STEM degrees.

Further, our results suggest that instruction in algebraic conceptual understanding is a critical equity issue. ACI score explained significant differences in math course grades and STEM versus non-STEM degree completion by URM status, suggesting that differences in opportunities to develop algebraic conceptual understanding explain a significant proportion

of college STEM outcome differences by race/ethnicity. It is already known that URM status correlates with differential access to high-quality mathematics instruction, including opportunities for reasoning and sensemaking (Schoenfeld, 2022; Stepter, 2023; Yeh et al., 2020). URM students may also have more marginalized experiences in mathematics classes (e.g., McGee, 2015; Ridgeway & McGee, 2018), which may reduce their opportunities to develop conceptual understanding. Thus, improving access for all students to meaningful opportunities to develop algebraic conceptual understanding in mathematics courses at all levels may be particularly crucial to addressing inequities in college STEM outcomes—this is a critical area for future research.

For gender, analysis revealed a suppressor effect, where women on average earned higher grades in math courses (and were more likely to successfully complete courses in the STEM major math sequence) despite having lower ACI scores. This may be because women on average spend more time on their studies than comparable men (Conway et al., 2021; Wladis et al., 2024a, 2024b), and therefore may earn higher grades because of the added time spent studying. Or it may relate to gendered norms that reinforce “compliance” for women and “risk-taking” for men during problem-solving (Lubienski & Ganley, 2017; Lubienski et al., 2021; Miller et al., 1996). When instruction stresses procedures (as some research suggests is common in algebra courses in college; Crooks & Alibali, 2014; Hammerman & Goldberg, 2003; Hodara, 2011), compliance may correspond to learning rote procedures that can then lead to higher grades but limit opportunities for developing conceptual understanding, whereas risk-taking may correspond to pursuing alternative problem-solving routes that are more likely to lead to increased conceptual understanding. More research is needed to explore these hypotheses and to better understand this trend.

In this study, differences in ACI scores also explained 10% of the gender gap in STEM versus non-STEM degree completion. Thus, it may be important to ensure that women and girls receive equal opportunities to develop algebraic conceptual understanding in mathematics courses. As with course grades, this may be related to marginalized experiences that women and girls have in mathematics classes (e.g., Reinholz et al., 2022), which may negatively impact their opportunities to develop conceptual understanding; or it might be related to patterns of “compliance” versus “risk-taking” in mathematics classrooms, where women and girls have been observed to be more “compliant” (Lubienski & Ganley, 2017; Miller et al., 1996) and to take fewer “risks” in mathematics problem-solving (Lubienski et al., 2021), which might reduce their opportunities to develop conceptual understanding.

### Implications

Because the ACI is based on a detailed framework describing specific concepts and conceptions (or ways of

thinking about those concepts) that constitute foundational algebraic conceptual understanding (Wladis et al., 2024a, 2024b; Wladis et al., manuscript under review), the ACI framework could be leveraged to inform instruction and curricula. Because ACI scores predict math and STEM outcomes in college, it may make sense to design curricula and instructional activities that focus on the particular concepts and conceptions that the ACI assesses. Thus, ACI sample items and details of the ACI framework could be an important tool for instructors, course and curriculum designers, and researchers who study undergraduate STEM education.

The fact that ACI scores predict STEM degree outcomes separately and significantly from course grade suggests that the ACI measures something that is distinct from assessments currently used in college mathematics courses but nonetheless important to STEM degree completion. This suggests that knowledge assessed by the ACI represents important algebra knowledge that is complementary to, rather than overlapping with, the knowledge measured by current course assessments. Because this conceptual knowledge appears not to be widely assessed in current college math courses, it may be that it is also not widely taught in the formal and enacted mathematics curriculum (Faulkner & Cook, 2006; Jennings & Jonathan, 2014; Sambell & McDowell, 1998). This might help to explain why mismatch has been identified between the algebra skills taught in the classroom and those perceived by workers as relevant on the job (Douglas & Attewell, 2017; Handel, 2016), even in STEM careers (Walkington et al., 2025). Research suggests that college graduates, even in STEM fields like engineering, often find more general skills like learning “how to think” (Moss-Pech, 2025, p. 117, 120), “problem-solving” (Moss-Pech, 2025, p. 120), or using the “scientific method” (Moss-Pech, 2025, p. 120) that they obtain from their college courses more relevant than narrower technical skills that they learn in their classes; thus, it is possible that algebraic conceptual understanding (which focuses on “problem-solving” and reasoning) may be particularly important in the STEM workforce. More research is needed to explore this possibility.

Relatedly, this study provides evidence that could be leveraged in ongoing developmental mathematics and college mathematics course sequence reform. These reforms have typically focused on redirecting students away from algebra into other courses (e.g., Ganga & Mazzariello, 2018; Logue et al., 2019), on compressing developmental and credit-bearing courses into a single combined course (Buckles et al., 2019; Kosiewicz et al., 2016; Merkin, 2023), or other interventions that typically keep the curriculum intact but provide students with alternate pacing or additional supports (Kalamkarian et al., 2015; Park et al., 2018; Twigg, 2011; The Century Foundation, 2016). These reforms may help with some of the barriers that students experience on their way to a degree, for example by allowing them to complete

the degree in fewer terms when they are able to be successful in accelerated or compressed courses. However, these reforms have not yet critically considered what kinds of algebraic knowledge students acquire in these courses, which instruction in these courses is providing students with rich opportunities to develop algebraic conceptual understanding, or whether these courses are fulfilling their original intention of providing access to STEM careers (Wladis, Bjorkman, et al., 2023; Wladis, Makowski, et al., 2023). Evidence from this study suggests that developmental mathematics and college math sequence reform should consider what kinds of opportunities students are getting in these classes to develop algebraic conceptual understanding, and should consider including instructional activities that attend specifically to the concepts and conceptions on which the ACI was built.

Further, the mediation analysis results for race/ethnicity and gender in this research show that access to opportunities to develop algebraic conceptual understanding predicts differential math course and STEM degree outcomes and is therefore a critical equity issue. Research that investigates the relationship between equitable learning opportunities in the classroom and subsequent development or improvement of conceptual understanding could be an important area for future research. In the meantime, instructional approaches that focus more explicitly on teaching algebraic conceptual understanding may have the potential to improve both college STEM outcomes and STEM equity, as long as they are implemented in such a way that every student has sufficient opportunity to engage in rich sense-making activities that support the development of algebraic conceptual understanding.

### Conclusion

The results of this study suggest that scores on the ACI predict math course grades, successful completion of core math courses typically required for STEM majors, and STEM versus non-STEM degree attainment. They also explain some differential outcomes in these measures by race/ethnicity and gender. An increase of one standard deviation in ACI score was associated with one-third to one-half letter grade increase in math course grade for students in elementary/intermediate/college algebra, precalculus, and Calculus I; with a roughly 8–16 percentage point increase in the probability of ever completing core math courses required for most STEM majors, such as elementary/intermediate/college algebra, pre-calculus, and Calculus I/II; and a 5–6 percentage point increase in the probability of ever completing a STEM versus non-STEM degree, depending on which courses are being compared and which controls are included. These relationships between ACI score and various future math and STEM

outcomes in college provide quantitative evidence to support existing descriptions of the importance of integrating instruction on algebraic conceptual understanding into both the K–12 and higher education mathematics curriculum. This suggests that a more explicit focus on conceptual understanding of foundational classical algebra may be a critical component of preparing students to succeed in college math courses and the degrees that require these courses, across many different majors.

ACI scores also explained significant portions of existing gaps in math course grades, course completion and STEM versus non-STEM completion by race/ethnicity and gender. ACI score accounted for 13%–17% of URM math grade differences, 100% of the STEM versus non-STEM degree completion gap by URM status, and 10% of the degree completion gap by gender. This suggests that access to high quality instruction in algebraic conceptual understanding is a critical equity issue, and that differences in opportunities to develop algebraic conceptual understanding explain a significant proportion of certain college STEM outcome differences by race/ethnicity and gender.

The fact that that ACI score was predictive of STEM degree completion in addition to and independently of math course grade suggests that (1) existing course assessment practices are likely not robustly capturing students' foundational algebraic conceptual understanding; and (2) existing curricula and instruction as currently enacted may not be focusing explicitly on foundational algebraic conceptual understanding. Given the predictive power of ACI scores for college math and STEM outcomes, this suggests that it may be critical to improve curricula and instructional practice so that students in college and college-preparatory mathematics courses have sufficient opportunities to develop robust foundational algebraic conceptual understanding. The framework used to generate the ACI could be used as a starting point for identifying the components of algebraic conceptual understanding that may be worth attending to in instruction.

This study provides some of the first large-scale quantitative evidence that algebraic conceptual understanding can predict college mathematics course outcomes and STEM degree attainment. In addition to informing potential instructional and curricular changes, it points towards a critical need for future research that (1) explores which instructional approaches improve students' algebraic conceptual understanding; (2) determines which students are provided access to instruction that supports the development of algebraic conceptual understanding; and (3) directly tests whether improved instruction in algebraic conceptual understanding can positively impact STEM degree attainment and math course completion. Better understanding these patterns appears to be critical not just to building a well-qualified STEM workforce, but also an equitable one.

## Appendix: Robustness Checks

TABLE A0

*Models Considering Moderation of Gender in the Relationship Between ACI Score and Outcomes*

---

Full Models From Table 4

---

Fixed Effects by Course Sequence Number, Regression Results

---

Model	Coefficient for Gender $\times$ ACI Score*	SE	<i>p</i>
Sequence	-0.05	0.05	.33

---

Separate Models for Each Course Type, Regression Results

---

Model	Coefficient for Gender $\times$ ACI Score*	SE	<i>p</i>
Intermediate/College Algebra	0.00	0.11	.99
Precalculus	-0.04	0.10	.68
Calculus I	0.13	0.15	.38
Calculus II	0.11	0.10	.25
Calculus III	-0.19	0.16	.26
Differential Equations	0.34	0.33	.29
Mathematics Education	-0.13	0.30	.65
Statistics	-0.04	0.09	.61
Quantitative Reasoning	-0.03	0.21	.87
Math For Health Majors	0.21	0.24	.37

---

Full Models From Table 5

---

Model	Coefficient for Gender $\times$ ACI Score*	SE	<i>p</i>
Elementary Algebra	0.04	0.06	.47
Intermediate/College Algebra	0.02	0.03	.58
Precalculus	0.00	0.03	.94
Calculus I	-0.01	0.05	.91
Calculus II	0.06	0.03	.05

---

Full Models From Table 6

---

Model	Coefficient for Gender $\times$ ACI Score*	SE	<i>p</i>
Without Course Grade	0.00	0.02	.80
With Course Grade	-0.02	0.02	.31

---

*Note.* ACI = Algebra Concept Inventory.

\*Coefficients are for the interaction term gender  $\times$  ACI score in each model.

TABLE A1

*Robustness Check: Comparing Full Model Results With and Without ELL Status, Example Using Table 4*

Full Models <sup>a</sup> (From Table 4)					Full Models <sup>a</sup> , ELL Measure Added				
Fixed Effects by Course Sequence Number, Regression Results									
	Coeff.	SE	<i>p</i>	sig.	Coeff.	SE	<i>p</i>	sig.	<i>n</i>
ACI Score	0.26	0.02	< .001	***	0.26	0.02	< .001	***	7,908
Sequence	0.05	0.03	.066	.	0.05	0.03	.072	.	
Separate Models for Each Course Type, Regression Results									
	Coeff.	SE	<i>p</i>		Coeff.	SE	<i>p</i>	sig.	<i>n</i>
Intermediate/College Algebra	0.40	0.04	< .001	***	0.40	0.04	< .001	***	1,960
Precalculus	0.36	0.04	< .001	***	0.36	0.04	< .001	***	2,167
Calculus I	0.24	0.05	< .001	***	0.24	0.05	< .001	***	1,020
Calculus II	0.06	0.11	.600		0.04	0.11	.712		495
Calculus III	0.01	0.08	.951		0.00	0.08	.961		201
Differential Equations	0.90	0.33	.006	**	0.87	0.32	.007	**	68
Mathematics Education	0.23	0.06	< .001	***	0.23	0.05	< .001	***	337
Statistics	0.20	0.05	< .001	***	0.20	0.05	< .001	***	2,166
Quantitative Reasoning	0.18	0.09	.048	*	0.18	0.09	.042	*	625
Math for Health Majors	0.18	0.07	.009	**	0.18	0.07	.009	**	606

*Note.* ACI = Algebra Concept Inventory; ELL = English language learner. Multi-Level Weighted Regression, fixed effects by course (first row) or separate models by courses (listed in first column), random effects by instructor, math course grade as dependent variable, ACI score as independent variable for which coefficients are reported.

<sup>a</sup>Full models include control variables: GPA at start of term, number of prior credits earned by start of term, race/ethnicity, gender, age, and median household income of zipcode.

\*\*\*  $\alpha = 0.001$ , \*\*  $\alpha = 0.01$ , \*  $\alpha = 0.05$ , .  $\alpha = 0.10$ .

TABLE A2

*Robustness Check: Comparing Results of Table 5 with Analysis That Excludes Students Currently Enrolled in the Course*

Base Models					Full Models <sup>a</sup>				
Ever Completed:	Coefficient ACI Score	SE	<i>p</i>	sig.	Coefficient ACI Score	SE	<i>p</i>	sig.	<i>n</i>
Intermediate/College Algebra, Including Currently Enrolled (From Table 5)	0.16	0.04	< .001	***	0.15	0.05	.001	**	499
Intermediate/College Algebra, not Currently Enrolled	0.19	0.10	.055	.	0.18	0.09	.033	*	193
Precalculus, Including Currently Enrolled (From Table 5)	0.14	0.02	< .001	***	0.11	0.02	< .001	***	2,597
Precalculus, not Currently Enrolled	0.14	0.02	< .001	***	0.10	0.02	< .001	***	2,163

*Note.* ACI = Algebra Concept Inventory. It is impossible to include elementary algebra in this robustness check because students below that course were not included in the study (i.e., students placed into arithmetic or not currently enrolled in math). Calculus courses were not included because of small *n* (i.e., less than 60). Linear probability models using multi-level weighted regression: Fixed-effects by course sequence number of current course, random effects by instructor, ever completed (if attempted) particular math course (listed in first column) as dependent variable, ACI score as independent variable for which coefficients are reported.

<sup>a</sup>Full models include control variables: GPA at start of term, number of prior credits earned by start of term, race/ethnicity, gender, age, and median household income of zipcode.

\*\*\*  $\alpha = 0.001$ , \*\*  $\alpha = 0.01$ , \*  $\alpha = 0.05$ , .  $\alpha = 0.10$ .

TABLE A3

Incremental Validity Analysis Robustness Check: Model Fit with and without ACI Score

With ACI Score		Without ACI Score						
Models From Table 4								
Fixed Effects by Course Sequence Number, Regression Results								
	AIC	BIC	AIC	BIC	diff AIC	diff BIC	<i>t</i> for BIC diff. <sup>a</sup>	<i>p</i> <sup>b</sup>
Sequence	256,489	256,622	259,065	259,190	2,575	2,568	50.7	<b>&lt;.001</b>
Separate Models for Each Course Type, Regression Results								
	AIC	BIC	AIC	BIC	diff AIC	diff BIC	<i>t</i> for BIC diff. <sup>a</sup>	<i>p</i> <sup>b</sup>
Intermediate/College Algebra	65,209	65,304	66,805	66,894	1,596	1,590	39.9	<b>&lt;.001</b>
Precalculus	70,999	71,095	72,536	72,627	1,538	1,532	39.2	<b>&lt;.001</b>
Calculus I	27,507	27,590	27,757	27,836	251	246	15.8	<b>&lt;.001</b>
Calculus II	12,774	12,841	12,780	12,843	6	2	2.1	.016
Calculus III	3,057	3,109	3,055	3,104	-2	-5	-2.7	.997
Differential Equations	1,215	1,248	1,483	1,514	268	266	16.4	<b>&lt;.001</b>
Mathematics Education	9,564	9,625	9,656	9,713	92	88	9.5	<b>&lt;.001</b>
Statistics	76,037	76,139	76,441	76,538	404	399	20.1	<b>&lt;.001</b>
Quantitative Reasoning	22,827	22,903	22,921	22,992	94	89	9.6	<b>&lt;.001</b>
Math for Health Majors	15,891	15,966	15,982	16,053	91	86	9.4	<b>&lt;.001</b>
Models From Table 5								
Ever Completed:	AIC	BIC	AIC	BIC	diff AIC	diff BIC	<i>t</i> for BIC diff. <sup>a</sup>	<i>p</i> <sup>b</sup>
Elementary Algebra	6,486	6,562	6,858	6,930	372	368	19.2	<b>&lt;.001</b>
Intermediate/College Algebra	26,459	26,565	27,727	27,827	1,268	1,262	35.6	<b>&lt;.001</b>
Precalculus	22,018	22,122	22,917	23,015	899	893	29.9	<b>&lt;.001</b>
Calculus I	9,508	9,600	9,876	9,963	368	363	19.1	<b>&lt;.001</b>
Calculus II	3,351	3,426	3,486	3,557	136	131	11.6	<b>&lt;.001</b>
Models from Table 6								
Ever Attained STEM Versus Non-STEM Degree	AIC	BIC	AIC	BIC	diff AIC	diff BIC	<i>t</i> for BIC diff. <sup>a</sup>	<i>p</i> <sup>b</sup>
ACI Score	2,728	2,830	2,755	2,851	27	21	5.0	<b>&lt;.001</b>

Note. <sup>a</sup>Criteria for *t*-statistic calculation and significance are taken from (Raftery, 1995, pp. 139–140). <sup>b</sup>Bolded *p* values indicate “strong evidence” (Raftery, 1995, p. 139) for added information due to additional variable (Akaike Information Criterion (ACI) score in this case), based on differences in Bayesian information criterion (BIC) greater than 10.

### Authors' Note

Geillan Aly is now the owner of Compassionate Math, LLC: Professional Consulting and Development in Math Education, 860-255-8783, [compassionatemath@gmail.com](mailto:compassionatemath@gmail.com).

### Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article:


This research was supported by grants from the National Science Foundation (#1760491, 2300725). Opinions reflect those of the authors and do not necessarily reflect those of the granting agency.




### Open Practices

The analysis files and information about obtaining the data for this article can be found at <https://doi.org/10.3886/E239171V1>.

### ORCID iDs

Claire Wladis  <https://orcid.org/0000-0003-3904-6113>

Benjamin Sencindiver  <https://orcid.org/0000-0001-6293-582X>

Alyse C. Hachey  <https://orcid.org/0000-0001-6503-7283>  
Nils Myszkowski  <https://orcid.org/0000-0003-1322-0777>  
Jason Samuels  <https://orcid.org/0000-0002-1600-5856>

### Notes

1. We note that we aim to describe an existing conceptualization used in the literature and do not endorse it. For example, students may be placed into developmental algebra because of flawed placement procedures, delayed college entry, or low K–12 mathematics instructional quality, all of which can be independent of student ability or interest in algebra, math or STEM.

2. There are likely many knowledge types that could be listed separately in the cognitive impacts box and studied in other research; however, because of the focus of this study we have streamlined the figure to highlight algebraic conceptual understanding versus other types of knowledge. This is not intended to imply that conceptual understanding is unrelated to other knowledge types or that other knowledge types all belong in a single category.

3. “Elementary” algebra refers to a first algebra course in college (typically classified by colleges as “developmental” and offered not-for-credit) that assumes no pre-requisite knowledge of algebra, although students who enroll have almost always already successfully completed Algebra I in high school.

4. Some have used the term “school” algebra, but we prefer the term “classical” algebra for reasons of clarity and equity. “School” algebra is a misnomer outside the K–12 setting. Students in algebra classes in college are adults, who are developmentally different from eighth/ninth graders and have typically already successfully completed Algebra I in secondary school. They may sometimes be asked to engage with similar mathematical objects as Algebra I in secondary school, but in college are often expected to do so in ways qualitatively different from younger students. For a more in-depth discussion of this, see (Wladis, Bjorkman, et al., 2023; Wladis, Makowski, et al., 2023).

5. For the purposes of test security, numbers and answer option order have been altered.

6. This is slightly less than the total number of students who took the BAI during the validation study because IR data was provided before data collection in the validation study was complete, and thus students who took the BAI after IR data was obtained were not included in this analysis.

7. A total of 89.7% of the missing zipcode cases belonged to students with at least one foreign address on file, whereas only 3.4% of students had at least one foreign address on file. Therefore, students with missing zipcode data were significantly more likely to be foreign students.

8. We also considered potential moderating relationships by gender for all regression models reported here; however, interaction terms (for gender  $\times$  ACI score) were not significant at the  $\alpha = 0.05$  level in any models, and so we did not include interaction terms in the mediation models. See Table A0 in the Appendix for results of moderation analysis.

9. We use the term “predict” to indicate that the BAI was administered to students before course grades were assigned and before students completed their degree; this is not intended to imply that every assessment that influenced students’ final course grades (and that in turn impacted their degree completion) occurred after the

BAI was administered. As noted in the method section, 90% of students took the ACI at least 1 month prior to final exam administration and none took it during the week of final exams and thus the ACI was taken temporally prior to final examinations and largely prior to various other coursework that determined final course grades. We note that we do not see any differences in overall predictive patterns if we only consider whether BAI scores predict course outcomes in a subsequent term compared to both current and subsequent terms (see, for example, Table A2 in the Appendix).

10. We were unable to include elementary algebra courses in this analysis due to small  $n$ , because very few elementary algebra courses had course grades beyond pass/fail designations during the study period.

11. As a robustness check, we also ran this analysis while only including students who were not currently enrolled in the course at the time of taking the ACI (see Table A2 in the Appendix); this substantially reduced sample size, reducing the magnitude of the significance for intermediate/college algebra, but produced even larger coefficients (suggesting a potentially stronger rather than weaker relationship between ACI score and course grade) than those reported in Table 5 and virtually identical results (coefficients, standard errors, and  $p$  values) to those for precalculus in Table 5. See Table A2 in the Appendix for details.

### References

- Ackerman, P. L., Kanfer, R., & Calderwood, C. (2013). High school advanced placement and student performance in college: STEM majors, non-STEM majors, and gender differences. *Teachers College Record*, *115*(10), 1–43. <https://doi.org/10.1177/016146811311501003>
- Adelman, C. (2006). *The toolbox revisited: Paths to degree completion from high school through college* (No. ED490195, p. 223). U.S. Department of Education. <https://eric.ed.gov/?id=ED490195>
- Alkhasawneh, R., & Hargraves, R. H. (2014). Developing a hybrid model to predict student first year retention in STEM disciplines using machine learning techniques. *LITEE*, *15*(3), 35–42. <https://doi.org/10.25772/6XXJ-7Q43>
- Al-Mutawah, M. A., Thomas, R., Eid, A., Mahmoud, E. Y., & Fateel, M. J. (2019). Conceptual understanding, procedural knowledge and problem-solving skills in mathematics: High school graduates work analysis and standpoints. *International Journal of Education and Practice*, *7*(3), 258–273. <https://doi.org/10.18488/journal.61.2019.73.258.273>
- Aly, G. (2022). Benny, Barbara, and the ethics of EdTech. *Journal of Humanistic Mathematics*, *12*(2), 98–127. <https://doi.org/10.5642/jhummath.HANB5943>
- Anderson, E., & Kim, D. (2006). *Increasing the success of minority students in science and technology* (No. 310736). American Council on Education. <http://hdl.handle.net/10919/86900>
- Bailey, T., Jeong, D. W., & Cho, S. (2010). Referral, enrollment, and completion in developmental education sequences in community colleges. *Economics of Education Review*, *29*(2), 255–270. <https://doi.org/10.1016/j.econedurev.2009.09.002>
- Balestra, C., & Fleischer, L. (2018). *Diversity statistics in the OECD: How do OECD countries collect data on ethnic, racial and indigenous identity?* (OECD Statistics Working

- Papers No. Vol. 2018/09). OECD Publishing. <https://doi.org/10.1787/89bae654-en>
- Berkowitz, M., & Stern, E. (2018). Which cognitive abilities make the difference? Predicting academic achievements in advanced STEM studies. *Journal of Intelligence*, 6(4), 48. <https://doi.org/10.3390/jintelligence6040048>
- Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord & M. R. Novick (Eds.), *Statistical theories of mental test scores* (pp. 397–479). Addison-Wesley. [https://faculty.ucmerced.edu/jvevea/classes/290\\_21/readings/week%209/Birnbaum.pdf](https://faculty.ucmerced.edu/jvevea/classes/290_21/readings/week%209/Birnbaum.pdf)
- Black, S. E., Muller, C., Spitz-Oener, A., He, Z., Hung, K., & Warren, J. R. (2021). The importance of STEM: High school knowledge, skills and occupations in an era of growing inequality. *Research Policy*, 50(7), Article 104249. <https://doi.org/10.1016/j.respol.2021.104249>
- Blanton Otto, K. (2018). *Transferability of algebra language and skills to algebra-based physics; Student and instructor perception and student skills*. Oklahoma State University. <https://www.proquest.com/openview/7872ff03574faba42aaf06eaf9deff6b/1?pq-origsite=gscholar&cbl=18750&diss=y>
- Boyce, S., & O'Halloran, J. (2020). Active learning in computer-based college algebra. *Problems, Resources, and Issues in Mathematics Undergraduate Studies*, 30(4), 458–474. <https://doi.org/10.1080/10511970.2019.1608487>
- Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., Stevens, M. T., & Welsh, M. E. (2016). A century of grading research: Meaning and value in the most common educational measure. *Review of Educational Research*, 86(4), 803–848. <https://doi.org/10.3102/0034654316672069>
- Buckles, E. L., Haydel, N. W., Thompson-Sanchez, J., & Page, Y. W. (2019). Implementing a corequisite algebra gateway course. *Association of American Colleges and Universities, Peer Review*, 21(1–2), 42–46. <https://www.aacu.org/peerreview/2019/winter-spring/Buckles>
- Buis, M. (2010). Stata tip 87: Interpretation of interactions in non-linear models. *Stata Journal*, 10(2), 305–308. <https://doi.org/10.1177/1536867X1001000211>
- Chalmers, R. P. (2012). Mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software*, 48(6), 1–29. <https://doi.org/10.18637/jss.v048.i06>
- Charleston, L. J., Adserias, R. P., Lang, N., & Jackson, J. F. (2014). Intersectionality and STEM: The role of race and gender in the academic pursuits of African American women in STEM. *Journal of Progressive Policy & Practice*, 2(3), 273–293. [https://www.researchgate.net/publication/280157009\\_Intersectionality\\_and\\_STEM\\_The\\_Role\\_of\\_Race\\_and\\_Gender\\_in\\_the\\_Academic\\_Pursuits\\_of\\_African\\_American\\_Women\\_in\\_STEM](https://www.researchgate.net/publication/280157009_Intersectionality_and_STEM_The_Role_of_Race_and_Gender_in_the_Academic_Pursuits_of_African_American_Women_in_STEM)
- Civil Rights Data Collection. (2023). *A first look: Students' access to educational opportunities in U.S. Public Schools. 2020-21 Civil Rights Data Collection*. U.S. Department of Education-Office of Civil Rights. <https://files.eric.ed.gov/fulltext/ED634099.pdf>
- Chen, X., & Simone, S. (2016). *Remedial Coursetaking at U.S. Public 2- and 4-Year Institutions: Scope, Experiences, and Outcomes. Statistical Analysis Report. NCES 2016-405* (Statistical No. NCIS 2017-405, pp. iii-f8). National Center for Education Statistics. <https://nces.ed.gov/pubs2016/2016405.pdf>
- Chen, X., & Weko, T. (2009). *Students who study science, technology, engineering, and mathematics (STEM) in postsecondary education* (No. Stats in Brief. NCES 2009-161; p. 25). U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. <https://eric.ed.gov/?id=ED506035>
- Cohen, R., & Kelly, A. M. (2019). Mathematics as a factor in community college STEM performance, persistence, and degree attainment. *Journal of Research in Science Teaching*, 57(2), 279–307. <https://doi.org/10.1002/tea.21594>
- Coltharp, J. (2020). Instruction type, student major, and success in developmental mathematics. *Journal of Developmental Education*, 44(10), 713–728. <https://files.eric.ed.gov/fulltext/EJ1320656.pdf>
- Conway, K. M., Wladis, C., & Hachey, A. C. (2021). Time poverty and parenthood: Who has time for college? *AERA Open*. Advance online publication. <https://doi.org/10.1177/23328584211011608>
- Cooke, R. L. (2008). *Classical algebra: Its nature, origins, and uses*. Wiley-Interscience. <https://doi.org/10.1002/9780470277980>
- Crisp, G., Nora, A., & Taggart, A. (2009). Student characteristics, pre-college, college, and environmental factors as predictors of majoring in and earning a STEM degree: An analysis of students attending a Hispanic serving institution. *American Educational Research Journal*, 46(4), 924–942. <https://doi.org/10.3102/0002831209349460>
- Crooks, N. M., & Alibali, M. W. (2014). Defining and measuring conceptual knowledge in mathematics. *Developmental Review*, 34(4), 344–377. <https://doi.org/10.1016/j.dr.2014.10.001>
- Douglas, D., & Attewell, P. (2017). School Mathematics as Gatekeeper. *The Sociological Quarterly*, 58(4), 648–669. <https://doi.org/10.1080/00380253.2017.1354733>
- Ellington, R. M. (2006). *Having their say: Eight high -achieving African -American undergraduate mathematics majors discuss their success and persistence in mathematics*. University of Maryland, College Park. <https://www.proquest.com/docview/305304604?pq-origsite=gscholar&fromopenview=true&sourcetype=Dissertations%20&%20Theses>
- Ellis, J., Fosdick, B. K., & Rasmussen, C. (2016). Women 1.5 times more likely to leave STEM pipeline after calculus compared to men: Lack of mathematical confidence a potential culprit. *PLoS One*, 11(7), Article e0157447. <https://doi.org/10.1371/journal.pone.0157447>
- Emblom-Callahan, M., Burgess-Palm, N., Davis, S., Decker, A., Diritto, H., Dix, S., Parker, C., & Styles, W. (2019). Accelerating student success: The case for corequisite instruction. *The Journal of the Virginia Community Colleges*, 22(1), Article 12. <https://commons.vccs.edu/inquiry/vol22/iss1/12>
- Erikson, R., Goldthorpe, J. H., Jackson, M., Yaish, M., & Cox, D. R. (2005). On class differentials in educational attainment. *Proceedings of the National Academy of Science (PNAS) of the USA*, 102, 9730–9733. <https://doi.org/10.1073/pnas.0502433102>
- Fagan, M. A., Hull, D. M., Gray, R., & Bolen, J. A. (2019). Predicting STEM performance in a Hispanic serving institution. *Personality and Individual Differences*, 141, 18–24. <https://doi.org/10.1016/j.paid.2018.12.017>
- Faulkner, S. A., & Cook, C. M. (2006). Testing vs. teaching: The perceived impact of assessment demands on middle grades instructional practices. *Research in Middle Level Education Online*, 29(7), 1–13. <https://doi.org/10.1080/19404476.2006.1462030>

- Frank, K., & Thompson, P. W. (2021). School students' preparation for calculus in the United States. *ZDM—Mathematics Education*, 53(3), 549–562. <https://doi.org/10.1007/s11858-021-01231-8>
- Freeman, R. B., & Huang, W. (2014). Collaboration: Strength in diversity. *Nature*, 513(7518), Article 305. <https://doi.org/10.1038/513305a>
- Ganga, E., & Mazzariello, A. (2018). *Math pathways: Expanding options for success in college math*. Education Commission of the States. <https://eric.ed.gov/?id=ED590584>
- Gipson, J. (2016). Predicting academic success for students of color within STEM majors. *Journal for Multicultural Education*, 10(2), 124–134. <https://doi.org/10.1108/JME-12-2015-0044>
- Givvin, K., Stigler, J., & Thompson, B. (2011). What community college developmental mathematics students understand about mathematics. Part II: The interviews. *MathAMATYC Educator*, 2, 4–18. [http://archive.carnegiefoundation.org/publications/pdfs/elibrary/MathAMATYC\\_Stigler.pdf](http://archive.carnegiefoundation.org/publications/pdfs/elibrary/MathAMATYC_Stigler.pdf)
- Goldrick-Rab, S. (2006). Following their every move: An investigation of social class differences in college pathways. *Sociology of Education*, 79(1), 61–70. <https://doi.org/10.1177/003804070607900104>
- Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, 20(1), 25–46. <https://doi.org/10.1093/pan/mpr025>
- Hammerman, N., & Goldberg, R. (2003). Strategies for developmental mathematics at the college level. *Mathematics and Computer Education*, 37(1), 79–95.
- Handel, M. J. (2016). What do people do at work? *Journal for Labour Market Research*, 49(2), 177–197. <https://doi.org/10.1007/s12651-016-0213-1>
- Hatfield, N., Brown, N., & Topaz, C. M. (2022). Do introductory courses disproportionately drive minoritized students out of STEM pathways? *PNAS Nexus*, 1(4), Article 167. <https://doi.org/10.1093/pnasnexus/pgac167>
- Herrera, F., & Hurtado, S. (2011). *Maintaining initial interests: Developing science, technology, engineering, and mathematics (STEM) career aspirations among underrepresented racial minority students*. 41. <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=1902394a863321accb2c38ad0af7b644f2470409>
- Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. Lester (Ed.), *Second Handbook of research on mathematics teaching and learning* (pp. 371–404). Information Age. <https://doi.org/10.1108/978-1-28141-200-320251013>
- Hodara, M. (2011). Improving pedagogy in the developmental mathematics classroom. *Community College Research Center: CCR Brief*, 51, 1–4. <https://eric.ed.gov/?id=ED516146>
- Hodara, M. (2019). Understanding the developmental mathematics student population: Findings from a nationally representative sample of first-time college entrants. *Workshop on Increasing Student Success in Developmental Mathematics*. <https://doi.org/10.17226/25547>
- Jennings, J. L., & Jonathan, J. M. (2014). “Teaching to the test” in the NCLB era: How test predictability affects our understanding of student performance. *Educational Researcher*, 43(8), 381–389. <https://doi.org/10.3102/0013189X14554449>
- Jensen, L. E., & Deemer, E. D. (2019). Identity, campus climate, and burnout among undergraduate women in STEM fields. *The Career Development Quarterly*, 67(2), 96–109. <https://doi.org/10.1002/cdq.12174>
- Juraev, D. A., & Bozorov, M. N. (2024). The role of algebra and its application in modern sciences. *IJFANS*, 3(1), 59–67. [https://www.researchgate.net/publication/378260241\\_The\\_role\\_of\\_algebra\\_and\\_its\\_application\\_in\\_modern\\_sciences](https://www.researchgate.net/publication/378260241_The_role_of_algebra_and_its_application_in_modern_sciences)
- Kalamkarian, H. S., Raufman, J., & Edgecombe, N. (2015). *Early implementation in Virginia and North Carolina* (Statewide Developmental Education Reform, p. 30). The Community College Research Center, Teachers College, Columbia University. <https://doi.org/10.7916/5xc6-gt93>
- Kosiewicz, H., Ngo, F., & Fong, K. (2016). Alternative models to deliver developmental math: Issues of use and student access. *Community College Review*, 44(3), 205–231. <https://doi.org/10.1177/0091552116651490>
- Larnell, G. V. (2016). More than just skill: Examining mathematics identities, racialized narratives, and remediation among black undergraduates. *Journal for Research in Mathematics Education*, 47(3), 233–269. <https://doi.org/10.5951/jresemathe-duc.47.3.0233>
- LeBeau, B., Harwell, M., Monson, D., Dupuis, D., Medhanie, A., & Post, T. R. (2012). Student and high-school characteristics related to completing a science, technology, engineering or mathematics (STEM) major in college. *Research in Science & Technological Education*, 30(1), 17–28. <https://doi.org/10.1080/02635143.2012.659178>
- Lin, L., Lee, T., & Snyder, L. A. (2018). Math self-efficacy and STEM intentions: A person-centered approach. *Frontiers in Psychology*, 9, Article 2033. <https://doi.org/10.3389/fpsyg.2018.02033>
- Logue, A. W., Douglas, D., & Watanabe-Rose, M. (2019). Corequisite mathematics remediation: Results over time and in different contexts. *Educational Evaluation and Policy Analysis*, 41(3), 294–315. <https://doi.org/10.3102/0162373719848777>
- Long, J. S. (1997). *Regression models for categorical and limited dependent variables*. Sage Publications. <https://uk.sagepub.com/en-gb/eur/regression-models-for-categorical-and-limited-dependent-variables/book6071>
- Lubienski, S. T., & Ganley, C. M. (2017). Research on gender and mathematics. In J. Cai (Ed.), *First compendium for research in mathematics education* (pp. 649–666). National Council of Teachers of Mathematics.
- Lubienski, S. T., Ganley, C. M., Makowski, M. B., Miller, E. K., & Timmer, J. D. (2021). “Bold problem solving”: A new construct for understanding gender differences in mathematics. *Journal for Research in Mathematics Education*, 52(1), 12–61. <https://doi.org/10.5951/jresemathe-duc-2020-0136>
- Maltese, A. (2008). *Persistence in STEM: An investigation of the relationship between high school experiences in science and mathematics and college degree completion in STEM fields* [PhD thesis, University of Virginia]. [https://www.researchgate.net/profile/Adam-Maltese/publication/234510692\\_Persistence\\_in\\_STEM\\_An\\_investigation\\_of\\_the\\_relationship\\_between\\_high\\_school\\_experiences\\_in\\_science\\_and\\_mathematics\\_and\\_college\\_degree\\_completion\\_in\\_STEM\\_fields/links/55255d880cf25d66dc9426e6/Persistence-in-STEM-An-investigation-of-](https://www.researchgate.net/profile/Adam-Maltese/publication/234510692_Persistence_in_STEM_An_investigation_of_the_relationship_between_high_school_experiences_in_science_and_mathematics_and_college_degree_completion_in_STEM_fields/links/55255d880cf25d66dc9426e6/Persistence-in-STEM-An-investigation-of-)

- the-relationship-between-high-school-experiences-in-science-and-mathematics-and-college-degree-completion-in-STEM-fields.pdf
- Manly, C. A., Wells, R. S., & Kommers, S. (2018). The influence of STEM definitions for research on women's college attainment. *International Journal of STEM Education*, 5(1), 45. <https://doi.org/10.1186/s40594-018-0144-1>
- McGee, E. O. (2015). Robust and fragile mathematical identities: A framework for exploring racialized experiences and high achievement among Black college students. *Journal for Research in Mathematics Education*, 46(5), 599–625. <https://doi.org/10.5951/jresmetheduc.46.5.0599>
- Melhuish, K., & Hicks, M. D. (2019). A validity argument for an undergraduate mathematics concept inventory. In J. D. Bostic, E. E. Krupa, & J. C. Shih (Eds.), *Quantitative measures of mathematical knowledge: Researching instruments and perspectives* (1st ed., pp. 121–152). Routledge. <https://www.taylorfrancis.com/chapters/edit/10.4324/9780429486197-6/validity-argument-undergraduate-mathematics-concept-inventory-kathleen-melhuish-michael-hicks>
- Merkin, R. (2023). Shortening the math sequence through a corequisite college algebra course with embedded personalized learning. *PRIMUS*, 34(2), 123–134. <https://doi.org/10.1080/10511970.2023.2250763>
- Miller, R. B., Greene, B. A., Montalvo, G. P., Ravindran, B., & Nichols, J. D. (1996). Engagement in academic work: The role of learning goals, future consequences, pleasing others, and perceived ability. *Contemporary Educational Psychology*, 21(4), 388–422. <https://doi.org/10.1006/ceps.1996.0028>
- Minaya, V. (2021). *Can dual enrollment algebra reduce racial/ethnic gaps in early STEM outcomes? Evidence from Florida*. Community College Research Center, Teachers College, Columbia University. <https://doi.org/10.7916/d8-gven-xq29>
- Moss-Pech, C. (2025). *Major trade-offs: The surprising truths about college majors and entry-level jobs*. University of Chicago Press. <https://press.uchicago.edu/ucp/books/book/chicago/M/bo246051255.html>
- National Research Council. (2001). *Adding it up: Helping children learn mathematics* (J. Kilpatrick, J. Swafford, & B. Findell, Eds.). National Academies Press. <https://doi.org/10.17226/9822>
- Ngo, F., Chi, W. E., & So Yun Park, E. (2018). Mathematics course placement using holistic measures: Possibilities for community college students. *Teachers College Record*, 120(2), 1–42. <https://doi.org/10.1177/016146811812000205>
- Nicholls, G. M., Wolfe, H., Besterfield-Sacre, M., & Shuman, L. J. (2013). Predicting STEM degree outcomes based on eighth grade data and standard test scores. *Journal of Engineering Education*, 99(3), 209–223. <https://doi.org/10.1002/j.2168-9830.2010.tb01057.x>
- Norton, E. C., & Dowd, B. E. (2018). Log odds and the interpretation of logit models. *Health Services Research*, 53(2), 859–878. <https://doi.org/10.1111/1475-6773.12712>
- Office for Civil Rights, US Department of Education. (2023). *A first look: Students' access to educational opportunities in U.S. public schools* (2020-21 Civil Rights Data Collection, p. 29). U.S. Department of Education. <https://eric.ed.gov/?id=ED664828>
- Palid, O., Cashdollar, S., Deangelo, S., Chu, C., & Bates, M. (2023). Inclusion in practice: A systematic review of diversity-focused STEM programming in the United States. *International Journal of STEM Education*, 10(1), 2. <https://doi.org/10.1186/s40594-022-00387-3>
- Park, T., Woods, C. S., Hu, S., Bertrand Jones, T., & Tandberg, D. (2018). What happens to underprepared first-time-in-college students when developmental education is optional? The case of developmental math and intermediate algebra in the first semester. *Journal of Higher Education*, 89(3), 318–340. <https://doi.org/10.1080/00221546.2017.1390970>
- Patel, A. (2024). *Exploring the role of gender, race, first-year college GPA and passing STEM Advanced Placement exams and the likelihood of graduating with a STEM degree: A multiple logistic regression analysis*. Western Kentucky University. <https://digitalcommons.wku.edu/cgi/viewcontent.cgi?article=4742&context=theses>
- Powell, M. G., Hull, D. M., & Beaujean, A. A. (2019). Propensity score matching for education data: Worked examples. *The Journal of Experimental Education*, 88(1), 145–164. <https://doi.org/10.1080/00220973.2018.1541850>
- Quarles, C. L., & Davis, M. (2017). Is learning in developmental math associated with community college outcomes? *Community College Review*, 45(1), 33–51. <https://doi.org/10.1177/0091552116673711>
- Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, 25, 111–163. <https://doi.org/10.2307/271063>
- Reardon, K., Bromley, K. W., & Unruh, D. (2021). The promise of universal design in postsecondary education: A literature review. *Journal of Postsecondary Education and Disability*, 34(3), 209–221. <https://eric.ed.gov/?id=EJ1325431>
- Rebello, N. S., Cui, L., Bennett, A., Zollman, D., & Ozimek, D. (2017). Transfer of learning in problem solving in the context of mathematics and physics. In D. H. Jonassen (Ed.), *Learning to solve complex scientific problems* (pp. 223–246). Routledge. <https://doi.org/10.4324/9781315091938-10>
- Redmond-Sanogo, A., Angle, J., & Davis, E. (2016). Kinks in the STEM pipeline: Tracking STEM graduation rates using science and mathematics performance. *School Science and Mathematics*, 116(7), 378–388. <https://doi.org/10.1111/ssm.12195>
- Reinholz, D., Johnson, E., Andrews-Larson, C., Stone-Johnstone, A., Smith, J., Mullins, B., Fortune, N., Keene, K., & Shah, N. (2022). When active learning is inequitable: Women's participation predicts gender inequities in mathematical performance. *Journal for Research in Mathematics Education*, 53(3), 204–226. <https://doi.org/10.5951/jresmetheduc-2020-0143>
- Richland, L. E., Stigler, J. W., & Holyoak, K. J. (2012). Teaching the conceptual structure of mathematics. *Educational Psychologist*, 47(3), 189–203. <https://doi.org/10.1080/00461520.2012.667065>
- Ridgeway, M. L., & McGee, E. O. (2018). Black mathematics educators: Researching toward racial emancipation of Black students. *The Urban Review*, 50(2), 301–322. <https://doi.org/10.1007/s11256-018-0452-2>
- Riegle-Crumb, C., King, B., & Irizarry, Y. (2019). Does STEM stand out? Examining racial/ethnic gaps in persistence across postsecondary fields. *Educational Researcher*, 48(3), 133–144. <https://doi.org/10.3102/0013189X19831006>
- Rittle-Johnson, B., & Schneider, M. (2014). 59. Developing conceptual and procedural knowledge of mathematics. In R. Cohen Kadosh & A. Dowker (Eds.), *The Oxford handbook of*

- numerical cognition (1st ed., pp. 1118–1134). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199642342.013.014>
- Roberts, M. T. (2020). Racism in remediation: How Black students navigate stereotypes to achieve success in developmental mathematics. *Community College Journal of Research and Practice*, 44(10–12), 701–721. <https://doi.org/10.1080/10668926.2019.1640143>
- Ryu, W., Schudde, L., & Pack, K. (2022). Constructing corequisites: How community colleges structure corequisite math coursework and the implications for student success. *AERA Open*, 8, 1–13. <https://doi.org/10.1177/23328584221086664>
- Sahin, A., Morgan, J., Capraro, M. M., & Capr, R. M. (2012). The effects of high school course taking and sat scores on college major selection. *Sakarya University Journal of Education*, 2(3), 96–109. [https://www.researchgate.net/publication/255682577\\_The\\_Effects\\_of\\_High\\_School\\_Course\\_Taking\\_and\\_SAT\\_Scores\\_on\\_College\\_Major\\_Selection](https://www.researchgate.net/publication/255682577_The_Effects_of_High_School_Course_Taking_and_SAT_Scores_on_College_Major_Selection)
- Sambell, K., & McDowell, L. (1998). The construction of the hidden curriculum: Messages and meanings in the assessment of student learning. *Assessment & Evaluation in Higher Education*, 23(4), 391–402. <https://doi.org/10.1080/0260293980230406>
- Sanabria, T., Penner, A., & Domina, T. (2020). Failing at remediation? College remedial coursetaking, failure and long-term student outcomes. *Research in Higher Education*, 61(4), 459–484. <https://doi.org/10.1007/s11162-020-09590-z>
- Schoenfeld, A. H. (2022). Why are learning and teaching mathematics so difficult? In M. Danesi (Ed.), *Handbook of cognitive mathematics* (pp. 1–35). Springer International Publishing. [https://doi.org/10.1007/978-3-030-44982-7\\_10-1](https://doi.org/10.1007/978-3-030-44982-7_10-1)
- Scott, A., & Martin, A. (2014). Perceived barriers to higher education in science, technology, engineering, and mathematics. *Journal of Women and Minorities in Science and Engineering*, 20(3), 235–256. <https://doi.org/10.1615/JWomenMinorScienEng.2014006999>
- Scott-Clayton, J. (2012). *Do high-stakes placement exams predict college success?* (CCRC Working Paper No. 41). Community College Research Center, Teachers College, Columbia University, 44. <https://ccrc.tc.columbia.edu/publications/high-stakes-placement-exams-predict.html>
- Snyder, J. L., & Cudney, E. A. (2017). A review of retention models for STEM majors and their alignment to community colleges. *Journal of STEM Education: Innovations and Research*, 18(3), 58–57. <https://www.jstem.org/jstem/index.php/JSTEM/article/view/2194>
- Stancher, A. (2019). *The impact of a holistic admissions process on college math completion rates at a Wisconsin Technical College*. University of Wisconsin. <https://minds.wisconsin.edu/bitstream/handle/1793/79670/2019stanchera.pdf?sequence=1&isAllowed=y>
- Stepter, C. M. (2023). *Truth-vetting & mathematics: Sense-making practices in an inequitable world*. Stanford University. <https://purl.stanford.edu/sr300hw4259>
- Stewart, S., & Reeder, S. (2017). Algebra underperformances at college level: What are the consequences? In S. Stewart (Ed.), *And the rest is just Algebra* (pp. 3–18). Springer International Publishing. [https://doi.org/10.1007/978-3-319-45053-7\\_1](https://doi.org/10.1007/978-3-319-45053-7_1)
- Stewart, S., Reeder, S., Raymond, K., & Troup, J. (2018). Could algebra be the root of problems in calculus courses? In C. Weinberg, J. Rasmussen, J. Rabin, M. Wawro, & S. Brown (Eds.), *Proceedings of the 21st annual conference on research in undergraduate mathematics education* (pp. 1023–1030), San Diego, CA. [http://sigmaa.maa.org/rume/crume2018/Abstracts\\_Files/Submissions/178\\_Could\\_Algebra\\_be\\_the\\_Root\\_of\\_Problems\\_in\\_Calculus\\_Courses.pdf](http://sigmaa.maa.org/rume/crume2018/Abstracts_Files/Submissions/178_Could_Algebra_be_the_Root_of_Problems_in_Calculus_Courses.pdf)
- Stigler, J. W., Givvin, K. B., & Thompson, B. J. (2010). What community college developmental mathematics students understand about mathematics. *MathAMATYC Educator*, 1(3), 4–16. [http://archive.carnegiefoundation.org/publications/pdfs/elibrary/MathAMATYC\\_Stigler.pdf](http://archive.carnegiefoundation.org/publications/pdfs/elibrary/MathAMATYC_Stigler.pdf)
- The Century Foundation. (2016). *When college students start behind (College Completion Series: Part Five)*. Higher Education. <https://tcf.org/content/report/college-students-start-behind/>
- Thomas, L. E., Thomas, S. M., Li, F., & Matsouaka, R. A. (2023). Addressing substantial covariate imbalance with propensity score stratification and balancing weights: Connections and recommendations. *Epidemiologic Methods*, 12(s1), Article 131. <https://doi.org/10.1515/em-2022-0131>
- Twigg, C. A. (2011). The math emporium: A silver bullet for higher education. *Change: The Magazine of Higher Learning*, 43(3), 25–34. <https://doi.org/10.1080/00091383.2011.569241>
- Tyson, W., Lee, R., Borman, K. M., & Hanson, M. A. (2007). Science, technology, engineering, and mathematics (STEM) pathways: High school science and math coursework and postsecondary degree attainment. *Journal of Education for Students Placed at Risk*, 12(3), 243–270. <https://doi.org/10.1080/10824660701601266>
- van den Hurk, A., Meelissen, M., & Van Langen, A. (2018). Interventions in education to prevent STEM pipeline leakage. *International Journal of Science Education*, 41(2), 150–164. <https://doi.org/10.1080/09500693.2018.1540897>
- VanderWeele, T. J., & Robinson, W. R. (2014). On the causal interpretation of race in regressions adjusting for confounding and mediating variables. *Epidemiology*, 25(4), 473–484. <https://doi.org/10.1097/EDE.0000000000000105>
- Villa Ross, C. A., Shin, H. B., & Marlay, M. C. (2021, October 27). *Pandemic impact on 2020 American Community Survey 1-year data*. <https://www.census.gov/newsroom/blogs/random-samplings/2021/10/pandemic-impact-on-2020-acs-1-year-data.html>
- Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, 101(4), 817–835. <https://doi.org/10.1037/a0016127>
- Walkington, C., Bernacki, M., Leyva, E., & Istas, B. (2025). *When am I (n)ever going to use this? How Algebraic functions are used in STEM-related careers*. <https://doi.org/10.5951/jresematheduc-2023-0091>
- Wang, X. (2013). Why students choose STEM majors: Motivation, high school learning, and postsecondary context of support. *American Educational Research Journal*, 50(5), 1081–1121. <https://doi.org/10.3102/0002831213488622>
- Webel, C., Krupa, E., & McManus, J. (2017). The math emporium: Effective for whom, and for what? *International Journal of Research in Undergraduate Mathematics Education*, 3(2), 355–380. <https://doi.org/10.1007/s40753-016-0046-x>

- Weinberg, A., Dresen, J., & Slater, T. (2016). Students' understanding of algebraic notation: A semiotic systems perspective. *The Journal of Mathematical Behavior*, 43, 70–88. <https://doi.org/10.1016/j.jmathb.2016.06.001>
- Winship, C., & Mare, R. D. (1984). Regression models with ordinal variables. *American Sociological Review*, 49(4), 512–525. <https://doi.org/10.2307/2095465>
- Wladis, C., Bjorkman, K., Duranczyk, I., Selbach-Allen, M., Schaub, B., & Tintera, G. (2023). What is college-level mathematics? A proposed framework for generating developmental progressions in mathematics up to and through college. In S. Cook, B. Katz, & D. Moore-Russo (Eds.), *Proceedings of the 25th Annual Conference on Research in Undergraduate Mathematics Education* (pp. 833–841). RUME. <http://sigmaa.maa.org/rume/Site/Proceedings.html>
- Wladis, C., Hachey, A. C., & Conway, K. M. (2024). It's about time, part II: Does time poverty contribute to inequitable college outcomes by gender and race/ethnicity? *AERA Open*, 10, 1–20. <https://doi.org/10.1177/23328584241237971>
- Wladis, C., Makowski, M., McKeown, K., & Williams, D. (2023). (Re)defining developmental mathematics: A critical examination of the research literature. In S. Cook, B. Katz, & D. Moore-Russo (Eds.), *Proceedings of the 25th Annual Conference on Research in Undergraduate Mathematics Education* (pp. 824–832). RUME. <https://doi.org/10.3102/2013099>
- Wladis, C., Murray, E., & Aly, G. D. (2025a). Undergraduate student conceptions of mathematical syntax meanings. In S. Cook, B. P. Katz, & K. Melhuish (Eds.), *Proceedings of the 27th Annual Conference on Research in Undergraduate Mathematics Education* (pp. 363–370). The Special Interest Group of the Mathematical Association of America (SIGMAA) for Research in Undergraduate Mathematics Education (RUME). <https://par.nsf.gov/servlets/purl/10627391>
- Wladis, C., Murray, E., & Aly, G. D. (2025b, February). *Predictive validity of the algebra concept inventory* [Conference session]. Proceedings of the 14th Congress of the European Society for Research in Mathematics Education (CERME14), European Society for Research in Mathematics Education (ERME); Bozen-Bolzano, Italy. <https://www.cerme14.it/proceedings/>
- Wladis, C., Murray, E., Hachey, A. C., Myszkowski, N., Aly, G. D., Sencindiver, B., Offenhholley, K., & Samuels, J. (2025c). Algebra conceptual understanding and predicting STEM degree completion. In S. Cook, B. P. Katz, & K. Melhuish (Eds.), *Proceedings of the 27th Annual Conference on Research in Undergraduate Mathematics Education* (pp. 328–337), The Special Interest Group of the Mathematical Association of America (SIGMAA) for Research in Undergraduate Mathematics Education (RUME); Alexandria, VA. <https://par.nsf.gov/servlets/purl/10627391>
- Wladis, C., Offenhholley, K., Licwinko, S., Dawes, D., & Lee, J. (2018). Development of the elementary algebra concept inventory for the college context. In C. Weinberg, J. Rasmussen, J. Rabin, M. Wawro, & S. Brown (Eds.), *Proceedings of the 21st Annual Conference on Research in Undergraduate Mathematics Education* (pp. 605–617), The Special Interest Group of the Mathematical Association of America (SIGMAA) for Research in Undergraduate Mathematics Education (RUME); San Diego, CA. [http://sigmaa.maa.org/rume/crume2018/Abstracts\\_Files/Submissions/75\\_Development\\_of\\_the\\_Elementary\\_Algebra\\_Concept\\_Inventory\\_for\\_the\\_College\\_Context.pdf](http://sigmaa.maa.org/rume/crume2018/Abstracts_Files/Submissions/75_Development_of_the_Elementary_Algebra_Concept_Inventory_for_the_College_Context.pdf)
- Wladis, C., Offenhholley, K., Sencindiver, B., Hachey, A., Myszkowski, N., Samuels, J., Mirin, A., Murray, E., & Aly, G. (manuscript under review). *Creation and validation of the Algebra concept inventory for college students* [Manuscript under review].
- Wladis, C., Offenhholley, K., Sencindiver, B., Myszkowski, N., & Aly, G. (2024a). The algebra concept inventory: Creation and validation with students across a range of math courses in college. In S. Cook, B. Katz, & D. Moore-Russo (Eds.), *Proceedings of the 26th Annual Conference on Research in Undergraduate Mathematics Education* (pp. 431–440). Proceedings of the 26th Annual Conference on Research in Undergraduate Mathematics Education. <https://par.nsf.gov/biblio/10558103>
- Wladis, C., Offenhholley, K., Sencindiver, B., Myszkowski, N., & Aly, G. D. (2024b). Creation and validation of the Algebra Concept Inventory in the tertiary context. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.), *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, p. 299), The International Group for the Psychology of Mathematics Education (IGPME); University Park, PA. <https://par.nsf.gov/servlets/purl/10558103>
- Wladis, C., Sencindiver, B., & Offenhholley, K. (2023). How do college students conceptualize algebraic properties? *Proceedings of the Thirteenth Congress of the European Society for Research in Mathematics Education (CERME 13)*, 676–683, The Special Interest Group of the Mathematical Association of America (SIGMAA) for Research in Undergraduate Mathematics Education (RUME); Omaha, NE. <https://hal.science/hal-04418311/document>
- Wladis, C., Sencindiver, B., Offenhholley, K., Jaffe, E., & Taton, J. (2022). Modeling student definitions of equivalence: Operational vs. structural views and extracted vs. stipulated definitions. In S. S. Karunakaran & A. Higgins (Eds.), *Proceedings for the 24th Annual Conference on Research in Undergraduate Mathematics Education* (pp. 708–715), The Special Interest Group of the Mathematical Association of America (SIGMAA) for Research in Undergraduate Mathematics Education (RUME); Boston, MA. <https://par.nsf.gov/biblio/10378561>
- Wladis, C., Hachey, A. C., & Conway, K. M. (2024). It's about time: The inequitable distribution of time as a resource for college, by gender and race/ethnicity. *Research in Higher Education*, 65(7), 1614–1646. <https://doi.org/10.1007/s11162-024-09796-5>
- Wladis, C., Hachey, A. C., & Conway, K. M. (2024). It's about time, Part II: Does time poverty contribute to inequitable college outcomes by gender and race/ethnicity. *AERA Open*, 10(1), 1–20. <https://doi.org/10.1177/23328584241237971>
- Wolniak, G. C. (2016). Examining STEM bachelor's degree completion for students with differing propensities at college entry. *Journal of College Student Retention: Research, Theory & Practice*, 18(3), 287–309. <https://doi.org/10.1177/1521025115622782>
- Wooldridge, J. M. (2002). *Econometric analysis of cross-section and panel data* (1st ed.). The MIT Press. <https://ipcid.org/evaluation/>

apio/Wooldridge%20-%20Cross-section%20and%20Panel%20Data.pdf

Wright, B., Gunther, O., & Bitar, J. (2023). Why STEM equity must address the experiences of women of color. *The Education Trust*. <https://edtrust.org/resource/why-stem-equity-must-address-the-experiences-of-women-of-color/>

Yeh, C., Ellis, M., & Mahmood, D. (2020). From the margin to the center: A framework for rehumanizing mathematics education for students with dis/abilities. *The Journal of Mathematical Behavior*, 58, Article 100758. <https://doi.org/10.1016/j.jmathb.2020.100758>

Zhang, Y. L. (2021). STEM persisters, switchers, and leavers: Factors associated with 6-year degree attainment for STEM aspiring community college transfer students. *Community College Journal of Research and Practice*, 46(11), 796–811. <https://doi.org/10.1080/10668926.2021.1906784>

### Authors

CLAIRE WLADIS is the Director of the CUNY Excellence Through Education Research Group, Professor of Mathematics at BMCC/CUNY and Urban Education at the CUNY Graduate Center. They currently lead several NSF-funded projects, including one to validate an algebra concept inventory and another to investigate factors that impact enrollment in or completion of math majors at two-year colleges.

BENJAMIN SENCINDIVER is an assistant professor at the University of Texas at San Antonio. His research interests include investigating college students' graphical and covariational thinking and factors that prompt students to change how they think about mathematical ideas.

ALYSE C. HACHEY is Chair of the Teacher Education Department, Director of the Division of Bilingual Education, Early Childhood Education, Literacy and Sociocultural Studies, and Professor-Lead Early Childhood Faculty in the College of Education at The University of Texas at El Paso. Her teaching and research interests focus on early childhood cognition and curriculum development and post-secondary online learning and retention, particularly for college populations often designated at high-risk of dropout.

NILS MYSZKOWSKI is an associate professor at Pace University. His research is focused on the application and improvement of psychometric methods to measure and understand creative, aesthetic and interpersonal skills, especially applied in occupational contexts.

KATHLEEN OFFENHOLLEY is a professor at the Borough of Manhattan Community College at the City University of New York. Her research interests include gaming in mathematics education and students' algebraic conceptual understanding; she is a former steering committee member of the CUNY Games Network.

JASON SAMUELS is a professor at the Borough of Manhattan Community College at the City University of New York. His research interests include investigating college students' thinking in calculus and algebra and developing innovative approaches to teaching calculus with an emphasis on quantitative reasoning.

GEILLAN ALY was previously a project manager for the Excellence Through Education Research Group at the City University of New York. She is currently focusing on her role as the founder Compassionate Math, LLC where she works to prevent and heal math trauma.