education sciences

MDPI

Article

The "Better Book" Approach to Addressing Equity in Statistics: Centering the Motivational Experiences of Students from Racially Marginalized Backgrounds for Widespread Benefit

Claudia C. Sutter [1,*], Matthew C. Jackson [2], Karen B. Givvin [1], James W. Stigler [1] and Ji Y. Son [2]

[1] Department of Psychology, University of California, Los Angeles, CA 90095, USA; karen.givvin@ucla.edu (K.B.G.); stigler@ucla.edu (J.W.S.)
[2] Department of Psychology, California State University, Los Angeles, CA 90032, USA; mjacks40@calstatela.edu (M.C.J.); json2@calstatela.edu (J.Y.S.)
* Correspondence: claudia@coursekata.org

Abstract: Although improving racial equity in critical college courses such as introductory statistics is a laudable goal, making research-based progress toward that aim in a scalable manner remains a challenge. To translate psychological insights to benefit racially marginalized students, we implemented the "Better Book" approach, where instructors, researchers, and developers work together to improve an online textbook used in introductory statistics. The "Better Book" approach to equity assumes that racially marginalized students are a "canary in the coal mine", alerting us to systemic issues that can affect a broader array of students. We started by finding places in the textbook where racially marginalized students reported higher perceptions of costs (the effort and time required to learn the content) than non-marginalized students. Then we drew upon suggestions from users to redesign the textbook where gaps in cost perceptions peaked. We then analyzed data from both the original and redesigned versions of the textbook to evaluate the impact on students who were subsequently enrolled in the course. Results showed that perceptions of cost were dramatically reduced in the experience of racially marginalized students but also the redesign resulted in an improved experience for all students.

Keywords: translational research; improvement science; introductory statistics; motivation; equity

Citation: Sutter, C.C.; Jackson, M.C.; Givvin, K.B.; Stigler, J.W.; Son, J.Y. The "Better Book" Approach to Addressing Equity in Statistics: Centering the Motivational Experiences of Students from Racially Marginalized Backgrounds for Widespread Benefit. *Educ. Sci.* **2024**, *14*, 487. https://doi.org/10.3390/educsci14050487

Academic Editors: Marie K. Norman and Michael W. Bridges

Received: 29 January 2024
Revised: 22 April 2024
Accepted: 25 April 2024
Published: 2 May 2024

1. Introduction

Statistics is a complex domain, hard to learn and teach, but required for advancement in many fields. Learning statistics can empower people to make informed decisions, critically evaluate research, and communicate findings effectively [1]. Moreover, in an increasingly data-driven world, proficiency in statistics can enhance problem-solving abilities and foster a deeper comprehension of complex phenomena. Introductory statistics is a required course for many popular STEM majors such as psychology [2,3], and student success in introductory statistics is critical for fulfilling transfer or college math requirements [4].

While statistics is considered a challenging topic of study by a variety of students [5], statistics courses may present unique challenges and barriers to students from racially marginalized backgrounds [6]. These challenges may lead to disparities in course outcomes such as performance [7] and future interest in statistics [8] and may also lead marginalized students to have lower levels of motivation and higher perceptions of emotional cost or cost related to effort [8].

Although performance and course grades are important outcomes, in this paper, we seek to understand racial equity gaps in statistics from a motivational lens. Not only does motivation predict a variety of student success and performance measures [9,10]), students' motivation serves as an indicator of the quality of their learning experiences. As part of a broader effort to improve racial equity in a variety of outcomes, motivational measures

can be conceptualized as a "pulse check" of their learning experiences. Motivation can also profoundly impact students' college and career journeys, including their persistence in pursuing a chosen field [9]. Students who leave their statistics course with negative experiences or low perceptions of its value are unlikely to apply what they have learned in their everyday lives [11]

To unpack the psychological construct of motivation, we turn to the situated expectancy-value [9] and expectancy-value-cost models of motivation [10]. These motivational frameworks suggest that student motivation comprises (1) expectancy for success (i.e., confidence in their ability to successfully complete a task), (2) value beliefs including intrinsic value (i.e., interest or enjoyment from engaging in a task), utility value (i.e., perceived usefulness of the course), and attainment value (i.e., importance of doing well) and (3) perceptions of cost (i.e., negative consequences of engaging in a task such as the time, energy, and resources required for learning).

Studies that disaggregate motivational and performance outcomes in introductory college-level statistics courses by race/ethnicity from an expectancy-value-cost perspective are rare e.g., [7,8,11,12]. One of the earlier studies to do so focused primarily on expectancy and value [13]. Researchers found that Black, Latine, and White students had largely similar expectations for their performance in an introductory statistics course (i.e., expected course grade prior to any assignments or examinations), and all three groups performed lower than their expectations. The gap between expectation and performance was larger, however, among Black and Latine students (both groups performing 0.82 semester GPA units below expectation) than among White students (who performed 0.57 semester GPA units below expectation). Asian students had higher expectations for their performance than did Black, Latine, and White students, but their performance was not significantly different from their expectations. Interestingly, researchers found no racial differences in the perceived value of learning statistics across the four groups, which all groups rated as high. This suggests that although students do not differ in how highly they value statistics skills and knowledge, many students struggle to meet their own expectations for success. Because this gap is largest among racially marginalized populations, understanding its causes may have the strongest benefit for these populations, though such insights have the potential to benefit other populations as well.

In a recent study that focused on cost (specifically effort cost and emotional cost), Sutter and colleagues [8] found that students from racially marginalized backgrounds experienced higher levels of cost in their introductory statistics course than did their non-marginalized peers. Further, incoming course concerns (e.g., concern about lack of prior knowledge) negatively predicted future interest in statistics, but only for racially marginalized students. And, importantly, students' perceptions of cost mediated this relationship (i.e., incoming concerns about their statistics class positively predicted perceptions of cost, which then negatively predicted future interest in statistics).

1.1. Why Are There Inequities in Intro Stats?

There are many possible reasons for why there are racial equity gaps in motivation (particularly in students' perceptions of cost). For example, economic disparities (e.g., racially marginalized students might also have less access to the internet or a computer at home), time constraints (e.g., racially marginalized students experience higher demands on their time due to responsibilities outside of school; [8,14]), and inequities in earlier educational opportunities may contribute to racial disparities in motivational experiences. These barriers might not only impede racially marginalized students' ability to engage with educational materials but also exacerbate feelings of disconnection and disengagement.

Other reasons include various features of instructional practices, policies, and content, which can negatively and disproportionately impact the learning experiences of students from racially marginalized groups and could cause these students to experience higher perceived cost, a lack of belonging, or stereotype threat [6,15,16]. For example, a key systemic factor contributing to racial inequities in introductory statistics and shaping

students' experiences is the representation—or lack thereof—in the course content, which may lead students from racially marginalized backgrounds to struggle to see themselves reflected in examples, applications, and datasets of statistics courses, perhaps making it harder to grasp abstract statistical concepts or see the applicability or relevance of statistical contexts to their lives and communities [17]. Format of a course may also play a crucial role. For instance, the length of chapters and overall workload can differentially affect marginalized students. Students with higher demands on their time might be less flexible when needing to adapt to fluctuations in assignment length. Therefore, it is essential not only to diversify the content but also to reconsider the course structure, ensuring that it does not impose disproportionate barriers on racially marginalized students.

Given the importance of statistics as a gateway course and the sparse literature that has disaggregated students' motivational experiences in introductory college-level statistics courses by race/ethnicity [7,8,12,13], understanding the current state of racially marginalized students' motivational experiences is a critical first step in our investigation. In trying to document equity gaps in expectancy, value, and cost, it is important to note that these motivational beliefs can be sensitive to context [9]. For example, within the learning context, perceptions of cost might change within the same week depending on the learning content, material, topic, task difficulty, or task type [18,19]. Thus, if we are to understand the current state and paint a clear picture of equity gaps, we must commit to collecting multiple motivation measures over the duration of a statistics course.

1.2. Approaches to Address Racial Equity Gaps in Motivation

There are many different approaches to reduce gaps in motivational experiences between racially marginalized and non-marginalized groups. We want to contrast two broad approaches: individual-change and systems-change (see Figure 1).

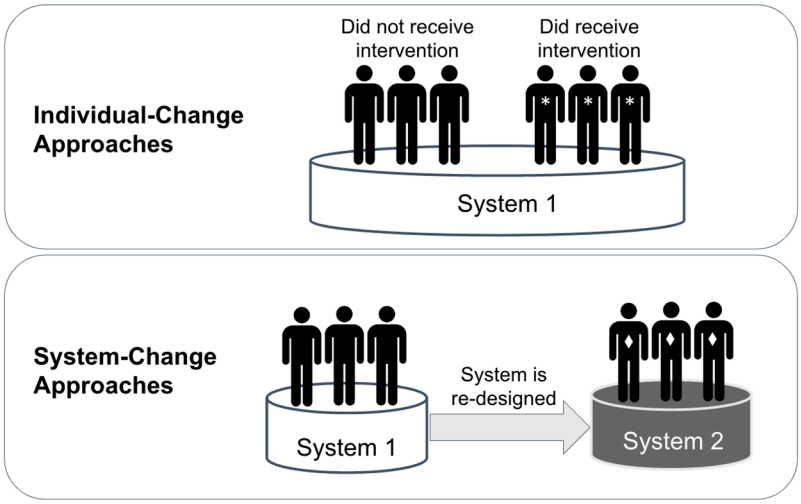


Figure 1. Individual-change and systems-change approaches to addressing equity gaps in motivation.

Individual-change approaches may assign students to either receive an intervention or not in order to ascertain whether a particular intervention improves students' motivational experiences. For example, utility-value interventions, which engage students in reflecting on how the course content relates to their lives (e.g., an essay about the personal relevance of biology; [20]), have increased students' perceptions of the usefulness of the material, course interest, and academic engagement. Although any student can benefit from utility-value interventions, they are particularly effective for students from traditionally underrepresented and racially marginalized backgrounds, including first-generation college students, Black and Latine students, and students from lower socioeconomic backgrounds [20–23]. However, systems-change approaches target systems by dismantling structural barriers that may generate educational disparities and inequities. In contrast to the individual-change

approach, there are no "interventions" given to students. Instead there is a re-design of the system. The system-change approach focuses on making changes to the system, which then leads to changes in the students' experience. Because system-change is difficult to accomplish, let alone conduct research on, these approaches are less common in teaching and learning research.

In the project reported here, we focus on one aspect of "the system" in which learning takes place: the textbook/curriculum. The textbook is a component that lies outside the control of both students and instructors, yet it can be studied and redesigned to promote equity. In the following section, we lay out the "Better Book" approach [24], which uses improvement science to continuously improve complex systems—in this case to guide and test changes in the textbook for the purpose of improving equity.

2. The "Better Book" Approach—A Quality Improvement Approach to Addressing Racial Equity Gaps in Motivation

Although improving inclusivity and promoting racial equity in an educational system are laudable goals, making research-based progress towards these and other aims in a scalable manner remains a challenge for three reasons: complexity, implementation, and silos. First, teaching and learning is a complex cultural system [24]. Thus, most research findings from a highly controlled and simplified environment (e.g., a research lab) fail to translate to improvements in the overall system. Second, improving an educational system is difficult because systems, especially innovative and novel systems, need to be implemented before they can be studied [25]. And third, the work of practitioners, researchers, and designers/developers tends to take place in silos instead of in the context of collaborative problem-based inquiry [24].

The "Better Book" approach [24], an innovative approach to doing research and development in educational contexts, addresses each of these challenges to making systemic change. The basic idea of the "Better Book" approach is that a community of designers/developers, researchers, instructors, and students focus their combined efforts on improving an online interactive textbook and its implementation. We consider the interactive textbook and the instruction that surrounds it a small-scale system on which all parties can focus their efforts. This approach also leverages the affordances of new technologies, which make it possible to collect data from enrolled students directly from within the textbook and over extended time frames (e.g., a semester or year). These data represent students' experiences and can not only point out problems in the system but also guide improvement.

Beyond simply collecting surveys and learning data from students, the textbook explicitly invites students to participate in the improvement process. For example, on one of the first pages of the textbook, students read the following statement: "Help Us Improve the Book. You can also create a ticket if you want to report a problem, or make a suggestion, that could help us to improve the quality of CourseKata Statistics and Data Science. We welcome your feedback! Please note that tickets are used for reporting technical issues and for offering ideas and suggestions for improving the book". Thus, students are aware that their responses and input is being solicited as a means of continuously improving the quality and effectiveness of the textbook and thus benefit future students.

The "Better Book" approach borrows quality improvement methodologies developed outside of education (such as improvement science and open-source software development), specifically designed to address problems of complex systems improvement. Because the textbook is an authentic setting in which students are learning, rather than a laboratory setting divorced from a broader course framework, any research-based insights gained from examining data from textbook use can be directly translated into changes in the book. Because all the data from the textbook interactions are collected in authentic settings by real students taking courses, issues of implementation within a broader course framework can be identified immediately. This model also gives designers/developers, researchers, instructors, and students a concrete system on which to collaborate. Despite being very

early in the application of the "Better Book" approach, we believe much can be learned from this novel method of doing research and development in education.

The focus of this paper will be CourseKata.org's instantiation of the "Better Book" approach, in which we engage a community in improving a statistics textbook [26] that contains over 1400 formative assessment questions, coding exercises, and measures of student motivation and engagement. The curriculum combines authentic data science tools used by professionals (i.e., R and Jupyter notebooks) with a pedagogy designed to help students connect problems in the world to core concepts and representations. The book, called *Introductory Statistics with R* (formerly: *Statistics and Data Science: A Modeling Approach*), is designed to present the content of introductory statistics under a unified framework of DATA = MODEL + ERROR ([26]; the entire textbook is publicly available at CourseKata.org accessed on 22 April 2024).

The "Better Book" approach has been used with CourseKata data to improve the teaching of specific statistics content [27,28] and to study the implementation of the instructional materials in a variety of institutional contexts. Researchers have only recently begun to use the "Better Book" approach to understand the experiences of racially marginalized students and implementation in minority-serving institutions [8]. In the next section, we describe how the Better Book approach generally uses improvement kata (or routine) to improve a textbook. Then, we describe the specific improvement kata directed to improve the textbook experience for racially marginalized students and how it benefited students.

3. The Improvement Kata

To improve teaching and learning, we turn to improvement science, a field that develops effective and replicable methods for translating research into the improvement of complex systems [24,29,30]. Improvement science works by (1) setting a direction or goal for the system; (2) understanding the current condition of the system; (3) establishing the next target for the system; and then (4) systematically experimenting with small changes designed to overcome barriers, studying results, and revising theories about how to reach a desired end. Improvement science is often implemented with routines of scientific thinking such as the Improvement Kata as described by Rother ([31,32], see Figure 2).

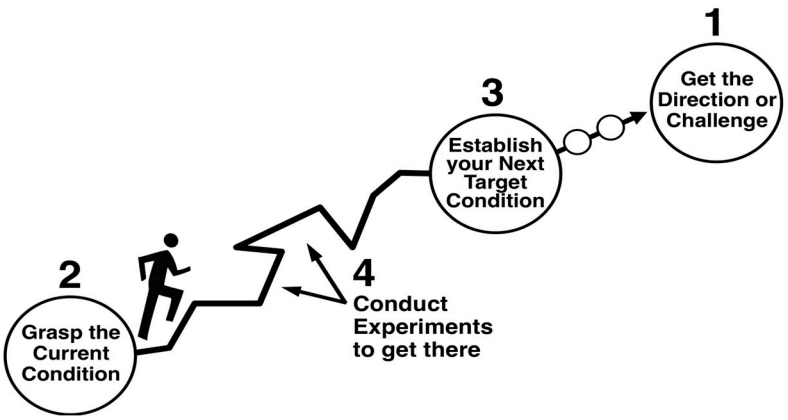


Figure 2. The Improvement Kata as described by Rother [31,32].

The overarching goal of this study is to apply the "Better Book" approach to increase racial equity within the context of an introductory statistics course. We began by examining student response data to identify the sections and features of the textbook that posed barriers to the motivational beliefs of racially marginalized students, threatening their potential engagement and achievement. Given the dynamic and situational nature of motivation, we measured students' motivational beliefs 10 times over a 10-week term in an introductory statistics course. Having identified a potential barrier to motivation, we set the intermediate goal of making changes to the textbook that would ameliorate it and benefit future students enrolled in the course. We then conducted an improvement cycle

wherein we examined and addressed deficits in the textbook by centering the experiences of racially marginalized students. Drawing on their voices (and those of their instructors), we redesigned the textbook and then evaluated the impact on racially minoritized students and their non-racially minoritized peers in the next cohort. The example demonstrates how data from students can point out flaws in a textbook, point the way toward possible design solutions, and help us evaluate whether the redesigned book is indeed better.

3.1. Get the Direction or Change

In the language of the Improvement Kata, the "direction" for the system is to promote racial equity in representation and achievement in STEM broadly, and in a statistics course that uses the CourseKata textbook, in particular. One can think of increasing equity as broadening opportunity by removing systemic barriers and ensuring access to the resources needed to thrive. An equitable learning environment is supportive of all students to develop their full academic potential.

3.2. Grasp the Current Condition (Study 1)

Step 2 of the Improvement Kata cycle is to grasp the current condition. Learning statistics is a complex task, and to begin to improve students' learning, we first need to understand students' experiences as they navigate the system in which they are learning. In our work, we will use student motivation (i.e., why individuals choose to engage in learning as well as why they might not want to engage), which can fluctuate over the course of a school term, as a key indicator of their experience. By tracking changes in students' motivation throughout the course, we can ascertain where changes to the system are needed. We can also disaggregate motivational measures to prioritize the experiences of racially marginalized students.

From situated expectancy-value theory and expectancy-value-cost theory, we have identified success expectancy, intrinsic value, utility value, and cost as key indicators of students' motivation that may influence learning and performance over the course of a full college term (i.e., 10 weeks during which students complete 12 textbook chapters). Most educators want their students to feel confident in their ability to succeed in the course, to perceive their learning as interesting and useful, and to have few barriers and negative consequences for engaging in learning tasks. Finding out whether students actually experience the course in these ways is the second step in our improvement cycle (see Figure 1).

3.2.1. Research Questions

Within the context of the online introductory textbook, we tried to understand the current condition by asking:

1. How are students experiencing the textbook? What trends and patterns are we noticing in their motivation (i.e., expectancy, value, cost)?
2. Are there any "hot spot" chapters where students' motivations are spiking or waning?
3. Do racially marginalized students show patterns of motivation that differ from their non-marginalized peers?

To get at these questions, we analyzed students' self-reported levels of motivation longitudinally in one large introductory statistics class.

3.2.2. Participants

Participants were 219 students from an introductory statistics class (79.0% female, 19.6% male, and 1.4% non-binary) at a large public university. We disaggregated the data to gain insight into the experiences and perspectives of students from diverse backgrounds, including different racial/ethnic backgrounds [33]. Of the students who indicated their race/ethnicity (n = 214), 39.3% identified as Asian or Asian American (n = 84), 22% as White (n = 47), 18.7% as Hispanic or Latine (n = 40), 3.7% as Black or African American (n = 8), and 16.4% as mixed or other races/ethnicities (n = 35). For the purpose of exploring students'

experiences by race/ethnicity, Black or African American, Hispanic or Latine, Native American, Native Hawaiian or Pacific Islander, and Greater Middle Eastern (e.g., Afghanistan, Pakistan) students were considered as belonging to a racially marginalized group, whereas White and Asian students were considered non-marginalized students. Students of mixed race were included in the racially marginalized group unless their race was a mix of White and Asian.

3.2.3. Measures

Students' motivational beliefs were measured eleven times over the course of the ten-week term (i.e., at the beginning of each of chapters 2–12). Students rated their perceptions of success expectancy ("I am confident in what I have learned so far in this course"), utility value ("I think what I have learned so far in this course is useful"), intrinsic value ("I think this class is interesting"), and cost ("I am unable to put in the time needed to do well in this course") on a six-point Likert scale ranging from 1 (strongly disagree) to 6 (strongly agree). This longitudinal data collection allowed us to identify particular pain points or hotspots (i.e., chapters in which students experienced a decline in expectancy or value or chapters that students perceived as particularly costly) at the chapter level in the textbook. Because the measures are embedded at the beginning of each chapter, they reflect students' (cumulative) experiences, particularly highlighting their beliefs about the prior chapter. For example, students' judgments of expectancy or cost at the beginning of chapter 8 reflect their motivational experiences upon completing chapter 7.

3.2.4. Analysis Plan

To grasp current conditions, we analyzed how students' levels of success expectancy, intrinsic value, utility value, and perception of cost developed at the chapter level, accounting for experiences by race. We used a 2 × 11, racially marginalized status × chapter, mixed repeated measures MANOVA (with racially marginalized status as the between-subjects variable and chapter as the within-subjects variable). We followed the overall analysis with post-hoc tests.

3.2.5. Results

Figure 3a–d depict the mean values of expectancy (Figure 3a), intrinsic value (Figure 3b), utility value (Figure 3c), and cost (Figure 3d) at the beginning of chapters 2–12, disaggregated by racially-marginalized status. The figures show that racially marginalized students and non-racially marginalized students differed in their ratings of cost, but their ratings of success expectancy, intrinsic value, and utility value were similar.

Mixed repeated measures MANOVA supports these observations. There was no statistically significant effect of racial marginalized status for expectancy (F (1, 188) = 0.111, p = 0.740), intrinsic value (F (1, 188) = 0.981, p = 0.323), and utility value (F (1, 188) = 0.226, p = 0.635). However, there was a significant effect of racially marginalized status on perceptions of cost (F (1, 188) = 8,783, p = 0.003). We also examined the effect sizes of racially marginalized status on perceptions of cost across each of the chapters (see Supplemental Table S1). The largest effect (Cohen's d = 1.262) was on ratings at the beginning of chapter 8. This helped us prioritize our investigation and redesign efforts on the contents of chapter 7, with more minor changes made in other parts of the textbook.

To further grasp the current condition and better understand why the ratings diverged at this point in the textbook, we started with a thorough review of comments and suggestions provided by students, instructors, researchers, and curriculum designers that were specifically relevant to chapter 7. One source of student and instructor comments was the interactive textbook's helpdesk ticketing process, where any user could submit issues, bugs, questions, comments, and ideas while working through the textbook. The ticketing system captures crucial information, such as the user's role (student or instructor) and the specific page on which the ticket was submitted. Most of the tickets students submitted while in chapter 7 referred to specific content they found difficult or confusing in this chapter

(e.g., General Linear Model notation, dummy coding). However, because helpdesk tickets are not connected to students' demographic information, we could not disaggregate them by racially marginalized status.

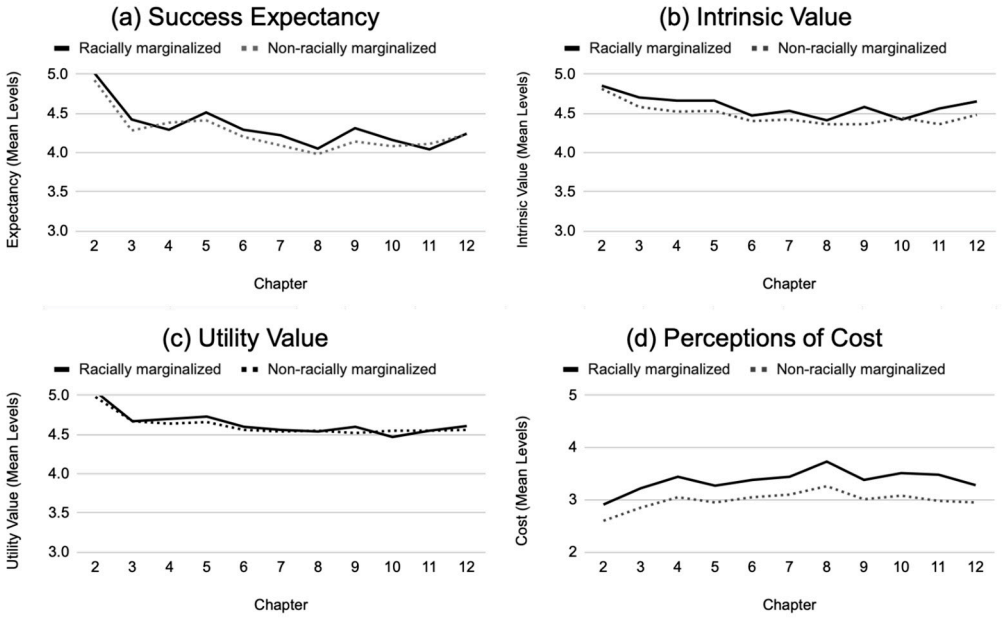


Figure 3. (**a**–**d**) Longitudinal trends in expectancy, intrinsic and utility value, and cost by chapter and racially marginalized status for students enrolled in the course using the original version of the textbook (v4.0).

We also gathered insights and information from marginalized students using the CourseKata materials and their instructors. We solicited and examined detailed notes from students and instructors at Hispanic-Serving Institutions and from community colleges with large numbers of marginalized students. In addition, students who completed the textbook in a class taught by its co-authors provided chapter-by-chapter feedback. Latine undergraduate student research assistants who were involved in data collection, or were conducting studies about statistics education, also contributed their insights. Lab meetings attended by predominantly marginalized students (but led by a non-marginalized faculty member) served as a focus group.

One Latino research assistant pointed out that some pages were much longer than others and suggested that could have an impact on student experiences. A Latina research assistant noticed that when she had more time, she was more likely to make repeated attempts to get the coding exercises correct instead of rushing to simply answer them. Instructors who had participated in professional development opportunities and subsequently taught the course at various institutions with sizable marginalized populations (from 40% up to 80%) also shared their suggestions. For example, we received valuable emailed notes about Chapter 7 from an instructor at a community college who was an early member of the CourseKata community.

3.3. Establish the Next Target Condition

Based on the initial step of grasping the current conditions in the Improvement Kata cycle, which revealed that all students experienced particularly high levels of cost in chapter 7 and students from racially marginalized backgrounds experienced significantly higher perceptions of cost throughout the textbook, the next target conditions could be established: (1) reducing perceptions of cost for all students and (2) reducing the gap in perceptions of cost between racially marginalized students and non-racially-marginalized students.

3.4. Conduct Quasi-Experiment on Redesigned Chapter 7 (Study 2)

After synthesizing the various observations and suggestions, the curriculum developers planned a revision to the textbook, applying the Plan-Do-Study-Act cycle (or PDSA; also called the Deming or Shewhart cycle) with the next target condition in mind.

When we established the current condition, the textbook was already in version 4.0, but for simplicity we will refer to it as original. The redesigned textbook was version 5.0. In the original version, chapter 7 began with the General Linear Model (GLM) notation for two-group models (in many introductory statistics courses, this would be presented as a *t*-test) and then presented the same model visually on a graph of data (as summarized in Table 1). Students and instructors both found the GLM notation highly abstract and difficult to follow but appreciated the graphs that came a few pages later. Some suggested flipping the order. Thus, in the revision, the idea of a two-group model was first presented visually and then the GLM notation was connected to those visual representations. This flow, called the "visualization-first" pedagogy (in contrast to the "algebra-first" pedagogy employed in the original), is also supported by research in the cognitive sciences that suggests that highly abstract symbols might be better understood if first given meaning in a visuospatial representation [34,35], making the material more intuitive for learners.

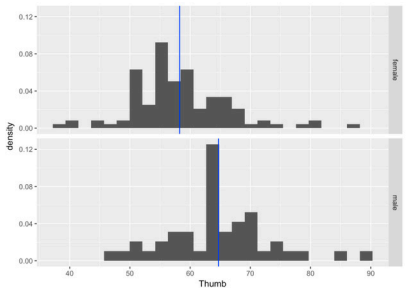
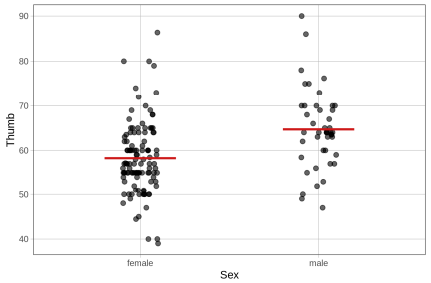
The revisions also led us to simplify the R code we used to teach students how to put models on their visualizations and to select a simpler visualization (scatter plots) as shown in the last row of Table 1. Statistics education research has documented the many misconceptions students have about histograms [36,37]; so to focus students' attentional resources on learning about the models rather than learning about interpreting difficult visualizations, we used scatter plots (also called jitter plots). This "visualization-first" approach and the simplified R code were first worked out for two-group models in chapter 7 but then implemented for the empty model (models with no explanatory variable) in chapters 5 and 6 as well as regression models in chapter 8.

Additionally, we followed up on the suggestion of the research assistant who wondered whether the variation in page lengths had an effect on students' flagging motivation. We conducted word counts of all pages in chapter 7 as well as the surrounding chapters. We found that chapter 7 was almost twice as long and its pages were much longer than those in other chapters. Thus, we split the content of Chapter 7 into two separate chapters (now Chapters 7 and 8). Splitting the chapter into two likely had implications for implementation, because college instructors tend to spend about one week covering each chapter of content. We hypothesized that these structural changes to the chapters and page lengths might induce instructors to slow down through this particular stretch of more dense content.

We also went through other parts of the textbook to bring other chapters (e.g., chapters 5 and 6) in parallel with the approach taken in chapter 7 and to reduce the variability in page lengths where possible. In the original version (4.0), the textbook had an average of 1341 words per page with a standard deviation of 627. The redesigned chapters (of version 5.0) had fewer words per page (M = 1095) and also a lower standard deviation (SD = 331), suggesting greater consistency of lengths across the redesigned pages. Note that word count is only a rough measure of how long students spend on a page because our method of counting words includes some non-content "words" (e.g., LaTeX tags) and does not include words from the interactive questions on a page. Complexity of the content also impacts the length of time spent on a page.

Overall, the redesign process incorporated a wide range of perspectives and inputs from various stakeholders, including students, instructors, researchers, and curriculum designers. Soliciting the insights of textbook users at Hispanic-Serving Institutions as well as Latine student and researcher perspectives led the way to major design changes. The resulting redesign was not the idea of one author or one instructor; it was the incorporation of many different viewpoints that all led to re-ordered and re-structured content. After making these changes, we were ready to ask: Are these changes an improvement?

Table 1. Summary of chapter 7 layout comparing original and redesigned versions.

Pages	Original Chapter 7 (Version 4.0) "Algebra-First" Pedagogy	Redesigned Chapter 7 (Version 5.0) "Visualization-First" Pedagogy
7.1	Introduced the context: modeling the variation in the thumb lengths of students with the explanatory variable sex (female versus male)	Introduced the context: modeling the variation in the thumb lengths of students with the explanatory variable sex (female versus male) with visual representations of the data (i.e., scatter plots)
7.2	Introduced the algebraic notation for the best-fitting General Linear Model for two groups ($Y_i = b_0 + b_1X_i + e_i$)	Introduced how to make a visual representation of the two-group model (as mean lines on a scatter plot) now presented before algebraic notation which moved to 7.3 Focused on interpreting the parameter estimates from R code in context of the visualization
7.3	Focused on interpreting parameter estimates from R code in context of the equation	Introduced algebraic notation of the two-group model ($Y_i = b_0 + b_1X_i + e_i$) and how it connects to the visual representation
7.4	Focused on using the algebraic equation to generate predictions	Focused on using the algebraic equation to generate predictions
7.5	Focused on calculating residuals from the predictions produced by the equation. At the end of the page, introduced how to make a visual representation of the two-group model (as mean lines on faceted histograms)	Focused on calculating residuals from the predictions produced by the equation.
Code used to visually represent a model on a graph	Emphasis on faceted histograms with slightly more complex code  Thumb_stats <- favstats(Thumb ~ Sex, data = Fingers) gf_dhistogram(~Thumb, data = Fingers) %>% gf_facet_grid(Sex ~ .) %>% gf_vline(xintercept = ~mean, data = Thumb_stats)	Emphasis on scatter plots with simplified code  model <- lm(Thumb ~ Sex, data = Fingers) gf_jitter(Thumb ~ Sex, data = Fingers, width = 0.1) %>% gf_model(model)

3.4.1. Research Questions

To assess the effectiveness of the redesign, we first compared students' responses to the motivation measures between the redesigned (version 5.0) textbook and the original (version 4.0) textbook. We were interested in the following questions:

1. Are there differences in perceptions of cost for students who completed the original textbook (version 4.0) versus students who completed the redesigned textbook (version 5.0)?
2. Is there an interaction between cost perceptions of the two textbook versions and students' racial background?

3.4.2. Participants

To make the comparison more meaningful, we analyzed student responses to the redesigned textbook in a statistics course taught by the same instructor at the same university as taught the materials in the original study. The sample of students who completed version 5.0 consisted of 235 students. Of those who indicated their gender (n = 231), 78.4% identified as female, 19% as male, and 2.6% as non-binary). Of the students who indicated

their race/ethnicity (n = 210), 39.5% identified as Asian or Asian American (n = 83), 27.1% as White (n = 57), 23.3% as Hispanic or Latine (n = 49), 4.8% as Black or African American (n = 10), and 5.2% as mixed or other races/ethnicities (n = 11). As in the original study, students who identified as a race other than White or Asian or a combination of those two races were considered racially marginalized (n = 70; 33.3%) and White or Asian students (n = 140; 66.7%) were considered non-racially marginalized.

3.4.3. Measures

The same measures that were embedded in the original version of the textbook were also embedded in the redesigned version at the same time points (i.e., at the beginning of chapters 2–12): success expectancy ("I am confident in what I have learned so far in this course"), utility value ("I think what I have learned so far in this course is useful"), intrinsic value ("I think this class is interesting"), and cost ("I am unable to put in the time needed to do well in this course").

3.4.4. Results

Figure 4a–d depict the mean values of expectancy (Figure 4a), intrinsic value (Figure 4b), utility value (Figure 4c), and cost (Figure 4d) at each chapter, disaggregated by racially marginalized status in Version 5.0 of the textbook. These figures are similar to Figure 2a–d but are for the redesigned textbook. Notably, in the redesigned textbook, the gap between perceptions of cost by racially marginalized and non-racially marginalized students seems to have shrunk.

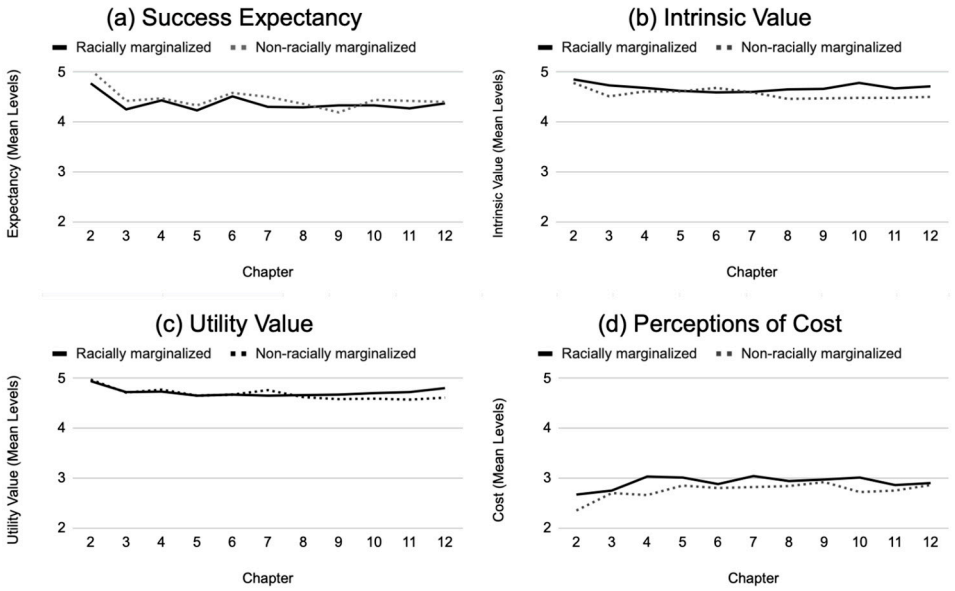


Figure 4. (**a**–**d**) Longitudinal trends in expectancy, intrinsic and utility value, and cost by chapter and racially marginalized status for students enrolled in the course using the redesigned textbook (version 5.0).

To compare the two versions of the textbook directly against one another, we ran a 2 × 2 × 11, version × racially marginalized status × chapter, mixed repeated measures ANOVA on the perceptions of cost, the main motivational measure hypothesized to change across the two versions. This analysis revealed three main effects. There was a significant effect of chapter, (F (10, 3680) = 11.583, $p \leq 0.001$), indicating that perceptions of cost varied at the chapter level. There was also a significant effect of version (F (1, 368) = 15.798, p <= 0.001) with the average perception of cost being significantly lower for the redesigned textbook compared to the original textbook (see Figure 5). Further, there was a significant effect of racially marginalized background (F (1, 368) = 4.145, p = 0.042), suggesting that although perceptions of cost decreased overall, students from racially marginalized back-

grounds still experienced higher levels of cost (see Figure 5). The main effect of racially marginalized background can be best understood as an artifact of collapsing versions 4.0 (where nearly every chapter had racial differences in perceived cost) and 5.0 (where only a few chapters exhibited such differences). There were no significant interactions, $p > 0.05$.

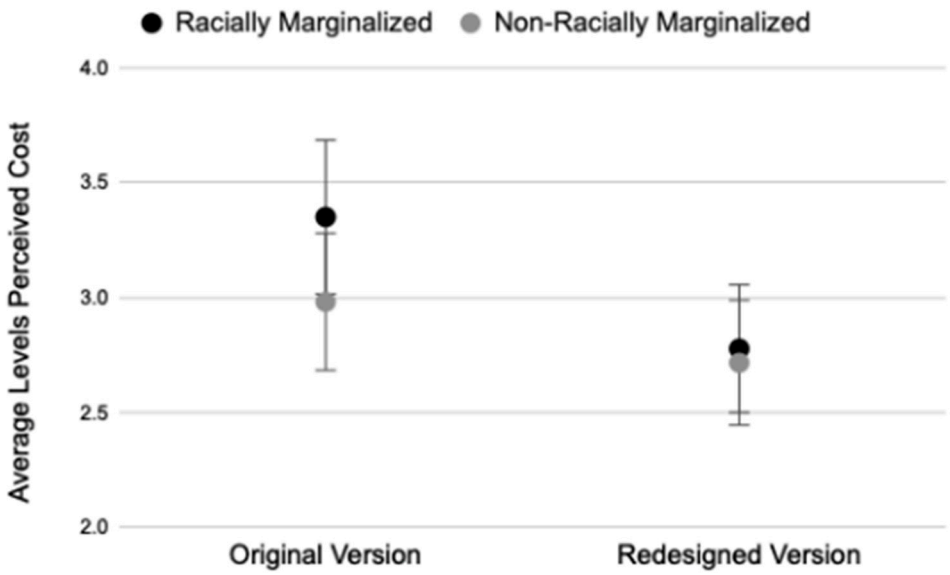


Figure 5. Average levels of perceived cost by textbook version and racially-marginalized background.

We also conducted an analysis of the redesigned version (5.0) alone. Like the original study, the mixed repeated measures MANOVA revealed no significant effect of racially marginalized status for expectancy (F (1, 98) = 0.605, p = 0.439), intrinsic value (F (1, 98) = 0.067, p = 0.797), or utility value (F (1, 98) = 0.057, p = 0.811). Unlike the original study, this time there was no statistically significant effect of racially marginalized status on perceptions of cost (F (1, 98) = 0.926, p = 0.338). There was a statistically significant interaction between chapter and racially marginalized status (F (40, 3920) = 1.412, p = 0.045).

Independent sample t-tests confirmed that the only remaining significant differences in perceptions of cost between racially marginalized students and non-racially marginalized students were at the beginning of chapters 2 and 4. There, racially marginalized students experienced significantly higher levels of cost than students from non-marginalized backgrounds (see Supplemental Table S2). This stands in contrast to the original textbook in which almost every chapter was perceived to be significantly more costly for racially marginalized students. In the comparison between versions 4.0 and 5.0, this manifested as a significant main effect of racially marginalized status on cost rather than an interaction between racially marginalized status and chapter.

The analysis of the redesigned results and the comparative data analysis revealed two encouraging findings. First, most of the disparate experiences related to perceptions of cost between racially marginalized and non-racially marginalized students, previously found in almost all of the chapters, were reduced in the redesigned textbook. Second, there were overall decreases in students' perceptions of cost, indicating that the changes made during the redesign process had a positive impact on students' perceptions of the textbook's workload, generally.

3.4.5. Discussion of Redesign

The final stage of our Plan-Do-Study-Act cycle included a thorough reflection on the prior cycle components and included discussions among the research team (the authors) and presentations of the findings to other developers, researchers, instructors, and students who work with CourseKata data. Although the data in this improvement cycle were collected in a quasi-experimental design, these findings provide promising evidence of

the potential benefits to subsequent generations of students who will use this textbook. Consequently, the decision was made to adopt the proposed changes (rather than to adapt or abandon them) and proceed with implementing them with all users of the textbook. We have plans to follow up and validate these results in the broader population of students using our textbook in other institutions with other instructors. (A summary of our implementation of the PDSA cycle, embedded in the Improvement Kata, is shown in Figure 6).

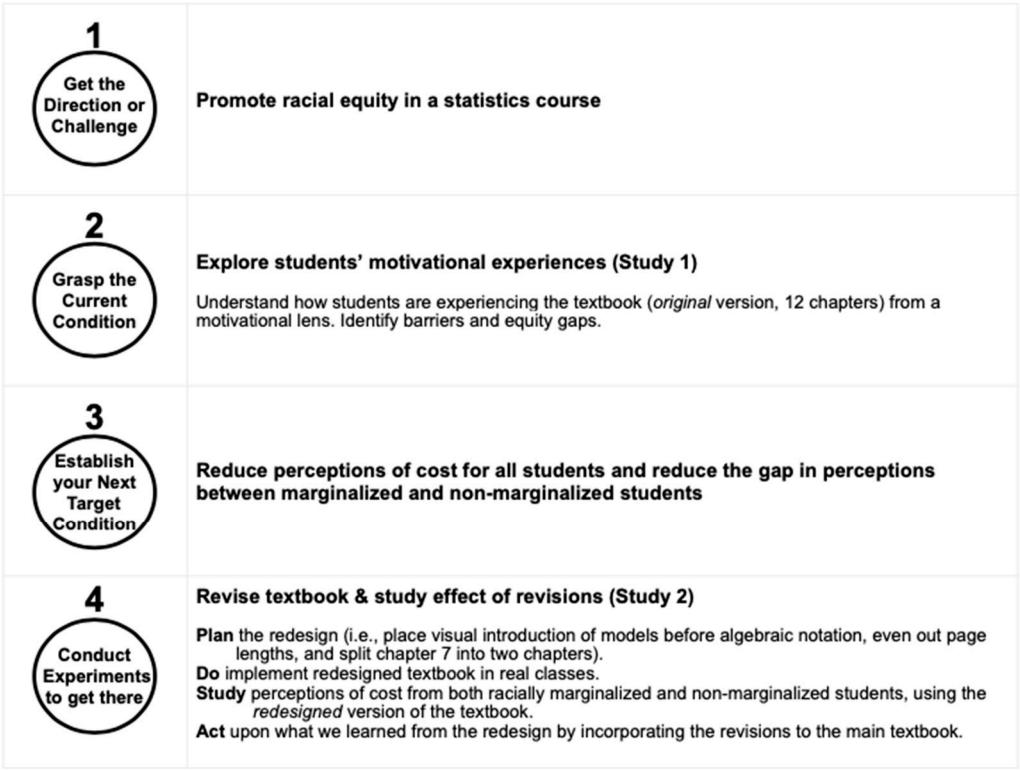


Figure 6. The Improvement Kata as described by Rother (2009, 2018), with a description of how each step was instantiated in the current studies.

A driving question during the 'act' process was: "What did we learn from doing this?" [32]. Engaging in this reflective exercise across a variety of stakeholders allowed us to recognize that certain textbook characteristics might have more substantial effects on students' perceptions of cost than previously anticipated. For instance, the length of the pages was almost invisible to the textbook authors who were concerned with coherence of the content on the page, but students expected pages to take a certain amount of time and may have found it discouraging when homework costs exceeded their expectations.

This realization led us to reconsider aspects of the textbook beyond the content itself, emphasizing the importance of how the material is experienced by students. The process of reflection triggered a cascade of new ideas. For example, because word counts are only a rough measure of the cost of a page, CourseKata's technology team developed new ways of measuring how long students spend on pages. In the future, this kind of metric will allow designers to adjust the combination of difficulty, word count, number of interactive components, and other design features. This may lead to a more consistent experience for students across the pages.

4. General Discussion

For decades, efforts to address disparities in students' experiences and outcomes in science related domains including statistics focused on trying to "fix students" through interventions targeted towards the individual student [38]. Such interventions—while

often well intended—assume deficits in students' preparation and aim to better align students with the existing higher education system. Through the collaborative efforts of the Better Book community (which includes students, instructors, researchers, curriculum designers, and technology developers), we engaged in a continuous improvement cycle. In this particular instance, we used this Plan-Do-Study-Act methodology towards improving equity in one component of a system that produces students' experiences, motivation, and learning—the textbook of an introductory statistics course. The research reported here is an attempt to apply the practices of improvement science toward the goal of equity through a motivational lens.

4.1. Original Version

The thorough examination of students' motivational beliefs allowed us to pinpoint potential hot spots or areas for improvement within the curriculum. One particularly noteworthy revelation from our exploration was the identification of a "hot-spot" in chapter 7. Upon closer examination, we found that it was the longest chapter in the textbook. By identifying where students perceived the content as particularly challenging and costly, we were able to zone in on critical aspects of the curriculum that required attention and refinement. Further, we were also able to identify who was experiencing the course and certain chapters as particularly challenging. The finding that students from racially marginalized backgrounds experienced higher levels of cost is in line with prior research [8]. This student-centered and data-driven approach ensured that our efforts for improvement were targeted and focused on addressing the specific needs of our diverse student population. Uncovering individual differences in students' experiences—particularly their perceptions of cost—resulted in research insights that led researchers and designers to further engage with teachers and students towards potential design solutions.

4.2. Redesigned Version

Because all students—but particularly students from racially marginalized backgrounds—experienced chapter 7 as a barrier to their course success, chapter 7 was redesigned and split into two more-manageable chapters. This revision and other design changes aimed to improve the learning experience and ensure that students could grasp the content effectively without feeling overwhelmed. When we returned to perceptions of cost post-redesign, we found that the differences between racially marginalized and non-racially marginalized students were reduced. Beyond reducing equity gaps, there were overall decreases in students' perception of cost, indicating that the changes made by focusing on marginalized students' experiences during the redesign process had a broader positive impact. In fact, mean levels of perceptions of cost for racially marginalized and non-racially marginalized students in the redesigned version of the textbook (v5.0) were both below the mean levels of perceptions of cost for the non-racially marginalized students in the pre-redesigned version (v4.0). Together, these results suggest that although the redesign did not reduce all the differences in motivational experiences between racially marginalized and non-racially marginalized students (i.e., students from racially marginalized backgrounds still experienced higher levels of costs initially and in one other chapter), the redesign may have contributed to a more equitable learning experience.

This study considers the potential role that textbook design plays in students' motivational experiences with a particular focus on the experiences of racially marginalized students. This exploration sheds light on potential structural barriers, such as page length, that can shape students' experiences, particularly their perceptions of cost. Whereas many may limit their conceptualization of structural barriers to societal-level challenges outside an individual's control, we argue that educators and institutions should consider their curricular choices and other small design decisions as having the power either to create new structural barriers or to mitigate the effects of existing ones. Addressing potential challenges in design and presentation of the textbook materials and how they impact students' motivation is crucial for identifying instructional improvements in the context of

teaching and learning. These improvements were made possible through collaborative efforts of diverse students, instructors, researchers, and curriculum designers in a continuous improvement process. Importantly, this work offers insights into how improvement science and translational research on motivation can be leveraged to guide us towards improving equity in a small system (i.e., the textbook of an introductory statistics course) and guiding the efforts of institutions of higher education in the pursuit of inclusive excellence.

Overall, creating spaces for racially marginalized students to describe their experiences and connecting those experiences to disaggregated data was vital to our improvement cycle. Even seemingly neutral expressions of their experience (noting that the pages took different amounts of time) eventually became design hypotheses. Because we knew from disaggregated data that perceptions of cost were a significant barrier, we took note when students mentioned page lengths and how they went through the textbook differently when they had more time. Also, what racially marginalized students say about their experiences may not, on the surface, seem like it has anything to do with race, ethnicity, or cultural background. Page lengths and chapter sequencing may seem far removed from a students' ethnic identity or socioeconomic status. However, we knew from other data collected via the textbook (in version 1.0) that students who attend institutions with higher proportions of racially marginalized students also report working more hours per week. Being able to have consistent homework experience (e.g., homework that takes roughly the same amount of time each week) might be particularly important to students who have other demands on their time, such as work or family commitments. Returning to our notion that racially marginalized students might serve as a canary in the coal mine, it is understandable how all students might benefit from more consistent expectations about the time it takes to complete their assignments, but those whose time is a more precious commodity would be the first ones to alert us of the problem.

5. Limitations and Future Directions

It is important to acknowledge that this study includes participants from two class sections taught by the same instructor at a single institution, limiting generalizability. We also do not suggest that adjusting textbook chapter and section lengths is the needed upgrade to improve equity in every classroom situation. Rather, methodologies such as the improvement kata may be generalizable to other educational settings. The process should be tailored to each setting [39], and its outcomes studied carefully for their broader applicability. Despite limits on generalizability, the current study provides evidence that redesigning systems to address the concerns of marginalized students may improve equity and reduce barriers for both marginalized students and their non-marginalized peers.

Following an approach consistent with the aims of improvement science, we hope in a future cycle to address the variability found in perceptions of cost found among racially marginalized students in the redesigned textbook. To reduce mean levels of perceived cost and to close the gap between subgroups of students is a start. But the long-term work of improvement science and translational research is to further reduce the variability in negative experiences (such as cost) for all students. Additional work will include an investigation of balancing measures to confirm that an improvement in perceived cost does not come at the sacrifice of performance in later chapters, for instance. We need to ensure that removing one barrier does not throw up new ones.

6. Conclusions

This study combines improvement science methodologies and research on motivation to nudge an educational system toward promoting equity. Our intensive exploration of students' motivational beliefs and experiences, comparing original and redesigned versions of a textbook, brought about valuable insights that led to targeted curriculum improvements. The identification of a cost-intensive chapter prompted a complete redesign, demonstrating our commitment to refining our learning materials based on data-driven hypotheses. The "Better Book" approach [24] guided by input from students, instructors,

researchers, curriculum designers, and technology developers is integral to our continuous improvement efforts. This collaborative approach and the underlying technology infrastructure ensures that changes to the textbook are evidence-based. It empowers us to see students' experiences through realistic time spans, respond to their evolving needs, and adjust the curriculum based on the latest research in educational psychology and pedagogy. By proactively adapting the instructional materials based on comprehensive feedback and data analysis, we create a more responsive and student-centered learning environment. Although the book is not ideal, there is a process for the book to get "better" and move towards providing a more equitable learning experience, where all students have the opportunity to thrive and develop their full academic potential.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/educsci14050487/s1, Table S1: Differences Between Racially Marginalized versus Non-Racially Marginalized Students for Perceptions of Cost at the Chapter Level for the Original Version (Version 4.0), Table S2: Differences Between Racially Marginalized versus Non-Racially Marginalized Students for Perceptions of Cost at the Chapter Level for the Redesigned Version (Version 5.0).

Author Contributions: Conceptualization, C.C.S., J.Y.S., M.C.J., J.W.S. and K.B.G.; methodology, C.C.S., M.C.J. and J.Y.S.; validation, C.C.S. and K.B.G. formal analysis, C.C.S.; investigation, C.C.S., M.C.J. and J.Y.S.; writing—original draft preparation, C.C.S., J.Y.S. and M.C.J.; writing—review and editing, K.B.G. and J.W.S.; visualization, C.C.S.; supervision, J.Y.S.; project administration, C.C.S., J.Y.S., J.W.S. and K.B.G.; funding acquisition, J.Y.S., J.W.S. and K.B.G. All authors have read and agreed to the published version of the manuscript.

Funding: This study has been made possible in part by a grant from the Chan Zuckerberg Initiative DAF, an advised fund of Silicon Valley Community Foundation (DRL-1229004), and a grant from the California Governor's Office of Planning and Research (OPR18115).

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board (or Ethics Committee) of the University of California, Los Angeles (IRB#20-001033 and 22 June 2020) for studies involving humans.

Informed Consent Statement: The {institution anonymized for review} IRB waived the requirement for informed consent, assent, and parental permission under 45 CFR 46.116(d) for the research. Students were notified at the outset of their course that their responses would be used for research purposes and that they could, at any time, opt to have their data excluded. Thus, the sample includes only students who consented to share their data for research purposes.

Data Availability Statement: Deidentified data pertaining to this study will be made publicly available on the Open Science Framework.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Tishkovskaya, S.; Lancaster, G.A. Statistical education in the 21st century: A review of challenges, teaching innovations and strategies for reform. *J. Stat. Educ.* **2012**, *20*, 1–56. [CrossRef]
2. Friedrich, J.; Buday, E.; Kerr, D. Statistical Training in Psychology: A national survey and commentary on undergraduate programs. *Teach. Psychol.* **2000**, *27*, 248–257. [CrossRef]
3. National Science Foundation, Division of Science Resources Statistics. *STEM Psychology as a Core Science, Technology, Engineering, and Mathematics Discipline*; National Science Foundation, Division of Science Resources Statistics: Alexandria, VA, USA, 2009. Available online: https://www.apa.org/pubs/reports/stem-discipline#:~:text=The%20National%20Science%20Foundation%20(NSF,Division%20of%20Science%20Resources%20Statistics, (accessed on 22 April 2024).
4. Mejia, M.C.; Rodriguez, O.; Johnson, H.; Perez, C.A. *A New Era of Student Access at California's Community Colleges*; Public Policy Institute of California: San Francisco, CA, USA, 2020. Available online: https://www.ppic.org/wp-content/uploads/a-new-era-of-student-access-at-californias-community-colleges-november-2020.pdf (accessed on 22 April 2024).
5. Murtonen, M.; Lehtinen, E. Difficulties experienced by education and sociology students in quantitative methods courses. *Stud. High. Educ.* **2003**, *28*, 171–185. [CrossRef]
6. Riegle-Crumb, C.; King, B.; Irizarry, Y. Does STEM stand out? Examining Racial/Ethnic gaps in persistence across postsecondary fields. *Educ. Res.* **2019**, *48*, 133–144. [CrossRef]

7. Sutter, C.C.; Hulleman, C.S.; Givvin, K.B.; Tucker, M.C. Utility value trajectories and their relationship with behavioral engagement and performance in introductory statistics. *Learn. Individ. Differ.* **2022**, *93*, 102095. [CrossRef]
8. Sutter, C.C.; Givvin, K.B.; Hulleman, C.S. Concerns and challenges in introductory statistics and correlates with motivation and interest. *J. Exp. Educ.* **2023**, 1–30. [CrossRef]
9. Eccles, J.S.; Wigfield, A. From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemp. Educ. Psychol.* **2020**, *61*, 101859. [CrossRef]
10. Barron, K.E.; Hulleman, C.S. Expectancy-value-cost model of motivation. *Psychology* **2015**, *84*, 261–271. [CrossRef]
11. Schau, C.; Emmioğlu, E. Do introductory statistics courses in the United States improve students' attitudes? *Stat. Educ. Res. J.* **2012**, *11*, 86–94. [CrossRef]
12. Starr, C.R.; Hunter, L.; Dunkin, R.; Honig, S.; Palomino, R.; Leaper, C. Engaging in science practices in classrooms predicts increases in undergraduates' STEM motivation, identity, and achievement: A short-term longitudinal study. *J. Res. Sci. Teach.* **2020**, *57*, 1093–1118. [CrossRef]
13. Van Es, C.; Weaver, M.M. Race, sex, and their influences on introductory statistics education. *J. Stat. Educ.* **2018**, *26*, 48–54. [CrossRef]
14. Barber, P.H.; Shapiro, C.; Jacobs, M.S.; Avilez, L.; Brenner, K.I.; Cabral, C.; Cebreros, M.; Cosentino, E.; Cross, C.; Gonzalez, M.L.; et al. Disparities in Remote Learning Faced by First-Generation and Under- represented Minority Students During COVID-19: Insights and Opportunities From a Remote Research Experience. *J. Microbiol. Biol. Educ.* **2021**, *22*, 1–25. [CrossRef]
15. Seo, E.; Lee, Y. Stereotype threat in high school classrooms: How it links to teacher mindset climate, mathematics anxiety, and achievement. *J. Youth Adolesc.* **2021**, *50*, 1410–1423. [CrossRef] [PubMed]
16. Totonchi, D.A.; Perez, T.; Lee, Y.; Robinson, K.A.; Linnenbrink-Garcia, L. The role of stereotype threat in ethnically minoritized students' science motivation: A four-year longitudinal study of achievement and persistence in STEM. *Contemp. Educ. Psychol.* **2021**, *67*, 102015. [CrossRef] [PubMed]
17. Bowman, N.; Logel, C.; Lacosse, J.; Canning, E.A.; Emerson, K.T.; Murphy, M.C. The Role of Minoritized Student Representation in Promoting Achievement and Equity Within College STEM Courses. *AERA Open* **2023**, *9*, 23328584231209957. [CrossRef]
18. Dietrich, J.; Viljaranta, J.; Moeller, J.; Kracke, B. Situational expectancies and task values: Associations with students' effort. *Learn. Instr.* **2017**, *47*, 53–64. [CrossRef]
19. Sutter, C.C.; Totonchi, R.A.; DeCoster, J.; Barron, K.E.; Hulleman, C.S. How does expectancy-value-cost motivation vary during a semester? An intensive longitudinal study to explore individual and situational sources of variation in statistics motivation. *Learn. Individ. Differ. Accepted for publication*.
20. Harackiewicz, J.M.; Canning, E.A.; Tibbetts, Y.; Priniski, S.J.; Hyde, J.S. Closing achievement gaps with a utility-value intervention: Disentangling race and social class. *J. Personal. Soc. Psychol.* **2016**, *111*, 745–765. [CrossRef]
21. Harackiewicz, J.M.; Tibbetts, Y.; Canning, E.A.; Hyde, J.S. Harnessing values to promote motivation in education. In *Advances in Motivation and Achievement*; Emerald Publishing Ltd.: Bingley, UK, 2014; pp. 71–105. [CrossRef]
22. Hulleman, C.S.; Harackiewicz, J.M. The utility-value intervention. In *Handbook of Wise Interventions: How Social Psychology Can Help People Change*; Walton, G.M., Crum, A.J., Eds.; The Guilford Press: New York, NY, USA, 2021; pp. 100–125.
23. Tibbetts, Y.; Harackiewicz, J.M.; Canning, E.A.; Boston, J.S.; Priniski, S.J.; Hyde, J.S. Affirming independence: Exploring mechanisms underlying a values affirmation intervention for first-generation students. *J. Personal. Soc. Psychol.* **2016**, *110*, 635–659. [CrossRef] [PubMed]
24. Stigler, J.W.; Son, J.Y.; Givvin, K.B.; Blake, A.B.; Fries, L.; Shaw, S.T.; Tucker, M.C. The Better Book approach for education research and development. *Teach. Coll. Rec.* **2020**, *122*, 1–32. [CrossRef]
25. Gallimore, R.; Santagata, R. Researching teaching: The problem of studying a system resistant to change. In *American Psychological Association eBooks*; American Psychological Association: Washington, DC, USA, 2006; pp. 11–28. [CrossRef]
26. Son, J.Y.; Stigler, J.W. Statistics and Data Science: A Modeling Approach. (2017–2023). Available online: https://coursekata.org/ (accessed on 22 April 2024).
27. Fries, L.; Son, J.Y.; Givvin, K.B.; Stigler, J.W. Practicing connections: A framework to guide instructional design for developing understanding in complex domains. *Educ. Psychol. Rev.* **2021**, *33*, 739–762. [CrossRef]
28. Zhang, I.Y.; Gray, M.E.; Cheng, A.X.; Son, J.Y.; Stigler, J.W. Representational-mapping strategies improve learning from an online statistics textbook. *J. Exp. Psychol. Appl.* 2023, *Advance online publication*. [CrossRef]
29. LeMahieu, P.G.; Grunow, A.; Baker, L.; Nordstrum, L.E.; Gomez, L.M. Networked improvement communities: The discipline of improvement science meets the power of networks. *Qual. Assur. Educ.* **2017**, *25*, 5–25. [CrossRef]
30. Lewis, C. What is improvement science? Do we need it in education? *Educ. Res.* **2015**, *44*, 54–61. [CrossRef]
31. Rother, M. *Toyota Kata*; McGraw-Hill Professional Publishing: New York, NY, USA, 2009.
32. Rother, M. *The Toyota KATA Practice Guide*; McGraw-Hill Education: New York, NY, USA, 2018; Volume 2022.
33. McNair, T.B.; Bensimon, E.M.; Malcom-Piqueux, L. *From Equity Talk to Equity Walk: Expanding Practitioner Knowledge for Racial Justice in Higher Education*; John Wiley & Son: Hoboken, NJ, USA, 2020.
34. De Vega, M.; Glenberg, A.; Graesser, A. *Symbols and Embodiment: Debates on Meaning and Cognition*; Oxford University Press: Oxford, UK, 2012. [CrossRef]
35. Goldstone, R.L.; Barsalou, L.W. Reuniting perception and conception. *Cognition* **1998**, *65*, 231–262. [CrossRef] [PubMed]

36. Cooper, L.L.; Shore, F.S. Students' misconceptions in interpreting center and variability of data represented via histograms and stem-and-leaf plots. *J. Stat. Educ.* **2008**, *16*, 1–14. [CrossRef]
37. Kaplan, J.J.; Gabrosek, J.G.; Curtiss, P.; Malone, C. Investigating Student Understanding of Histograms. *J. Stat. Educ.* **2014**, *22*, 1–30. Available online: https://www.tandfonline.com/doi/abs/10.1080/10691898.2014.11889701 (accessed on 22 April 2024). [CrossRef]
38. Hatfield, N.; Brown, N.; Topaz, C.M. Do introductory courses disproportionately drive minoritized students out of STEM pathways? *PNAS Nexus* **2022**, *1*, pgac167. [CrossRef]
39. Ramirez, G.; Covarrubias, R.; Jackson, M.C.; Son, J.Y. Making hidden resources visible in a minority serving college context. *Cult. Divers. Ethn. Minor. Psychol.* **2021**, *27*, 256–268. [CrossRef]