

Research Article

Attitudinal changes in face-to-face and online statistical reasoning learning environments

Daniel A. Showalter

Eastern Mennonite University, USA

Attitudes toward statistics play an important role in statistical understanding, postsecondary decisions, and a lifelong relationship with statistics. Unfortunately, the average undergraduate student tends to view statistics as less interesting and less valuable after completing an introductory statistics course. The product of several decades of statistics education reform, a statistical reasoning learning environment (SRLE) has shown some positive early results in cognitive domains and may impact attitudes as well. In this study, four classes of introductory undergraduate statistics (two fully online, two face-to-face) were designed as SRLEs. Students ($n = 83$) completed a pretest and posttest version of the Survey of Attitudes Toward Statistics-36[®]. Both online and face-to-face sections showed average gains in Interest and Value that were higher than those reported in a large reference group, and these gains were practically significant.

Keywords: Statistics education research; Attitudes; Online education; Undergraduate statistics

Article History: Submitted 12 February 2021; Revised 11 May 2021; Published online 15 May 2021

1. Introduction

Undergraduate statistics courses for non-majors are an odd creature. First, consider the potential of the content itself. There are growing amounts of data (Baraniuk, 2011; Thayne, 2016) and computational power to analyze data with freely-available, user-friendly software (Nolan & Lang, 2010). Essentially anything that students are interested in can be analyzed through a statistical lens (Bergen, 2016; Gould, 2010; Libman, 2010; Stern et al., 2020), and widespread calls have been made to statistics educators to incorporate these student-relevant data into their classrooms and assessments while making the material more accessible to those without strong mathematical backgrounds (Chew & Dillon, 2014; Engel, 2017; Gould, 2017; Hall, 2011; Lesser, 2007; Neumann et al., 2013). Moreover, statistics is one of the most in-demand and sought after skills on the job market (Bombaci-Bilgin et al., 2020).

With all of this, statistics should be one of the highlights in a college student's academic path. Yet, the reality is that non-statistics majors often lose interest in statistics (Ramirez et al., 2012; Swanson et al., 2014), value it less (Ramirez et al., 2012; Swanson et al., 2014), and have a more

Address of Corresponding Author

Daniel A. Showalter, PhD., Eastern Mennonite University, 1200 Park Rd, United States.

✉ daniel.showalter@emu.edu

ORCID [0000-0003-1769-3088](https://orcid.org/0000-0003-1769-3088)

How to cite: Showalter, D. A. (2021). Attitudinal changes in face-to-face and online statistical reasoning learning environments. *Journal of Pedagogical Research*, 5(2), 1-18. <http://dx.doi.org/10.33902/JPR.2021269257>

negative affect towards it (Murtonen & Lehtinen, 2003) by the time they finish an introductory course. These negative attitudinal outcomes have been linked to lower achievement (Emmioğlu & Capa-Aydin, 2012; Tishkovskaya & Lancaster, 2012) and a damaged lifelong relationship with statistics (Leavy et al., 2013; Ramirez et al., 2012; Sowe, 2020). The widespread rejection of scientific findings during the COVID-19 pandemic has illuminated the potentially lethal consequences of a citizenry with a damaged relationship with statistics. What's wrong, and what can we do about it?

In this study, a promising approach to undergraduate statistics education was implemented in four class sections of introductory statistics. This approach, known as a statistics reasoning learning environment (SRLE), has shown potential to counter the national trends of negative attitudinal outcomes. If students can see the relevance of the material to their personal and professional lives, it stands to reason that they will be more likely to view statistics as valuable and interesting. A validated instrument was used to assess the change in students' attitudes towards statistics from the beginning to the end of the SRLE.

1.1. Literature Review

This section begins with an overview of the research demonstrating the importance of student attitudes toward statistics in an introductory statistics course and beyond. It then discusses several studies that describe the current situation of what attitudes students have towards statistics and how these attitudes shift during an introductory statistics course. Results of research initiatives to improve attitudes in statistics courses are covered. Finally, a brief history is given on how the statistics education reform movement of the past decades has culminated in statistical reasoning learning environments, which may hold potential for positively impacting student attitudes toward statistics.

1.1.1. Importance of attitudes toward statistics

Statistics education is a relatively new field, and much of the initial research on it focused on cognitive performance outcomes (Tishkovskaya & Lancaster, 2012). However, even as early as 1980, some researchers began to suspect that attitudes toward statistics could be important as well (Roberts & Bilderback, 1980). By the turn of the century, these suspicions had developed into a more structured exhortation to educators and researchers to begin paying attention to students' attitudes toward statistics. Gal et al. (1997) authored a chapter filled with illustrative quotes from students about their experiences with statistics. These quotes framed a narrative about how attitudes and cognitive outcomes were inextricably linked.

This relationship between attitudes and cognitive performance was formalized in a meta-analysis conducted by Emmioğlu and Capa-Aydin (2012). The researchers identified 17 studies that included measures of both attitudes toward statistics and statistics achievement. In particular, they looked at studies which measured statistics attitudes with the *Survey of Attitudes Toward Statistics*® (SATS; Schau, 2003) in the domains of *Affect*, *Cognitive Competence*, *Value*, and *Difficulty*. *Affect* measured the feelings students held towards statistics with items such as "I will feel (felt) insecure when I have to do statistics problems" and "I am scared by statistics." *Cognitive Competence* measured student attitudes towards how their knowledge could be applied to statistics with items such as "I can learn statistics" and "I will find (found) it difficult to understand statistical concepts." *Value* measured how useful, relevant, and worthwhile students viewed statistics both personally and professionally with items such as "Statistical thinking is not applicable in my life outside my job" and "Statistics should be a required part of my professional training." *Difficulty* measured how difficult students viewed statistics as a subject with items such as "Statistics is a complicated subject" and "Statistics involves massive computations." It is worth noting that, while most instructors would probably prefer that their students have higher scores on the first three domains, there might be less agreement on what the ideal score is on the *Difficulty* subscale. The 17 studies included a total of over 4,000 students from around the U.S. and Europe. Emmioğlu and Capa-Aydin found that statistics achievement had a moderately-sized positive

correlation with both *Affect* and *Cognitive Competence* and a small positive correlation with *Value* and *Difficulty*. Zimmerman and Austin (2018) extended these results using a version of the *Statistics Anxiety Rating Scale* (STARS; Cruise et al., 1985) with a sample of around 1,000 students enrolled in an introductory statistics course. They found that both *Self-Concept* and *Worth of Statistics* were significant predictors of a student's final grade. In a much smaller pilot study ($n = 41$) that also made use of STARS, Bourne and Nesbit (2018) found suggestive evidence that anxiety towards the statistics requirement of a psychology major may cause some students to choose a different major. This is not hard to fathom when even student teachers preparing to be high school mathematics teachers have been shown to avoid teaching statistics due to its perceived difficulty (Leavy et al., 2013).

But perhaps the most concerning consequence of attitudes toward statistics in an introductory statistics course—either positive or negative—is the foundation it lays for a student's lifelong relationship with statistics (Ramirez et al., 2012). In a literature review on challenges, innovations, and strategies in statistics education over two decades, Tishkovskaya and Lancaster (2012, p. 2) denote as “most critical” the “fact that [statistical courses] affect life-long perceptions of and attitudes toward the value of statistics for many students, and hence many future employees, employers, and citizens.”

1.1.2. Current landscape of attitudes toward statistics

Given the importance of attitudes toward statistics to course-level, program-level, and lifelong outcomes, it is useful to examine the distribution of attitudes toward statistics. One of the most widely used assessment tools in this domain is the aforementioned *Survey of Attitudes Toward Statistics*® (SATS; Schau, 2003). Candace Schau developed the original multiple-response survey to measure self-reported attitudes toward statistics with subscales for *Affect*, *Cognitive Competence*, *Difficulty*, and *Value* as described in the previous subsection. She later revised the survey to include *Effort*, *Interest*, and some global attitude items. *Effort* measured the amount of work the student planned to (did) invest in learning statistics, and *Interest* measured the student's level of interest in statistics. Each subscale composite value is an average of several items measured on a 7-point Likert scale.

Schau and Emmioğlu (2012) administered the SATS-36 to a sample of roughly 2,200 U.S. university students both before and after the students had taken an introductory course in undergraduate statistics. They found that average *Affect*, *Cognitive Competence*, and *Difficulty* all increased slightly from pretest to posttest (4.16 to 4.30, 4.94 to 5.03, and 3.75 to 3.90, respectively). In other words, over the course of their statistics class, students (on average) gained a more positive attitude towards statistics, grew more confident about their ability to apply statistical knowledge, and felt that statistics was more difficult than they had originally anticipated. Meanwhile, *Value*, *Interest*, and *Effort* decreased (5.04 to 4.72, 4.51 to 4.00, and 6.32 to 5.84, respectively), indicating that (on average) students grew less likely to view statistics as useful, relevant, worthwhile, and interesting and also reported investing less effort into learning statistics than they had thought they would. The researchers pointed out that these results are not necessarily representative. By the mere fact that instructors agreed to participate in the study, it was likely that they were already more motivated to learn about, and impact, student attitudes than the average statistics instructor; thus, U.S. university students in general are likely to experience even more sobering attitudinal changes during an introductory statistics course. These types of results are not limited to instructors using traditional curricula; similar results were found on a sample of 425 students whose instructors used a randomization-based curriculum (Swanson et al., 2014) and in a study where the same instructor taught both lecture- and discussion-based versions of the same course (Bateiha et al., 2020).

To better understand the point during an introductory statistics courses at which these (mostly discouraging) changes take place, Kerby and Wroughton (2017) administered the SATS-36 to a sample of 292 undergraduate students at three points in time: after the first day of class, at the

midpoint of the semester, and during the final week of the course. From pretest to posttest, the attitudinal changes were similar—but modestly better—than those reported by Schau and Emmiöglu (2012), ranging from 0.04 more in average *Interest* gain to 0.25 more in average *Cognitive Competence* gain. However, this “improvement” still represented a decrease in average *Interest* of 0.46 points, from 4.57 to 4.11. The researchers discovered that, although the mean changes in the subscale scores were greater in the first half of the semester, the majority of individual student scores varied in both halves of the semester. Lawton and Taylor (2020) took an even closer look at when the student attitudes change by asking how engaging the class was at the end of each class. Although their self-designed attitude scale can’t be compared directly to the SATS-36, they found that (a) the most drastic changes occurred near the beginning and end of the course, and (b) the instructor consistently rated classes as more engaging than the average student.

The landscape of attitudinal changes toward statistics appears even graver for certain groups of students and for students in particular types of courses. Dierker et al. (2016) surveyed 333 students across four years of a semester-long multidisciplinary, project-based statistics course. They found that 31.9% of Black and Hispanic students ($n = 74$) rated the course as difficult while only 11.0% of their peers ($n = 259$) did. Similarly, only 76.7% of the Black or Hispanic students reported a desire to take a follow-up course, whereas 83.7% of the peers did. On the other hand, Black and Hispanic students were more likely to report an increased interest in conducting research (42.3% vs. 29.1%). Another study conducted across three years of an introductory statistics course at Cornell University ($n = 611$) found no significant differences by race in average SATS-28 subscales scores, but did find that females reported significantly lower *Affect* and *Cognitive Competence* averages than their male counterparts, despite the fact that the instructor for all sections was female (van Es & Weaver, 2018). Opstad (2020) found somewhat different results when administering the SATS-36 to a sample of 140 business students enrolled in a macroeconomics course. In his study, males and females reported similar scores on *Affect* and *Cognitive Competence*, although the males did find statistics more interesting, more valuable, and easier than the females. These differences held up even after controlling for personal traits and mathematical ability, but were substantially weaker.

The degree to which an introductory statistics course operates online may also be a factor related to attitudes. Gundlach et al. (2015) analyzed data on a sample of 462 students who chose to register in a traditional ($n = 331$), fully online ($n = 75$), or a flipped ($n = 56$) version of an introductory statistics class. Of these 462 students, 261 (56%) completed both a pretest and posttest version of the SATS. Despite all sections being taught by the same instructor, the traditional students had statistically significant gains beyond what their online counterparts did in *Affect* (+0.42), *Cognitive Competence* (+0.61), and *Interest* (+0.27), although their perception of *Difficulty* also increased (+0.33). In another study, researchers examined outcome differences between 605 students across six campuses randomly assigned to either a traditional introductory statistics course with 163 minutes of face-to-face instruction per week or a hybrid one with 70 minutes of face-to-face instruction per week (Bowen et al., 2014). Although performance on a standardized assessment of conceptual understanding was similar between the groups, 11% fewer of the students in the hybrid sections reported liking their course. There are some indications that studies like these don’t capture the full difference in attitudes between traditional and hybrid or fully online courses due to higher dropout rates in courses that have substantial online components. For example, Tu (2014) began a hybrid introductory statistics class with 40 students. Despite using real-life data and modeling her instructional design on evidence-based principles, 32 (80%) of the students had dropped out or stopped engaging by the end of the course—a dropout rate she claimed was typical of online hybrid statistics courses at her institution.

In summary, the research literature paints a bleak picture of the impact of introductory statistics courses on the typical student. With few exceptions, studies have found that undergraduate students tend to leave their introductory statistics course less interested in statistics and viewing the content as less valuable for their lives than when they started the course.

1.1.3. Research on improving affective outcomes

Given the importance of attitudes toward statistics and the current landscape of what those attitudes are in various contexts and for different groups of students, educators and researchers have made substantial efforts to improve student attitudes toward statistics.

In a review of literature on statistics anxiety, a construct interrelated with attitudes, Chew and Dillon (2014) found that discouraging procrastination through regular assessment, using humor, and reducing emphasis on mathematical computations were three effective ways of improving attitudes. For example, inserting fun items into the readings of randomly-assigned students increased average SATS *Interest* levels by 0.20 more than the control group (Lesser et al., 2016). Herman (2020) redesigned her statistics class to emphasize connections between statistics and the real world. Example activities included “laptop days” where students paired up to analyze authentic data on topics of interest. Herman then administered the SATS-36 to her 96 students before and after the redesigned course. She found significant average gain scores in *Cognitive Competence* (+0.25), *Affect* (+0.48), and *Difficulty* (+0.27), but minor drops in *Value* (-0.09) and *Interest* (-0.25).

The use of more serious topics in statistics classes has also shown some potential for improving attitudes. Science education researchers have found it motivating to show students, especially first generation college students, ways in which the material they are learning can help people make a positive impact on their community and society (Allen et al., 2015). Statistics educators have made efforts in this direction by using statistics to raise awareness and reflect on social injustices (Bergen, 2016; Lesser, 2007; Lesser, 2010). These efforts help ensure that students exit an introductory statistics course viewing statistics as more than “a disconnected collection of theorems and plug-and-chug recipes.” (Lesser, 2007, p. 8). Although integrating statistics and social justice efforts can bring its own set of challenges (see, for example, Garii & Appova, 2013), it is likely much easier for teachers to do in statistics courses than in mathematics courses in general (Showalter, 2013).

Studies such as the ones reported on in this subsection provide hope for techniques that could improve student attitudes in an introductory college statistics course. However, some researchers have expressed concern that these isolated changes are not enough to impact the multi-dimensional field of statistics education that spans areas such as culture, pedagogy, content, and assessment; they call instead for a more comprehensive shift (Ben-Zvi et al., 2018). The following subsection describes a classroom learning environment that has evolved over the past couple decades and which is steadily becoming central to the field of statistics education.

1.1.4. Statistical reasoning learning environments

Partly in response to what they saw as an overemphasis on computation and a lack of statistical reasoning, and partly due to the way in which John Tukey’s exploratory style of data analysis supplanted the need for probability-centered statistics (Carver et al., 2016), statistics educators began calling for reformation in the final decades of the 20th century. These calls culminated in a focus group report detailing a new way of teaching statistics that emphasized statistical thinking, real data, and active learning (Cobb, 1992). A decade later, these recommendations evolved into a set of design principles addressing five aspects of a statistics class environment essential for developing students’ statistical reasoning: core statistical ideas, instructional activities, classroom activity structure, computer-based tools, and classroom discourse (Cobb & McClain, 2004). These design principles were then formalized into six recommendations of the original Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report (Aliaga et al., 2005), which quickly became the cornerstone for the undergraduate statistics education reform movement.

To give statistics educators a more tangible framework, Garfield and her colleagues (2008) developed the idea of a statistical reasoning learning environment (SRLE) that embodied Cobb and McClain’s (2004) design principles and the original GAISE College Report. They describe an SRLE

as “an effective and positive statistics classroom that develops in students a deep and meaningful understanding of statistics and helps students develop their ability to think and reason statistically” (Garfield & Ben-Zvi, 2009, p. 73). The use of the term “learning environment” underscores the idea that this interrelated set of design principles is more than an isolated technique or tool. The six pillars of an SRLE are as follows:

- (1) Help students develop an understanding of the central ideas in statistics.
- (2) Use authentic, motivating data as much as possible.
- (3) Engage students with inquiry-based activities centered on interesting problems.
- (4) Train students to use technological tools that empower them.
- (5) Prioritize classroom discourse above lecture.
- (6) Assess understanding of central ideas using measures aligned with learning goals.

These same six pillars appear in the executive summary of a second GAISE College Report (Carver et al., 2016), which updated and revised the recommendations of the original report. Endorsed by the American Statistical Association, both of these reports have become a guide for many statistics educators from Pre-K through the postsecondary level (Wood et al., 2018).

The idea of an SRLE is clearly positioned in the direction statistics education is moving, but it is less clear how to implement it in practice. Ben-Zvi et al. (2018) devoted a chapter of the *International Handbook of Research in Statistics Education* to exploring ways for educators to implement an SRLE through a variety of design perspectives, theoretical frameworks, and levels of schooling. They provide some examples of early findings from researchers on SRLE implementation but caution that the complex nature of a learning environment will require a more concerted research effort.

Although an SRLE would plausibly have some impact on student attitudes, the existing research has focused mainly on either the cognitive impact of an SRLE (e.g., Conway et al., 2019; Hidayah et al., 2015) or on the attitudinal impact of isolated elements of an SRLE (Neumann et al., 2013). One notable exception is an article about a graduate education course that includes some participant quotes about positive attitudinal impact of both the face-to-face and online version of the course (Garfield & Everson, 2009). Another was a study with 280 pre-service teachers enrolled in a GAISE-centered course (not explicitly an SRLE) that found positive average change on all six of the SATS subscales (Leavy et al., 2019). The current study addresses a gap in the literature by examining the impact of implementing an SRLE on student attitudes.

1.1.5. Research questions

Although research studies often focus on one aspect of pedagogy at a time, the broad nature of a statistical reasoning learning environment requires a more holistic, even if messier, approach (Ben-Zvi et al., 2018). This study addressed the following “messy” questions, in order to contribute to the discussion about how an SRLE could potentially impact student attitudes in face-to-face and online environments.

RQ1: What changes are observed in student attitudes from the beginning to the end of a face-to-face SRLE-inspired course?

RQ2: What changes are observed in student attitudes from the beginning to the end of an online asynchronous SRLE-inspired course?

2. Method

2.1. Description of Course

The study took place at a small liberal arts university in the southeastern region of the United States. Four semester-long 3-credit-hour class sections of an elementary statistics course for non-math majors were included in the study. Two of these sections were conducted fully online and asynchronous, and two were conducted face-to-face. All four sections took place during the 2017–2018 school year. The author was the sole instructor for each of the four sections.

The course was designed to follow Garfield et al.'s (2008) statistical reasoning learning environment (SRLE), which centers on Cobb and McClain's (2004) principles of instructional design. Although an SRLE is a holistic learning model rather than a collection of practices or techniques, some specific examples of how the course adhered to the SRLE principles are provided below to give the reader a more accurate picture. A more complete description of the assessment aspects of the course can be found in Showalter (2019).

2.1.1. Help Students develop an understanding of the central ideas in statistics

Both the online and face-to-face sections were reverse-engineered starting with the central concepts described in the GAISE College Report (Carver et al., 2016) and similar statistics education articles. For example, the weekly online modules included "The Power and Beauty of Statistics", "Collecting and Organizing Data", "Decisionmaking: Data Vs. Anecdotes", "Correlation and Causation", and "Data Visualization" in addition to more traditional topics such as "The Middle" and "Variability." In the face-to-face sections, the same topics were interwoven into a more conventional sequencing that followed David Moore's (2010) *Basic Practice of Statistics* textbook.

2.1.2. Use authentic, motivating data as much as possible

Most modern statistics textbooks use real datasets, but the data—and the way in which they are presented—are not always motivating. Leading up to the semester in which the SRLEs were implemented, students in previous statistics sections had been surveyed regularly about their interests so that the datasets used in the course would be relevant. Despite the deluge of available data, finding and preparing meaningful datasets to address these student interests was the most time-consuming part of the course design. In addition to student-generated data, datasets covered topics such as food, exercise, music, faith, race, gender, social media, sports, social justice, and mental health.

2.1.3. Use authentic, motivating data as much as possible

Both the online and face-to-face sections had weekly spreadsheet assignments where students had the freedom to either choose their own variables within a dataset or even choose their own dataset and then follow an investigative path to answer a set of questions. Examples included making a motion chart to compare their high school with rival high schools on demographic indicators, creating and analyzing a world map of gender inequity, setting up a Monte Carlo simulation, and running a regression on a nationally-representative longitudinal dataset to see what could be predicted about someone in their mid-twenties based on data from their sophomore year of high school.

2.1.4. Train students to use technological tools that empower them

Students were required to use Google Sheets for most labs but also had optional labs in which they could learn other software packages (e.g., SPSS, Excel, JASP, Tableau, Data Studio). These labs were incentivized by reducing the weight of the final exam, although students often enjoyed learning them for the sake of discovery (and perhaps a resume boost!). With both required and optional labs, the focus was on developing students' abilities to "think with data" (Horton et al., 2014).

2.1.5. Prioritize classroom discourse above lecture

In the face-to-face sections, each class began with a kickstarter, which was usually a thought-provoking data visualization like those found on Nathan Yau's *Flowing Data* (<https://flowingdata.com>), David McCandless's *Information is Beautiful* (<https://informationisbeautiful.net>), or the NY Times' *What's Going On in This Graph?* ([nytimes.com/column/whats-going-on-in-this-graph](https://www.nytimes.com/column/whats-going-on-in-this-graph)). In randomly-assigned partners, and later as a class, students discussed what messages were being presented, the potential for bias, and how

the information could be used. In the online sections, classroom discourse was more difficult to stimulate. Attempts were made at discussion forums, discussion partners, and shared journals, and these were successful for some students.

2.1.6. Assess understanding of central ideas using measures aligned with learning goals

To obtain student buy-in, Garfield et al. (2008) emphasize that the assessments must be aligned with the learning goals. In the face-to-face sections, midterms were replaced with frequent "Authentic Quizzes" that aligned with the ways in which statistical reasoning was being developed throughout the course. During the first 20 minutes of a quiz day, students were randomly placed with a discussion partner and the instructor would lead a discussion on a topic of interest ("Out of school children", "Measuring poverty", "Racial discrimination in hiring", "Eating habits and depression", "The ethics of 'mathing' students", "Embracing diversity", "Sex and gender", "Identity and suicide", and "A better life"). For the remaining 30 minutes, students individually took a quiz where they were provided with actual data on the topic. They then analyzed and interpreted the data, often combining their life experience with the contents of the 20-minute discussion. In the online class, weekly quizzes included similar types of questions on a smaller scale.

2.2. Description of Course Participants

There were 61 total students enrolled in the two face-to-face sections. Of these, 3 withdrew and 5 did not complete the posttest survey for an analytic sample of 53 students. In the two online sections, there were 40 total students. Of these, 1 withdrew, 3 did not complete the pretest, and 6 did not complete the posttest, yielding an analytic sample of 30 students. Within the face-to-face analytic sample, there were a total of 14 (26%) students who reported having no parent with a bachelor's degree (referred to in this study as "first generation college students"); 33 females (62%); 11 (21%) students identifying as Black or Hispanic; and 2 (4%) students above the age of 23. Within the online analytic sample, there were a total of 13 (43%) students who reported having no parent with a bachelor's degree; 24 females (80%); 5 (17%) students identifying as Black or Hispanic; and 6 (20%) students above the age of 23.

2.3. Instruments

To assess attitudinal changes, the *Survey of Attitudes Toward Statistics-36*® (SATS-36; Schau 2003) instrument was used. The SATS-36 includes 36 items measured on a 7-point scale that are then used to form a composite measure of six subscales. The SATS-36 posttest is nearly identical to the pretest, with the wording of some items slightly adjusted to make grammatical sense. The subscales are *Affect* (e.g., "I will like (like) stats"), *Cognitive Competence* (e.g., "I can learn statistics"), *Value* (e.g., "I use statistics in my everyday life"), *Difficulty* (e.g., "Statistics is a complicated subject"), *Interest* ("I am interested in using statistics"), and *Effort* ("I plan to work hard (worked hard) in my statistics course"). In assessing several instruments that measure attitudes toward statistics, Nolan et al. (2012) found the SATS to have the strongest evidence of construct validity and reliability. Some concerns have been raised about the six-factor structure and whether certain items should be excluded (Hommik & Luik, 2017; Persson et al., 2019). However, given the debatable nature of these concerns (Xu & Schau, 2019), and the value of being able to position the results of this study within the bulk of the research literature, no changes were made to the SATS-36.

Students in all four sections completed the SATS-36 as an anonymous online Qualtrics survey. Students in all four sections were required to complete the SATS-36 by the end of the first week of class and then again in the final week of class before taking the final exam. Regardless of their participation in the study, students were required to make a sincere attempt for both the pretests and the posttests; approximately 1% of their overall grade was based on completion of the surveys in an effort to encourage robust participation.

2.4. Data Analysis

In order to examine the gains (or losses) in attitudes and understanding, paired difference gain scores were analyzed for the SATS-36. Although the SATS-36 items are measured on an ordinal scale, the researcher treated the composite subscale scores as quantitative for several reasons.

First, the majority of the research using the SATS-36, including the largest study to date (Schau & Emmioğlu, 2012), has reported means and standard deviations. Following this convention more accurately positions the results of the current study within the landscape of the research literature. Second, although the intervals between measurement points are not exactly equal, it is reasonable to view them as approximately equal. Moreover, the composite nature of the six subscales makes this assumption even stronger as they change from Likert-type to Likert scale data (Boone & Boone, 2012). Finally, the assumptions of the Wilcoxon signed-rank test (a nonparametric equivalent of the paired *t*-test), would not be met even if the data were treated as ordinal, because the analysis would involve taking the sums of individual items to form the composite measures.

In light of this measurement limitation, boxplots are provided in the results section to show the medians and interquartile spreads. Moreover, the reader is invited to request from the author a deidentified version of the original data for separate analysis.

There is disagreement over how familywise Type I error rate should be controlled when analyzing the SATS; this study followed the conservative recommendation to use a Bonferroni correction on an .05 alpha level of significance (Millar & Schau, 2010). Practical significance was defined as an average gain score of at least 0.50 points (Millar et al., 2013). Only seven students were missing data on either the pretest or posttest, and no students were missing data on more than one item on a test. These missing data points were imputed with the mean for the subscale on which the data were missing.

3. Results

3.1. Summary Results

Pretest, posttest, and gain score means and standard deviations for the face-to-face and online sections are provided in Table 1 for each of the six SATS attitude subscales. A more detailed analysis of these scores, along with boxplots of the respective distributions of subscale scores, can be found in the following subsections.

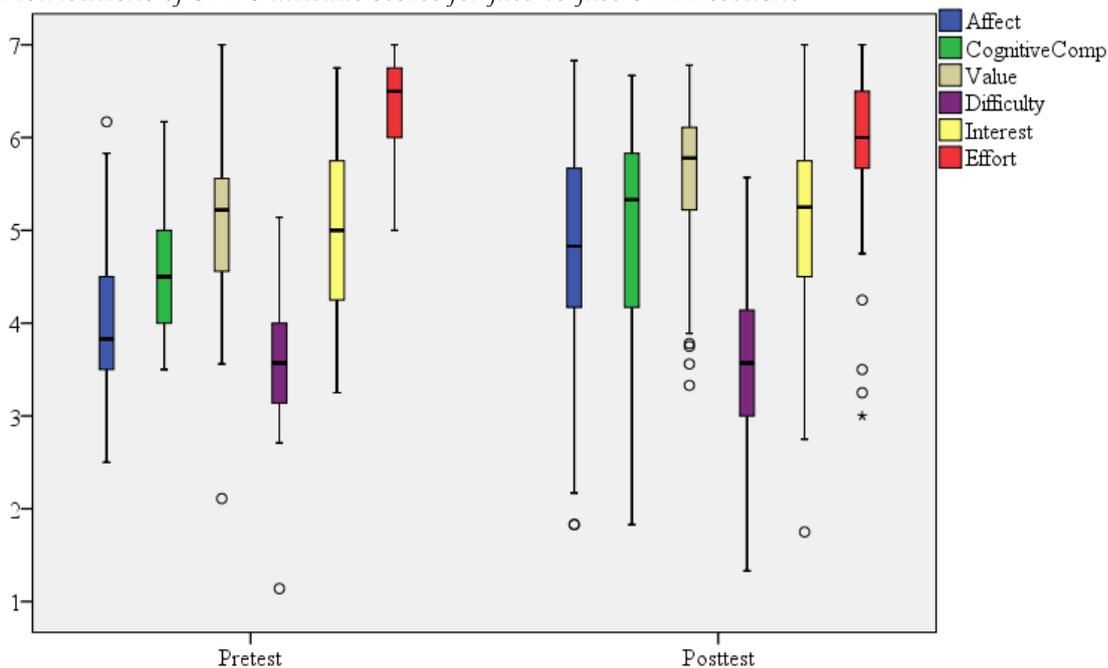
3.2. Findings Regarding Changes in the Face-To-Face Sections

The boxplots in Figure 1 show the distribution of scores on each SATS subscale for the pretest and posttest in the two face-to-face SRLE sections. Within the analytic sample, the median subscale score increased on *Affect* (3.83 to 4.83), *Cognitive Competence* (4.50 to 5.78), *Value* (5.22 to 5.78), and *Interest* (5.00 to 5.25). It remained constant for *Difficulty* (3.57) and decreased on *Effort* (6.50 to 6.00). Except for *Difficulty*, posttest medians were all above a neutral score of 4. So, although students ended the SRLE course feeling, on average, that statistics was a difficult subject, they had positive overall attitudes toward statistics. Range increased on all subscales except *Value*, and the interquartile range increased most on *Affect* and *Cognitive Competence*. This tells us that the changes in average student attitudes were not experienced uniformly; though most students had a more positive attitude and felt more competent to do statistics by the end of the course, some students followed the prevalent trend of feeling even more negative towards statistics and less competent in the subject than when they started.

Table 1
Means and standard deviations for the six subscales on the SATS-36

	<i>n</i>	Pretest	Posttest	Mean Gain
Affect				
Face-to-face	53	4.03 (0.80)	4.67 (1.25)	0.64 (1.27)
Online	30	4.13 (0.90)	4.33 (1.02)	0.19 (1.19)
Cognitive Competence				
Face-to-face	53	4.58 (0.73)	5.01 (1.25)	0.43 (1.14)
Online	30	4.87 (0.67)	5.01 (0.79)	0.14 (0.90)
Value				
Face-to-face	53	5.13 (0.79)	5.56 (0.85)	0.43 (0.79)
Online	30	5.30 (0.83)	5.71 (0.71)	0.40 (0.73)
Difficulty				
Face-to-face	53	3.58 (0.64)	3.63 (0.94)	0.05 (0.89)
Online	30	3.42 (0.64)	3.44 (0.60)	0.02 (0.63)
Interest				
Face-to-face	53	5.06 (0.91)	5.13 (1.06)	0.07 (1.08)
Online	30	4.88 (1.01)	5.21 (0.91)	0.33 (0.89)
Effort				
Face-to-face	53	6.37 (0.56)	5.94 (0.93)	-0.43 (0.86)
Online	30	6.50 (0.46)	6.11 (0.81)	-0.39 (0.58)

Figure 1
Distributions of SATS attitude scores for face-to-face SRLE sections



Note: Outliers are depicted as open dots and extreme outliers as asterisks

Table 2 displays the average change in attitudes for students in the face-to-face SRLE sections. Using a gain score of 0.50 as the threshold for practical significance (Millar et al., 2013), only the gains in *Affect* would be practically significant when comparing the face-to-face students' posttest scores to their pretest scores. However, when comparing the mean gains reported by students in the face-to-face sections with the large national reference group from Schau and Emmioğlu's (2012) study, there were statistically and practically significant gains on *Affect*, *Value*, and *Interest*.

Table 2

Mean attitude gain scores for face-to-face SRLE sections and reference group

	Mean gain score (reference)	Mean gain score (SRLE)	95% CI for difference in mean gain score	<i>p</i> -value
Affect	0.13	0.64	(0.02, 0.99)	.006*
Cognitive Competence	0.10	0.43	(-0.10, 0.76)	.042
Value	-0.32	0.43	(0.45, 1.06)	< .001*
Difficulty	0.15	0.05	(-0.44, 0.23)	.407
Interest	-0.50	0.07	(0.16, 0.98)	< .001*
Effort	-0.48	-0.43	(-0.28, 0.38)	.676

Note. Reference group is Schau and Emmioğlu's (2012) sample of 2,000+ university students.

*Due to a Bonferroni correction, only *p*-values below .0083 are considered statistically significant.

The average pretest score on *Affect* for the face-to-face sections in this study was 4.03, which is 0.13 lower than the 4.16 reported by Schau and Emmioğlu (2012). This means that they had more room for growth than the reference sample, and thus regression to the mean may mute the reported difference in mean gain scores slightly (Millar & Schau, 2010). Similarly, compared to the reference group, the pretest means for the face-to-face students in this study were 0.09 higher on *Value* and 0.55 higher on *Interest*, suggesting that any regression to the mean would only strengthen the effects observed here. In other words, although the face-to-face students in the SRLE sections did not see significant gains in *Value* or *Interest*, they did experience significantly more growth in these areas than the national reference group, which itself is probably a more optimistic sample than the population of all undergraduate statistics students (Schau & Emmioğlu, 2012).

At the individual student level, only 1 of the 53 (1.7%) students in the face-to-face SRLE sections dropped from having a positive (above 4.5) score on the *Value* subscale to a neutral or negative score. This contrasts with Schau and Emmioğlu's (2012) observation that 25% of section *means* dropped from a positive to a neutral/negative score on *Value*.

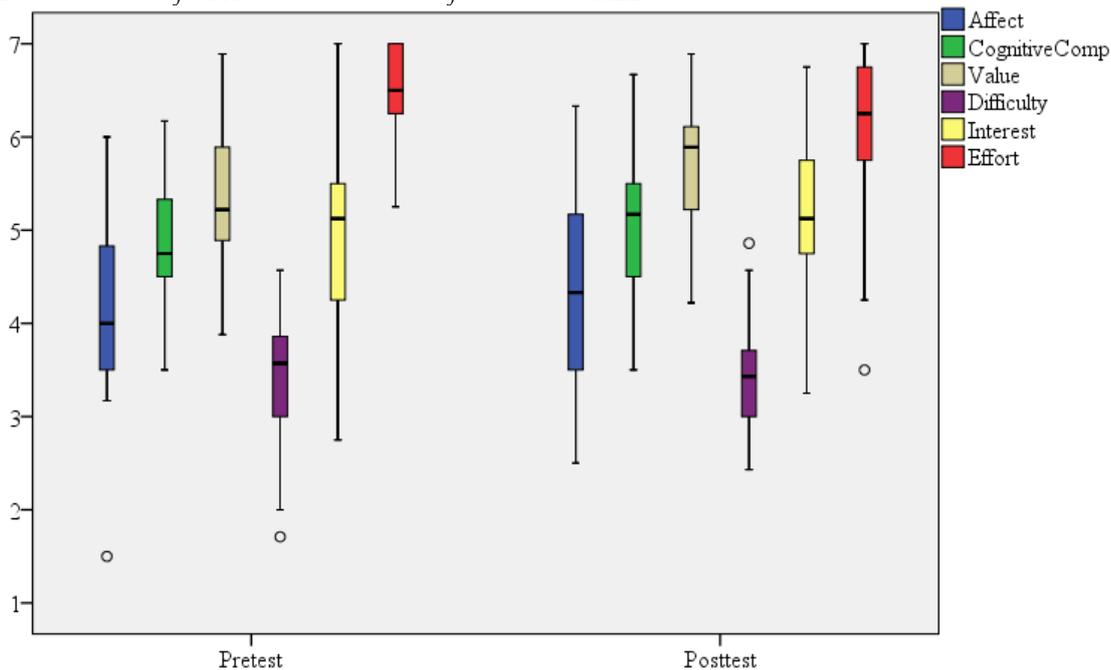
3.3. Findings Regarding Changes in the Online Sections

The boxplots in Figure 2 show the distribution of scores on each SATS subscale for the pretest and posttest in the two online SRLE sections. Within the analytic sample, the median subscale score increased on *Affect* (4.00 to 4.33), *Cognitive Competence* (4.75 to 5.17), and *Value* (5.22 to 5.89). It remained constant for *Interest* (5.13) and decreased on *Difficulty* (3.57 to 3.43) and *Effort* (6.50 to 6.25). Except for *Difficulty*, posttest medians were all above a neutral score of 4. As with the students in the face-to-face SRLE sections, although the online students ended the SRLE course feeling, on average, that statistics was a difficult subject, they had positive overall attitudes toward statistics. Range increased most notably on *Effort*, and the interquartile ranges for the pretest subscales were similar to those for the posttest. Thus, students in the online sections experienced attitudinal changes somewhat more uniformly than their counterparts in the face-to-face sections.

Table 3 displays the average change in attitudes for students in the online SRLE sections. None of the average gain scores would be considered practically significant at the 0.50 level. However, when comparing the mean gains reported by students in the online sections with the large national reference group from Schau and Emmioğlu's (2012) study, there were statistically and practically significant gains on *Value* and *Interest*. The average *Interest* gain was even higher for the online sections than it was for the face-to-face (0.33 for online and 0.07 for face-to-face compared to the reference group -0.50), and the average *Value* gain was similar (0.40 for online and 0.43 for face-to-face compared to the reference group -0.32).

Figure 2

Distributions of SATS attitude scores for online SRLE sections



Note: Outliers are depicted as open dots

Table 3

Mean attitude gain scores for online SRLE sections and reference group

	Mean gain score (reference)	Mean gain score (SRLE)	95% CI for difference in mean gain score	<i>p</i> -value
Affect	0.13	0.19	(-0.56, 0.68)	.772
Cognitive Competence	0.10	0.14	(-0.43, 0.51)	.790
Value	-0.32	0.40	(0.34, 1.11)	< .001*
Difficulty	0.15	0.02	(-0.46, 0.20)	.281
Interest	-0.50	0.33	(0.36, 1.29)	< .001*
Effort	-0.48	-0.39	(-0.22, 0.39)	.433

Note. Reference group is Schau and Emmioğlu's (2012) sample of 2,000+ university students.

*Due to a Bonferroni correction, only *p*-values below .0083 are considered statistically significant.

As with the face-to-face students, the online students in the SRLE sections may not have experienced significant gains in *Value* or *Interest*, but they did experience significantly more growth in these areas than the national reference group. Moreover, compared to the reference group, the pretest means for the online students in this study were 0.26 higher on *Value* and 0.37 higher on *Interest*, suggesting that any regression to the mean would likely only strengthen the effects observed here.

At the individual student level, only 1 of the 30 (3.3%) students in the face-to-face SRLE sections dropped from having a positive (above 4.5) score on the *Value* subscale to a neutral or negative score. This contrasts with Schau and Emmioğlu's (2012) observation that 25% of section *means* dropped from a positive to a neutral/negative score on *Value*.

4. Discussion and Limitations

The statistics education research literature has numerous studies where student interest in statistics decreased through taking a course in introductory statistics (Bateiha et al., 2020; Bond et al., 2012; Gundlach et al., 2015; Kerby & Wroughton, 2017; Paul & Cunnington, 2017; Schau & Emmioğlu, 2012; Swanson et al., 2014). This is unacceptable in an era when an aversion to statistics can serve as a lifelong barrier to the increased quality of life and citizenship offered by competent

statistical reasoning. Moreover, nearly all of these same studies reveal a loss in students' perceived value of statistics during the introductory statistics course. The more that society's collective trust in the utility of statistics wanes, the more we sink into the warring anecdotes that polarized us in the crucial early stages of the COVID-19 pandemic. The current study adds to the handful of studies that suggest potential ways for statistics educators to counter these discouraging trends within an introductory statistics course (e.g., Bayer, 2016; Lesser et al., 2016; Neumann et al., 2013). Statistical reasoning learning environments have already shown promising results for increasing students' understanding of statistical concepts (Conway et al., 2019; Hidayah et al., 2019; Wei Chan et al., 2015). This study suggests that they may also help maintain student interest in, and value of, statistics.

4.1. Value and Interest within the SRLE

What was it about the SRLE sections that led students to perceive statistics as more valuable? Among his recommendations for designing statistics courses that are valued by minority students, Davidson (2007) suggests using examples of social importance. Discussions on issues like racial discrimination, gender equality, and poverty, followed up by student conjectures and data investigations (as per Garfield and Ben-Zvi's (2009) description of an SRLE) appeared to engage all the students, but particularly the Black and Hispanic students. Perhaps most valuable was the followup; after the data and discussions had illuminated the severity of certain injustices, the spreadsheet assignments provided tools to embolden students to create and communicate their own perspective on the injustices. As one student described it, "I thought that this class would be a huge waste of time ... just one of those dumb things you had to do to get all the required general ed courses in. I completely changed my mind after the first couple of spreadsheet assignments. The spreadsheets were challenging sometimes, but I always felt really accomplished after I was done and I also felt like I learned a useful tool."

In terms of interest, a primary factor was the kickstarter session to open each face-to-face class. These ranged from 5 to 25 minutes long, but they hooked even the texting-prone students in the class through controversy, challenged beliefs, and beautiful data visualizations. More importantly, the questions that arose from, and perpetuated, these discussions had no "right" answer and were thus more welcoming (Garfield & Ben-Zvi, 2009). The online sections had a similar element with a provoking article or data visualization followed by a discussion forum, although anecdotal student comments suggest that the personal reflections on these elements were more valuable than the discussions. These discussions were a time investment, and that time was made possible by following Cobb's (1992) advice to remove certain mathematical barriers in order to focus on the core statistical ideas. In the evaluations, an online student mentioned that "the most valuable part of the course was that I learned things about statistics, not just numbers and formulas." It was clear that students were surprised (mostly pleasantly) upon realizing that they could engage with rich statistical concepts with only basic mathematics.

But these individual elements of the course are insufficient in capturing the full experience. Just as most people would struggle to define completely why they value their closest friend, a learning environment is more than the sum of its parts – precisely the point made by Ben-Zvi et al. (2018).

4.2. Discussion on Face-to-face vs. Online SRLE Sections

While students across all four sections tended to like statistics more (*Affect*) by the end of the course, this gain was much more prominent among the face-to-face sections than it was in the online sections. At the same time, students in the online sections experienced higher average increases in how interested (*Interest*) they were in statistics. What can explain this apparent contradiction of liking stats more but being less interested in it? First, note that the finding runs against the halo effect often observed when attitudes are self reported. Second, since all four sections had the same instructor, the instructor effect should be relatively constant (unless, perhaps, one views the online presence of an instructor as a different entity than the face-to-face version).

Ironically, at the item level of the *Affect* subscale, there was a greater average gain for “liking” statistics among the online students than those in the face-to-face sections (although the reader is cautioned that single Likert-type items do not enjoy the same quantitative properties as the Likert scale composite scores analyzed in the Results section). At the same time, the two largest gain differences between the two groups across all 36 items were both on the *Affect* subscale: (a) the face-to-face students felt less frustrated than they expected when going over statistics tests whereas the online students felt more frustrated, and (b) the face-to-face students felt much less overall stress in the statistics course than they had expected whereas the online students felt about the same amount as what they had expected.

The first difference can be explained by the fact that the instructor rarely reviewed the weekly quizzes with the online sections aside from an occasional piece of feedback in a class-wide email. It could be argued that this was a lapse in fidelity with the SRLE principle of aligning assessments with learning objectives. The second difference was likely connected to the spreadsheet assignments; when online students ran into an issue, they often wrestled with it for hours (a stressful experience!) whereas the face-to-face students would often ask the instructor at the beginning of the following class. On the *Interest* subscale, online students reported higher gains than the face-to-face students on all four items—interest in understanding, learning, using, and communicating statistics. This could be related to the maturity that comes with age: Only 4% of the face-to-face students were over the age of 23, whereas 20% of the online students were. As Emmioğlu (2011) found in her dissertation, age tends to correlate positively with the *Interest* subscale.

One final difference worth highlighting is that the face-to-face students grew much less likely to say that statistics has no applications in their respective professions whereas the online students were much less likely to say that statistics is not useful to the typical professional. This may indicate that the online students bought into the idea that statistics was indeed useful in general but that the face-to-face students were more likely to take this a step further and claim that statistics was useful *to them*.

Despite the differences between the face-to-face and online sections, there were strong similarities as well. Both groups averaged around an entire point lower on the items, “I will have (had) no idea of what’s going on in this statistics course” and “I am scared by statistics.” In other words, both groups carried into the course some fear and confusion towards statistics that was relieved as they engaged in the course.

4.3. Limitations

Millar and Schau (2010) describe seven statistical issues that can arise when measuring attitudinal changes: the Likert scale nature of the composite scores can truncate the distribution, data are not independent, self-reported attitude scores only approximate the true attitude, gain scores depend on pretest scores, there is plausible risk of regression to the mean, the six-subscale nature of the SATS opens up the potential for inflated Type I error rates, and there are often missing data to account for.

In this study, the third issue is mitigated by the way the SATS uses composite scores to measure each subscale; the fourth and fifth issues would likely only mute the gains observed in *Value* and *Interest*, due to the pretest averages being higher than in Schau and Emmioğlu’s (2012) study; the sixth issue was addressed by applying a conservative Bonferroni correction to account for familywise Type I error; and the seventh was a minor issue since none of the participants were missing data on more than one item. However, the first two issues remain as limitations for the results of this study with the lack of independence being particularly troublesome for generalization.

Additional major limitations include the potential selection bias incurred by the fact that students were allowed to select their own section, the low sample size (especially within the online class sections), survivorship bias (i.e., how different were the 83 students who completed the

course and surveys from the 13 who did not?), the ordinal nature of the survey items, and the inability of the researcher to measure or account for any instructor effect (Xu, 2019). Moreover, unlike statistical significance, effect size is independent of sample size, and so it is quite possible that the large effect sizes found in this study would not hold up on a larger scale. Collectively, these limitations reduce the impact of this study to a cautiously optimistic discussion starter of the potential attitudinal changes that occur when implementing an SRLE. At the same time, they underscore the need for larger-scale research that examines the attitudinal impacts of various ways of implementing an SRLE.

5. Conclusion

In general, after learning introductory statistics within a statistical reasoning learning environment, students were more likely to view statistics as interesting and valuable than what is typically found in undergraduate introductory statistics classrooms. This finding held true whether the student experienced the SRLE in a face-to-face classroom or online. One quote that summed up many students' experience with the SRLE came in the end-of-course evaluations. When asked to describe the most valuable aspect of the course, the student simply wrote, "Learning that I love stats."

Even more encouraging is that, in the two years since the courses took place, several of the students have contacted the instructor in excitement about a data visualization they came across, a media claim they were able to critique in an informed way, or a story about using their statistical knowledge to better their life. This excitement seems to serve as a magnet drawing them towards anything statistics-related that they may stumble across in their personal or professional lives. This is precisely the sort of approach needed in our modern society to develop, and enjoy, a higher level of statistical reasoning.

Acknowledgements. This work was supported by the National Science Foundation under Grant #1611713.

Disclosure statement. No potential conflict of interest was reported by the author.

References

- Aliaga, M., Cobb, G., Cuff, C., Garfield, J., Gould, R., Lock, R., ... & Velleman, P. (2005). *Guidelines for assessment and instruction in statistics education (GAISE) (College Report)*. <http://www.amstat.org/education/gaise>
- Allen, J. M., Muragishi, G. A., Smith, J. L., Thoman, D. B., & Brown, E. R. (2015). To grab and to hold: Cultivating communal goals to overcome cultural and structural barriers in first-generation college students' science interest. *Translational Issues in Psychological Science*, 1(4), 331-341. <https://doi.org/10.1037/tps0000046>
- Baraniuk, R. G. (2011). More is less: Signal processing and the data deluge. *Science*, 331(6018), 717-719. <https://doi.org/10.1126/science.1197448>
- Bayer, T. J. (2016). *Effects of guided project-based learning activities on students' attitudes toward statistics in an introductory statistics course* (Doctoral dissertation). Old Dominion University, USA. https://digitalcommons.odu.edu/stemps_etds/14
- Bergen, S. (2016, July). *Melding data with social justice in undergraduate statistics and data science courses*. Paper presented at the IASE Roundtable Conference 2016: Promoting Understanding of Statistics about Society, Berlin, Germany.
- Bateiha, S., Marchionda, H., & Autin, M. (2020). Teaching style and attitudes: a comparison of two collegiate introductory statistics classes. *Journal of Statistics Education*, 28(2), 154-164. <https://doi.org/10.1080/10691898.2020.1765710>
- Ben-Zvi, D., Gravemeijer, K., & Ainley, J. (2018). Design of statistics learning environments. In D. Ben-Zvi, K. Makar & Joan Garfield (Eds.) *International handbook of research in statistics education* (pp. 473-502). Springer, Cham.

- Bombaci-Bilgin, A. A., Bulger, D., & Fung, T. (2020). Statistics: Your ticket to anywhere. *Statistics Education Research Journal*, 19(1), 11-20.
- Bond, M. E., Perkins, S. N., & Ramirez, C. (2012). Students' perceptions of statistics: An exploration of attitudes, conceptualizations, and content knowledge of statistics. *Statistics Education Research Journal*, 11(2), 6-25.
- Boone, H. N., & Boone, D. A. (2012). Analyzing Likert data. *Journal of Extension*, 50(2), 1-5.
- Bourne, V. J., & Nesbit, R. J. (2018). Do attitudes towards statistics influence the decision to study psychology at degree level? A pilot investigation. *Psychology Teaching Review*, 24(2), 55-63.
- Bowen, W. G., Chingos, M. M., Lack, K. A., & Nygren, T. I. (2014). Interactive learning online at public universities: Evidence from a six-campus randomized trial. *Journal of Policy Analysis and Management*, 33(1), 94-111. <https://doi.org/10.1002/pam.21728>
- Carver, R., Everson, M., Gabrosek, J., Horton, N., Lock, R., Mocko, M., ... & Wood, B. (2016). *Guidelines for assessment and instruction in statistics education: College Report 2016*. https://www.amstat.org/asa/files/pdfs/GAISE/GaiseCollege_Full.pdf
- Chew, P. K. H., & Dillon, D. B. (2014). Statistics anxiety update: Refining the construct and recommendations for a new research agenda. *Perspectives on Psychological Science*, 9(2), 196-208. <https://doi.org/10.1177/1745691613518077>
- Cobb, G. (1992). Teaching statistics. In L. A. Steen (Ed.), *Heeding the call for change: Suggestions for curricular action* (pp. 3-43). The Mathematical Association of America.
- Cobb, P., & McClain, K. (2004). Principles of instructional design for supporting the development of students' statistical reasoning. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 375-395). Springer.
- Conway, B., IV, Martin, W. G., Strutchens, M., Kraska, M., & Huang, H. (2019). The statistical reasoning learning environment: A comparison of students' statistical reasoning ability. *Journal of Statistics Education*, 27(3), 171-187. <https://doi.org/10.1080/10691898.2019.1647008>
- Cruise, R. J., Cash, R. W., & Bolton, D. L. (1985). Development and validation of an instrument to measure statistical anxiety. *American Statistical Association Proceedings of the Section on Statistical Education*, 4(3), 92-97.
- Davidson, M. A. (2007). *Understanding the burdens of race at a predominantly white university: The experiences of underrepresented students in an introductory statistics course* (Unpublished doctoral dissertation). The Ohio State University, Columbus, OH.
- Dierker, L., Alexander, J., Cooper, J. L., Selya, A., Rose, J., & Dasgupta, N. (2016). Engaging diverse students in statistical inquiry: A comparison of learning experiences and outcomes of under-represented and non-underrepresented students enrolled in a multidisciplinary project-based statistics course. *International Journal for the Scholarship of Teaching and Learning*, 10(1), 2. <https://doi.org/10.20429/ijsofl.2016.100102>
- Emmioğlu, E. (2011). *A structural equation model examining the relationships among mathematics achievement, attitudes toward statistics, and statistics outcomes* (Unpublished doctoral dissertation). Middle East Technical University, Ankara, Turkey.
- Emmioğlu, E., & Capa-Aydin, Y. (2012). Attitudes and achievement in statistics: A meta-analysis study. *Statistics Education Research Journal*, 11(2), 95-102.
- Engel, J. (2017). Statistical literacy for active citizenship: A call for data science education. *Statistics Education Research Journal*, 16(1), 44-49.
- Gal, I., Ginsburg, L., & Schau, C. (1997). Monitoring attitudes and beliefs in statistics education. *The assessment challenge in statistics education*, 12, 37-51.
- Garfield, J. B., Ben-Zvi, D., Chance, B., Medina, E., Roseth, C., & Zieffler, A. (2008). Creating a statistical reasoning learning environment. In J. B. Garfield & D. Ben-Zvi (Eds.) *Developing Students' Statistical Reasoning* (pp. 45-63). Springer.
- Garfield, J., & Ben-Zvi, D. (2009). Helping students develop statistical reasoning: Implementing a statistical reasoning learning environment. *Teaching Statistics*, 31(3), 72-77. <https://doi.org/10.1111/j.1467-9639.2009.00363.x>
- Garfield, J., & Everson, M. (2009). Preparing teachers of statistics: A graduate course for future teachers. *Journal of Statistics Education*, 17(2). <https://doi.org/10.1080/10691898.2009.11889516>
- Garii, B., & Appova, A. (2013). Crossing the great divide: Teacher candidates, mathematics, and social justice. *Teaching and Teacher Education*, 34, 198-213. <https://doi.org/10.1016/j.tate.2012.07.004>
- Gould, R. (2010). Statistics and the modern student. *International Statistical Review*, 78(2), 297-315. <https://doi.org/10.1111/j.1751-5823.2010.00117.x>

- Gould, R. (2017). Data literacy is statistical literacy. *Statistics Education Research Journal*, 16(1), 22–25. <https://doi.org/10.1111/j.1751-5823.2010.00117.x>
- Gundlach, E., Richards, K.A.R., Nelson, D., & Levesque-Bristol, C. (2015). A comparison of student attitudes, statistical reasoning, performance, and perceptions for web-augmented traditional, fully online, and flipped sections of a statistical literacy class. *Journal of Statistics Education*, 23(1). <https://doi.org/10.1080/10691898.2015.11889723>
- Hall, J. (2011). Engaging teachers and students with real data: Benefits and challenges. In C. Batanero, G. Burrill, & C. Reading (Eds.) *Teaching statistics in school mathematics-challenges for teaching and teacher education* (pp. 335–346). Springer.
- Herman, J. (2020). Student attitudes in a real-world inspired second statistics course. In Marchionda, H. and Bateiha, S. (Eds.). *Proceedings of the 48th Annual Meeting of the Research Council on Mathematics Learning* (pp. 85-92). Virtual.
- Hidayah, N., Wahyudin, W., & Turmudi, T. (2019, September). Building ability of sport students statistical reasoning through statistical reasoning learning environment (SRLE). Paper presented at *3rd International Conference on Sport Science, Health, and Physical Education (ICSSHPE 2018)*. Atlantis Press.
- Hommik, C., & Luik, P. (2017). Adapting the Survey of Attitudes Toward Statistics (SATS-36) for Estonian secondary school students. *Statistics Education Research Journal*, 16(1), 228-239.
- Horton, N. J., Baumer, B. S., & Wickham, H. (2014). Teaching precursors to data science in introductory and second courses in statistics. *arXiv preprint arXiv:1401.3269*.
- Kerby, A. T., & Wroughton, J. R. (2017). When do students' attitudes change? Investigating student attitudes at midterm. *Statistics Education Research Journal*, 16(2), 476–486.
- Lawton, S., & Taylor, L. (2020). Student perceptions of engagement in an introductory statistics course. *Journal of Statistics Education*, 28(1), 45–55. <https://doi.org/10.1080/10691898.2019.1704201>
- Leavy, A. M., Hannigan, A., & Fitzmaurice, O. (2013). If you're doubting yourself then, what's the fun in that? An exploration of why prospective secondary mathematics teachers perceive statistics as difficult. *Journal of Statistics Education*, 21(3). <https://doi.org/10.1080/10691898.2013.11889684>
- Leavy, A. M., Hourigan, M., Murphy, B., & Yilmaz, N. (2019). Malleable or fixed? Exploring pre-service primary teachers' attitudes towards statistics. *International Journal of Mathematical Education in Science and Technology*, 52(3), 1–25. <https://doi.org/10.1080/0020739X.2019.1688405>
- Lesser, L. M. (2007). Critical values and transforming data: Teaching statistics with social justice. *Journal of Statistics Education*, 15(1). <https://doi.org/10.1080/10691898.2007.11889454>
- Lesser, L. M. (2010). The necessity of equity in teaching statistics. <http://www.math.utep.edu/Faculty/lesser/LesserPMEJ.pdf>
- Lesser, L. M., Pearl, D. K., & Weber III, J. J. (2016). Assessing fun items' effectiveness in increasing learning of college introductory statistics students: Results of a randomized experiment. *Journal of Statistics Education*, 24(2), 54–62. <https://doi.org/10.1080/10691898.2016.1190190>
- Libman, Z. (2010). Integrating real-life data analysis in teaching descriptive statistics: A constructivist approach. *Journal of Statistics Education*, 18(1). <https://doi.org/10.1080/10691898.2010.11889477>
- Murtonen, M., & Lehtinen, E. (2003). Difficulties experienced by education and sociology students in quantitative methods courses. *Studies in Higher Education*, 28(2), 171–185. <https://doi.org/10.1080/0307507032000058064>
- Millar, A. M., White, B. G. J., & Romo, R. (2013). Comparing apples with apples: Assessing student attitudes in the presence of regression to the mean. *Proceedings of the Joint Statistical Meetings*. <http://ec.msvu.ca/xmlui/bitstream/handle/10587/1284/JSM2013millar.pdf>
- Millar, A. M., & Schau, C. (2010). Assessing students' attitudes: The good, the bad, and the ugly. 2010 Joint Statistical Meetings - Section on Statistical Education, 1133-1143. Vancouver, Canada: American Statistical Association. <http://www.statlit.org/pdf/2010MillarSchauASA.pdf>
- Moore, D. S. (2010). *The basic practice of statistics*. W. H. Freeman and Company.
- Neumann, D.L., Hood, M., & Neumann, M.M. (2013). Using real-life data when teaching statistics: Student perceptions of this strategy in an introductory statistics course. *Statistics Education Research Journal*, 12(2), 59–70.
- Nolan, M. M., Beran, T., & Hecker, K. G. (2012). Surveys assessing students' attitudes toward statistics: A systematic review of validity and reliability. *Statistics Education Research Journal*, 11(2), 103-123.
- Nolan, D., & Temple Lang, D. (2010). Computing in the statistics curricula. *The American Statistician*, 64(2), 97–107. <https://doi.org/10.1198/tast.2010.09132>

- Opstad, L. (2020). Attitudes towards statistics among business students: Do gender, mathematical skills and personal traits matter? *Sustainability*, 12(15), 6104. <https://doi.org/10.3390/su12156104>
- Paul, W., & Cunningham, R. C. (2017). An exploration of student attitudes and satisfaction in a GAISE-influenced introductory statistics course. *Statistics Education Research Journal*, 16(2), 487–510.
- Persson, I., Kraus, K., Hansson, L., & Wallentin, F. Y. (2019). Confirming the structure of the Survey of Attitudes Toward Statistics (SATS-36) by Swedish students. *Statistics Education Research Journal*, 18(1), 83–93.
- Ramirez, C., Schau, C., & Emmioğlu, E. (2012). The importance of attitudes in statistics education. *Statistics Education Research Journal*, 11(2), 57–71.
- Roberts, D. M., & Bilderback, E. W. (1980). Reliability and validity of a statistics attitude survey. *Educational and Psychological Measurement*, 40(1), 235–238.
- Schau, C. (2003). Survey of attitudes toward statistics (SATS-36). <http://evaluationandstatistics.com>
- Schau, C., & Emmioğlu, E. (2012). Do introductory statistics courses in the United States improve students' attitudes? *Statistics Education Research Journal*, 11(2), 86–94.
- Showalter, D. A. (2013). Place-based mathematics education: A conflated pedagogy. *Journal of Research in Rural Education*, 28(6), 1-13.
- Showalter, D. (2018). Replacing torture tests with Broccoli quizzes. *The Journal for Research and Practice in College Teaching*, 3(2), 160-165.
- Sowey, E. R. (2020). Statistics education that inspires long-term retention of learning—a qualitative model. *Statistics Education Research Journal*, 19(1), 21-28.
- Stern, D., Stern, R., Parsons, D., Musyoka, J., Torgbor, F., & Mbasu, Z. (2020). Envisioning change in the statistics-education climate. *Statistics Education Research Journal*, 19(1), 206-225.
- Swanson, T., VanderStoep, J., & Tintle, N. (2014). Student attitudes toward statistics from a randomization-based curriculum. In K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July, 2014)* (pp. 1-5). International Statistical Institute.
- Thayne, J. L. (2016). *Making statistics matter: Using self-data to improve statistics learning* (Unpublished doctoral dissertation). Utah State University, Logan, UT.
- Tishkovskaya, S., & Lancaster, G. A. (2012). Statistical education in the 21st century: A review of challenges, teaching innovations and strategies for reform. *Journal of Statistics Education*, 20(2). <https://doi.org/10.1080/10691898.2012.11889641>
- Tu, W. (2014). *Designing for statistical reasoning and thinking in a technology-enhanced learning environment* (Unpublished doctoral dissertation). Nova Southeastern University, Davie, FL.
- van Es, C. & Weaver, M. M. (2018) Race, sex, and their influences on introductory statistics education. *Journal of Statistics Education*, 26(1), 48–54. <https://doi.org/10.1080/10691898.2018.1434426>
- Wei Chan, S., Ismail, Z., & Sumintono, B. (2015). The impact of statistical reasoning learning environment: A Rasch analysis. *Advanced Science Letters*, 21(5), 1211–1215. <https://doi.org/10.1166/asl.2015.6077>
- Xu, C. (2019). *Examining instructor and instructional effects on students' statistics attitudes* (Unpublished doctoral dissertation). The University of Houston–Clear Lake, Houston, TX.
- Xu, C., & Schau, C. (2019). Exploring method effects in the six-factor structure of the Survey of Attitudes Toward Statistics (SATS-36). *Statistics Education Research Journal*, 18(2), 39–53.
- Wood, B. L., Mocko, M., Everson, M., Horton, N. J., & Velleman, P. (2018). Updated guidelines, updated curriculum: The GAISE College Report and introductory statistics for the modern student. *Chance*, 31(2), 53–59. <https://doi.org/10.1080/09332480.2018.1467642>
- Zimmerman, W. A., & Austin, S. R. (2018). Using attitudes and anxieties to predict end-of-course outcomes in online and face-to-face introductory statistics courses. *Statistics Education Research Journal*, 17(2), 68–81.