International Journal of Educational Technology and Learning
ISSN: 2523-0581
Vol. 17, No. 2, pp. 64-79, 2024
DOI: 10.55217/101.v17i2.874

Equipping college students with statistical literacy under the massification of higher education

Weichen Teng[1]
Yueh-Hsia Huang[2]♣
Mei-Hui Peng[3]
Tien-Tien Liao[4]

[1,2]*Department of International Trade, Chinese Culture University, Taipei 11114, Taiwan.*
[1]*Email: dwc@ulive.pccu.edu.tw*
[2]*Email: hyx3@ulive.pccu.edu.tw*
[3]*Institute of Information Management, Minghsin University of Science and Technology, Hsinchu 300044, Taiwan.*
[3]*Email: clare4260@gmail.com*
[4]*Department of International Business Administration, Chinese Culture University, Taipei 11114, Taiwan.*
[4]*Email: ttliao@sce.pccu.edu.tw*

Abstract

Big data analysis has become an essential decision-making tool across various sectors, with statistics serving as the critical knowledge base. However, many college students, especially those with weaker mathematical skills, experience anxiety about statistics. This study explores strategies to improve statistical literacy among such students. Using a questionnaire survey, it focuses on an introductory statistics course at a university in Taiwan, applying the Partial Least Squares method to test the structural equation model. The findings suggest that integrating statistical software into instruction, conducting frequent formative assessments, and designing test content conducive to learning can enhance students' perceived learning outcomes and self-efficacy. Despite limitations like small sample size and representativeness, the study highlights that statistical software not only aids learning but also develops practical skills for real-world application. Open-book exams are recommended, but multiple assessment methods should be used to ensure fairness. The study concludes that while statistics anxiety is difficult to alleviate, enhancing self-efficacy can improve students' attitudes toward statistics. It proposes scaffolding strategies to support students with low mathematical skills, ultimately improving their statistical literacy and confidence.

Keywords:
Attitude toward statistics
Statistical literacy and sustainability
Statistics anxiety
Statistics self-efficacy
Supplementary material and statistical Software learning.

Publisher:
Scientific Publishing Institute

Received: *12 September 2024*
Revised: *28 November 2024*
Accepted: *16 December 2024*
Published: *30 December 2024*

(♣ Corresponding Author)

Funding: This research is supported by Ministry of Education, Taiwan (Grant number: PBM1080215).
Institutional Review Board Statement: The Ethical Committee of the MOE Teaching Practice Research Program, Taiwan has granted approval for this study on 1 August 2019.
Transparency: The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.
Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: All authors contributed equally to the conception and design of the study. All authors have read and agreed to the published version of the manuscript.

1. Introduction

Big data analysis has become an important decision-making tool for governments, businesses, and non-profit organizations. The critical knowledge base for data analysis is statistics. For example, statistics are crucial to realizing the Sustainable Development Goals (SDGs) proposed by the United Nations (UN). Because many environmental, ecological, and social issues require data to assist decision-making, SDG indicators are mainly quantitative. Reliable global data depend on countries using the same statistical definitions and methodologies.

Statistics is a well-defined and structured discipline applicable to both the natural and social sciences and has become an indispensable tool in academic research and management practice. However, statistics involves mathematical symbols, formulas, and algorithms, which are not easy to understand; therefore, many college students who dislike mathematics think they will fail statistics courses (Sherwood & Kwak, 2017). In addition, the massification of higher education in many countries, including Taiwan, has led to low entry barriers for universities, resulting in more college students facing statistical learning problems.

Improving the effectiveness of statistics teaching has long been a common concern among colleges. Conners, Mccown, and Roskos-Ewoldsen (1998) identified four significant challenges college teachers face in teaching statistics: encouraging students to read textbooks they deem uninteresting, facing mathematics anxiety, managing polarized student performance, and facilitating deep learning and memory.

Because in an era of big data, the daily operation of the public and private sectors is increasingly dependent on statistical analysis, even entry-level employees need to understand statistics to provide, process, interpret, and analyze data appropriately. Therefore, the pedagogy of statistics for students who have low performance in math to equip them with basic statistical literacy has become a responsibility and challenge in higher education.

How to improve statistics teaching has been widely discussed (e.g., (Cheng, Ferris, & Perolio, 2018; Johanns, Dinkens, & Moore, 2017; Wathen & Rhew, 2019; Zimmerman & Austin, 2018)); and in the current era of a data-driven society, statistical literacy has also been discussed and advocated (Johannssen, Chukhrova, Schmal, & Stabenow, 2021; Sharma, 2017) but few studies provide specific teaching strategies for students with lower mathematical skills and courses aimed at developing statistical literacy in students.

It is more important to help students with lower mathematical abilities learn statistics than to assess their statistical proficiency. Therefore, this study aims to explore teaching strategies to fill the research gap in enriching the statistical literacy of college students with low mathematical skills.

We propose that statistical software lessons, formative assessments, open-book exams, and learning-enhancing test questions will help students who lack mathematical skills improve their perceived learning outcomes and course satisfaction, increasing their self-efficacy, reducing statistics anxiety, and improving their attitude towards statistics.

Self-reported questionnaire results from undergraduates at a university in Taiwan were used to test the research model and hypotheses proposed in this study. The expected results contribute not only to higher education but also to adult education to improve the statistical literacy of citizens.

2. Literature Review

2.1. Statistical Literacy and Sustainability

Statistics plays a crucial role in an era of big data. An understanding of statistical language and techniques has become essential for both daily life and the workplace, leading to calls for increased attention to the statistical literacy of citizens (Makar & Rubin, 2009).

Gal (2002) described statistical literacy as the ability to interpret, evaluate, and communicate statistical information and data-based messages. He also identified two interrelated components of statistical literacy: knowledge and disposition. The knowledge component consists of literacy, statistical, mathematical, contextual, and critical foundations, while the disposition component relates to a critical stance, beliefs, and attitudes (Gal, 2002). Statistical literacy is promoted as a critical competence expected of citizens in modern society, an essential outcome of education, and a necessary component of both literacy and adult numeracy (Garfield, 2011). The UN has echoed the call for statistical literacy. In 2015, the UN launched its 2030 Agenda for sustainable development. To implement the agenda, countries must invest in data production and the required skills and literacies to improve their capacity to effectively use and understand data to generate knowledge and evidence for effective policymaking.

Statistics education has increasingly shifted its focus from procedural understanding—such as statistical techniques, formulas, computations, and algorithms—to developing a conceptual understanding that underpins statistical literacy and higher-order statistical reasoning and thinking (Garfield & Ben-Zvi, 2007). The Australian Bureau of Statistics' Education Services identifies four essential criteria for statistical literacy: data awareness, the ability to understand statistical concepts, the ability to analyze, interpret, and evaluate statistical information, and the ability to communicate statistical information and understanding.

The GAISE (2016) recommends that instructors of introductory statistics courses (a) teach statistical thinking, (b) focus on conceptual understanding, (c) integrate real data with context and purpose, (d) foster active learning, (e) use technology to explore concepts and analyze data, and (f) use assessments to improve and evaluate student learning (GAISE, 2016).

2.2. Statistics Anxiety

Research has demonstrated that college students are prone to anxiety when studying statistics (Onwuegbuzie & Leech, 2003; Zeidner, 1991). Statistics anxiety refers to intense fear, disgust, confusion, tension, and other emotions when students work with statistical concepts, problems, learning, and evaluation

(Zeidner, 1991). Statistics anxiety stems from students' frustration with learning mathematics (Onwuegbuzie, Da Ros, & Ryan, 1997). High anxiety may interfere with students' ability to learn, making it difficult for them to focus on the course content (Hanoch & Vitouch, 2004). They may lose confidence in their statistical abilities, causing them not to complete the course (Malik, 2015). Pan and Tang (2005) indicated that students who lack basic mathematical skills are more likely to experience statistics anxiety; DeVaney (2010) finds that statistics anxiety affects students in fields such as business, psychology, and education.

Some studies have indicated that establishing scaffolding (Wood, Bruner, & Ross, 1976) for students can reduce learning anxiety and help students learn by promoting social interaction between students and teachers, students and peers, and student interaction with teaching tools (e.g., (Murtagh & Webster, 2010; Zackariasson, 2020)). Scaffolding theory was initially developed for early, primary, and secondary education but was later applied to higher education (Murtagh & Webster, 2010; Zackariasson, 2020).

Teaching strategies that can help students alleviate statistics anxiety and promote learning have been explored by a few studies, such as establishing a peer-to-peer learning atmosphere (Calderwood, 2002) adopting digital learning (Suanpang, Petocz, & Kalceff, 2004) involving students in the collection and analysis of statistical data (Bolen, 2006) and making full use of repetition and teachers' immediate feedback (Marson, 2007; Williams, 2010).

2.3. Supplementary Materials and Statistical Software

From the perspective of social constructivism, instructional scaffolding includes a soft framework for teacher–student interaction and student–peer interaction and a hard framework for student–tool interaction (Saye & Brush, 2002; Shin, Brush, & Glazewski, 2020). In the soft framework, teachers or peers provide appropriate support to students in a dynamic form according to the student's needs (Pea, 2004; Saye & Brush, 2002). The hard framework comprises predesigned supplementary teaching materials, such as prerecorded demonstration videos, which can help students overcome anticipated learning obstacles (Lee & Calandra, 2004).

2.4. Formative Assessment for Learning and Summative Evaluation of Learning

In the teaching process, assessments provide teaching feedback. Teachers can use assessments to understand student learning outcomes and adjust and improve teaching strategies to achieve optimal results.

The general form of formative assessment is classwork, homework, and unit tests, usually taken from textbooks, which are smaller in scope and emphasize the fundamentals of each chapter. There is a high frequency of formative assessments in the course.

Formative assessment is crucial for teaching statistics and is often linked to assessment for learning (Earl, 2003). Statistics textbooks typically use sample questions to explain statistical methods and steps to solving a problem, and they provide exercises for classwork or homework, highlighting the value of formative assessment. Classwork and homework are learning promotion mechanisms by which students can practice and internalize statistical techniques to develop their knowledge and skills.

On the other hand, summative assessment is often linked with the assessment of learning. It evaluates students' knowledge, abilities, and skills to understand whether course objectives or standards are met (Stiggins, 2006). Typical forms of summative assessment are midterm, final, and graduation exams. Exam questions are usually taken from textbooks used throughout the semester. Summative evaluation emphasizes the comprehensive application of knowledge, and its primary purpose is to judge students' ability in a particular subject.

2.5. Constructive Alignment with Open-Book Examinations

The mainstream method of summative assessment is a closed-book examination in the classroom under supervision and a time limit (Bengtsson, 2019). Closed-book examinations are advantageous because students spend more time preparing for them (Durning et al., 2016). They can reduce the risk of assessment results being affected by student cheating (Williams & Wong, 2009). However, they are disadvantageous because the unrealistic time limit increases pressure on students and adversely affects their performance. Closed-book examinations also encourage students to cram study to memorize rote answers, thus producing short-term memory effects rather than the long-term effects of knowledge acquisition (Anaya, Evangelopoulos, & Lawani, 2010). Krathwohl (2002) argued that traditional closed-book examinations typically focus on low-level cognitive processing, such as reproduction and description. In addition, closed-book examinations in a classroom are conducted in a deliberately arranged environment that is inconsistent with the operation of real society (Oakleaf, 2008; Simkin, 2005). The effects of closed-book examinations also do not conform to the mainstream teaching ideas of constructive alignment in higher education, which posits that expected learning outcomes, teaching methods, learning activities, and assessment methods should be consistent (Biggs, 1996).

The open-book examination differs from the closed-book examination in that it allows students to use textbooks, notes, or reference materials in the classroom or at home to complete the examination. In open-book examinations, students collect and analyze data from various sources to increase their engagement and depth of thinking (Sato, He, Warschauer, & Kadandale, 2015). Open-book examinations are advantageous because they reduce test anxiety and unnecessary rote memorization, encourage action learning, cultivate students' ability to learn independently, allow students to develop advanced thinking skills and deep learning, and increase student achievement; students can master course materials by improving their study skills, increasing their learning effectiveness, simulating real situations, and promoting constructive learning (Block, 2012; Boniface, 1985; Johanns et al., 2017). Krathwohl (2002) argued that open-book examinations require cognitive processing at elevated levels in Bloom's taxonomy (e.g., analysis, evaluation, and creation).

However, some studies have indicated that open-book examinations do not significantly improve student performance, and the long-term retention of learning outcomes from open-book tests is not significantly different from those achieved with closed-book tests (Agarwal, Karpicke, Kang, Roediger III, & McDermott, 2008; Theophilides & Dionysiou, 1996). (Ioannidou, 1997) found no significant difference between the results of open-book and closed-book tests in assessing students' higher-level analytical thinking abilities. Theophilides and Koutselini (2000) identified the following shortcomings of open-book examinations: (a) students spend a substantial amount of time looking for information instead of formulating and writing answers, and (b) students often spend less time preparing for open-book tests. Despite the shortcomings of an open-book examination, it can promote deep learning and is still recognized by scholars (Teodorczuk, Fraser, & Rogers, 2018).

Open-book examinations can be divided into two forms according to testing location: classroom and home (Heijne-Penninga, Kuks, Hofman, & Cohen-Schotanus, 2010). Take-home examinations resemble real-world problem-solving methods and promote collaborative learning (American Association for the Advancement of Science, 2011; Handelsman et al., 2004). Compared with traditional examinations, take-home examinations can cause less test anxiety, cultivate people skills, enable students to spend more time and effort answering questions and enable students to apply and integrate knowledge to deepen their understanding (Johnson, Green, Galbraith, & Anelli, 2015).

2.6. Statistics Self-Efficacy

According to research on learning anxiety, student self-efficacy crucially affects anxiety (Onwuegbuzie, 2000). Self-efficacy refers to an individual's belief that they have sufficient ability to complete tasks (Bandura, 1977). Self-efficacy is unrelated to personal skills but is related to self-judgment of ability. Self-efficacy is a person's belief in their ability, which determines their behavior, thinking, and emotional response in a specific situation. People with higher self-efficacy can achieve more, treat complex tasks as challenges and exercises, experience less stress, and avoid depression. Finney and Schraw (2003) defined statistics self-efficacy as an individual's confidence in their ability to learn the skills required for statistical courses, and they proved that statistics self-efficacy is positively correlated with learning performance.

2.7. Attitude toward Statistics

According to the theory of learning and cognition, learning attitude is a crucial factor affecting student learning effectiveness (Schau & Emmioğlu, 2012). In research on teaching statistics, attitude toward statistics is typically used as a dependent variable to evaluate the effectiveness of teaching strategies (e.g., (Abbiati et al., 2021; Cladera, Rejón-Guardia, Vich-I-Martorell, & Juaneda, 2021)). Studies have confirmed that attitude toward statistics is related to and can predict learning achievement (Emmioğlu & Capa-Aydin, 2012; Zimmerman & Austin, 2018). The more positive the attitude toward statistics, the more learning achievement can be improved (Chiesi & Primi, 2009). Therefore, attitude toward statistics indicates learning achievement; improving students' attitudes toward statistics is crucial in teaching statistics (Schau, 2003).

Attitude toward statistics is multidimensional and represents the learning trends of students with positive or adverse reactions to statistics (Emmioğlu & Capa-Aydin, 2012). To measure attitudes toward statistics, scholars have developed relevant scales, of which the Survey of Attitudes Toward Statistics (SATS); Schau, Stevens, Dauphinee, and Del Vecchio (1995)) is one of the most used. The SATS-28 scale consists of four dimensions: emotion, cognitive ability, value, and difficulty (Hilton, Schau, & Olsen, 2004; Schau et al., 1995).

3. Research Model and Hypotheses

3.1. Effects Of Statistics Self-Efficacy and Statistics Anxiety

Studies have demonstrated that students usually have a negative attitude toward statistics (Onwuegbuzie, 2004; Sizemore & Lewandowski, 2009) regarded as a significant obstacle to learning statistics (Waters, Martelli, Zakrajsek, & Popovich, 1988). In addition, previous studies confirmed that statistics attitude indicates academic performance (Chiesi & Primi, 2009) and belongs to the dispositional component of statistical literacy (Gal, 2002). Therefore, this study uses statistics attitude as an indicator of the effectiveness of the proposed statistics pedagogy. Improving students' statistics attitudes is believed to help them develop statistical literacy and enhance their ability to learn statistics.

Since many college students who enroll in an introductory statistics course react with anxiety, and statistics anxiety negatively affects learning achievement (Macher, Paechter, Papousek, & Ruggeri, 2012; Onwuegbuzie, 2004; Zare, Rastegar, & Hosseini, 2011) and the indicator of statistics attitude can predict the learning achievement, the following hypothesis is proposed:

H_1: Statistics anxiety will reduce the attitude toward statistics.

In addition to statistics anxiety, statistics self-efficacy has also been used to validate the effectiveness of teaching strategies in statistics courses. McGrath, Ferns, Greiner, Wanamaker, and Brown (2015) observed a significant positive correlation between students' statistics self-efficacy and learning performance.

(Onwuegbuzie, 2000) suggested that statistics anxiety is higher in students with lower self-efficacy. According to the study of Finney and Schraw (2003) statistics self-efficacy significantly affects attitude toward statistics. Based on the findings mentioned above, the following hypotheses are proposed:

H_2: Statistics self-efficacy will reduce statistics anxiety.

H_3: Statistics self-efficacy will improve the attitude toward statistics.

3.2. Effects of Perceived Learning Outcomes and Course Satisfaction

The effectiveness of teaching depends on the learning outcome achieved by students. Student achievement is a cognitive variable that includes grades and test scores, often viewed as a primary learning outcome (Doménech-Betoret, Abellán-Roselló, & Gómez-Artiga, 2017). However, students' grades are easily affected by variables such as evaluation criteria and methods, which makes it difficult to generalize the research results. Mitra (2023) found that students' perception of their learning is strongly associated with direct assessment grades. Therefore, this study uses perceived learning outcomes to verify the effectiveness of the proposed teaching strategies.

Furthermore, when discussing the pedagogy of Introduction to Statistics, course satisfaction is used as another indicator for evaluating teaching effectiveness (Haughton & Kelly, 2015; Soesmanto & Bonner, 2019). Therefore, this study uses students' perception of learning outcomes and course satisfaction with statistics courses as indicators of teaching effectiveness.

We assume that the perceived learning outcomes and course satisfaction of students, which are the indicators of teaching effectiveness, will affect the attitude toward statistics, self-efficacy, and anxiety and propose the following hypotheses:

H_{4-a}: Perceived learning outcomes of the statistics course will reduce statistics anxiety.

H_{4-b}: Perceived learning outcomes of the statistics course will improve the attitude toward statistics.

H_{4-c}: Perceived learning outcomes of the statistics course will improve statistics self-efficacy.

H_{5-a}: Satisfaction with the statistics course will reduce statistics anxiety.

H_{5-b}: Satisfaction with the statistics course will improve the attitude toward statistics.

H_{5-c}: Satisfaction with the statistics course will improve statistics self-efficacy.

3.3. Effects of Supplementary Materials

This study selected a statistics textbook containing practical cases in the field of business management, and students were guided to understand the statistical foundation by using practical examples in the textbook to solve statistical problems.

However, since the first-year students are not familiar with business and management, to enhance students' interest in learning statistics and build learning scaffolding, in addition to textbook cases, classroom discussions on supplementary materials were also held. Examples include news articles, video clips, and statistical reports from market research organizations or governments. Supplementary topics include opinion polls, price indices, rainfall probabilities, and COVID-19 virus testing. These examples encourage students to understand statistical terminology, statistical fundamentals commonly used in everyday life, and the scope and value of statistics.

Supplementary materials were assumed to improve learning outcomes and satisfaction with the statistics course; therefore, the following hypotheses are proposed:

H_{6-a}: Supplementary materials will improve the student's perceived learning outcomes of the statistics course.

H_{6-b}: Supplementary materials will increase the student's satisfaction with the statistics course.

3.4. Effects of Statistical Software Teaching

Statistics is a strictly structured knowledge system developed from mathematics, involving numerous mathematical terms, symbols, and formulas. Using statistical software to teach statistics can eliminate tedious statistical calculations (Harrington & Schibik, 2004) thus providing teachers with more time to explain key concepts and examples.

Microsoft Excel has the essential functions required for statistical data processing—including data creation and editing, function calculations, and graph generation. Moreover, Excel is one of the most popular spreadsheet software programs in Taiwan, often used on personal computers; therefore, this study uses Excel to assist teaching.

Because Excel can calculate functions, students are no longer required to memorize formulas; however, they must understand when to apply each function, how to use software tools to perform them, how to interpret the results, and how to use statistical graphs and tables to present results.

Statistical software teaching was presumed to improve student's learning outcomes and increase course satisfaction; therefore, the following hypotheses are proposed:

H_{7-a}: Teaching statistical software will improve the student's perceived learning outcomes of the statistics course.

H_{7-b}: Teaching statistical software will increase the student's satisfaction with the statistics course.

3.5. Effects of Formative Assessment

Vaessen, Prins, and Jeuring (2017) stated that formative assessments are considered motivational factors for learning. Regarding learning burden, homework assignments with a small assessment scope were presumed to help reduce anxiety, and students could repeat and explain the homework exercises in subsequent classes to strengthen their understanding. Because this study aimed to improve statistics teaching effectiveness for students with low mathematical skills, weekly formative assessments were used as homework assignments after each class.

Implementing formative assessment was presumed to increase students' learning outcomes and satisfaction with the statistics course, enhancing their statistics attitude. Therefore, the following hypotheses were proposed:

$H_{8\text{-}a}$: Formative assessment will improve the student's perceived learning outcomes of the statistics course.

$H_{8\text{-}b}$: Formative assessment will increase the student's satisfaction with the statistics course.

3.6. Effects of Perceived Test Content on Learning

Regarding learning assessment, Salcedo (2014) finds that teacher-made tests remain the preferred means of assessing learning outcomes in the higher education system. He collected and analyzed test questions prepared by teachers in college statistics courses. He found that most were designed to see whether students understood basic statistical concepts and could perform the basic statistical calculations taught in the classroom.

This study proposes that the test content is essential to the assessment for learning to help students evaluate their learning outcomes. The following hypotheses are stated:

$H_{9\text{-}a}$: Perceived test content that facilitates learning will improve the student's perceived learning outcomes of the statistics course.

$H_{9\text{-}b}$: Perceived test content facilitating learning will increase the student's satisfaction with the statistics course.

3.7. Effects of Open-Book Exams

Since the Internet is ubiquitous and people can use keyword queries on web search engines to search for answers to questions, people's learning strategies can change from traditional passive acceptance of knowledge to active construction of knowledge. Receptive learning strategy relies on narrative and memory, whereas constructive learning strategy relies on understanding and organization. Therefore, the role of textbooks in learning can be transformed from a "bible" that provides students with the correct answers to a reference material with which students can locate answers, and teachers should encourage students to use textbooks. Accordingly, this study's formative and summative assessments were based on the open-book and take-home exam models. Students could read textbooks repeatedly to locate and consider answers without a time limit, enabling them to familiarize themselves with the meanings of statistical terms and symbols. Students could also discuss freely with classmates and seek answers to establish a collaborative learning scaffold.

This study hypothesizes that open-book exams can improve students' learning outcomes and increase their satisfaction with statistics courses. Therefore, the following hypotheses were proposed:

$H_{10\text{-}a}$: Open-book exams can improve the student's perceived learning outcomes of the statistics course.

$H_{10\text{-}b}$: Open-book exams can increase the student's satisfaction with the statistics course.

Open-book and take-home exams also have disadvantages. Plagiarism is likely because take-home exams are unsupervised (Bengtsson, 2019). In other words, teachers need help distinguishing between students' authentic answers, copying answers from other sources, and answers completed by others on take-home exams. If students perceive that some classmates have engaged in such dishonest behavior, they will feel that the examination is unfair. Therefore, without a proper method for detecting and responding to cheating on take-home exams, students will be less supportive of open-book and take-home exams, affecting their satisfaction with the course. Conversely, the more students support open-book and take-home exams, the less likely they are to doubt the fairness of the examinations. Therefore, the following hypotheses were proposed:

H_{11}: The more positive perceptions of open-book exams, the less doubt students have about the fairness of the exams.

H_{12}: The higher the level of doubt about the fairness of open-book exams, the greater the negative impact on course satisfaction.

Based on the hypotheses, the research model for pedagogical scaffolding for statistical literacy is presented in Figure 1.

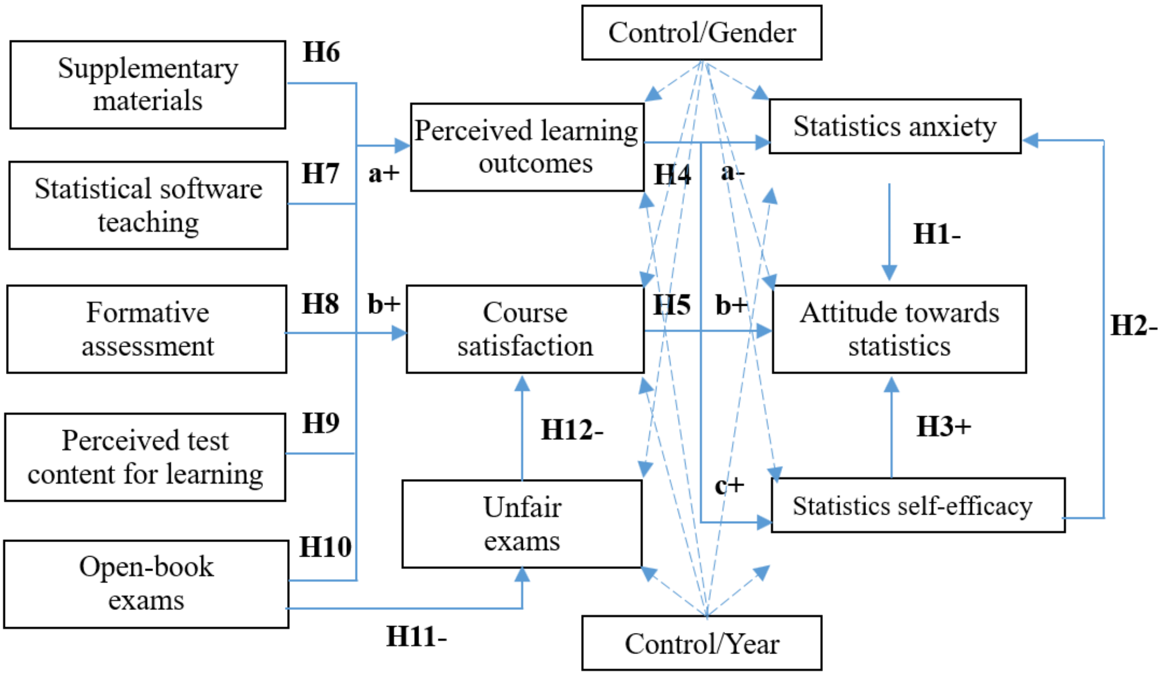

Figure 1. Research model.

4. Methodology

This study adopts the method of questionnaire survey to conduct empirical research. The author uses an introductory statistics course taught at a university in Taiwan as the sample source. The course was awarded three credits, 18 weeks per semester, for two semesters a year. Most students in this course are first-year students, but a few students retake the course because of previous failures. This study collected students' questionnaire responses in two years, from the fall of 2020 to the spring of 2022, for data analysis.

The first-semester course introduced data description (including frequency distribution, numerical measurement, and data display) and probability distributions (including an overview of probability, discrete probability distributions, and continuous probability distributions). The second-semester course included sampling methods and the central limit theorem, estimation and confidence intervals, hypothesis testing, analysis of variance, correlation analysis, and linear regression.

The questionnaire contains items scored using a 5-point Likert scale to measure respondents' perceptions, beliefs, and attitudes. Questionnaires were distributed in class at the end of the two semesters. Respondents voluntarily and anonymously responded to the questionnaire to ensure that respondents would not worry that their answers would affect their course grades.

The construct items in the questionnaire were based on the statistics anxiety scale and construct definition developed in previous studies. The wording of some items was adjusted based on pilot test results and the discussion of two statistics teachers to enhance the validity of the questionnaire. The questionnaire items are presented in Table 1.

Table 1. Construct items.

Construct	Item	
Statistics anxiety (SA)	SA1	I fear statistics.
	SA2	Statistics make me anxious.
	SA3	I feel under pressure when I study statistics.
Attitude toward statistics (ATS)	ATS1	I am interested in learning statistics.
	ATS2	I am interested in analyzing statistical information.
	ATS3	I am interested in being able to communicate statistics to others.
Self-efficacy (SE)	SE1	I am confident that I understand statistics.
	SE2	I am confident that I can learn statistics.
	SE3	I am confident that I can apply statistics.
Statistical software teaching (SST)	SST1	The teacher used Excel to explain concepts in class, which helped me to learn statistics.
	SST2	The teacher asked me to complete my homework in Excel, which helped me study statistics.
Supplementary materials (AM)	AM1	The teacher played video clips in class that helped me learn statistics.

Construct	Item	
	AM2	The teacher gave examples of news and current statistics-related cases, which helped me study statistics.
Formative assessment (FA)	FA1	After each class, the textbook exercises selected by the teacher for us are helpful for me to learn statistics.
	FA2	I agree with the practice of doing homework after each class.
Test content for learning (TCL)	TCL1	I think the exam questions can effectively help me learn statistics.
	TCL2	I think the exam questions will help me assess my learning outcomes in studying statistics.
	TCL3	I think the exam questions can help me assess my level of statistical knowledge.
Open-book exams (OBE)	OBE1	The teacher conducted open-book exams, which were helpful for me to learn statistics.
	OBE2	The teacher allowed students to complete the exam at home, which was helpful for me to study statistics.
Unfair exams (UE)	UE1	I think the way the teacher designed the examination caused unfairness.
	UE2	I think the examination method used by the teacher was unfair to the students who studied hard.
Course satisfaction (CS)	CS1	Overall, the experience of taking the statistics course was better than I expected.
	CS2	Overall, I liked the statistics course.
	CS3	Overall, I am satisfied with the statistics course.
Perceived learning outcomes (PLO)	PLO1	I think learning to use Excel gives me a sense of accomplishment.
	PLO2	The statistics course gave me a sense of achievement.
	PLO3	I learned a lot from the statistics course.

High school students in Taiwan who apply for university and college admission must complete the General Scholastic Ability Test (GSAT) developed by the College Entrance Examination Center. The purpose of the test is to help schools understand the readiness of applicants for undergraduate education. The percentile scores publicly reported by the college entrance examination center were 88%, 75%, 50%, 25%, and 12%. GSAT scores are often used as a typical threshold for the initial screening of applicants by universities and college admission offices.

This study aims to develop effective teaching strategies to help students with weaker mathematical abilities learn statistics. Table 2 shows that most respondents in the class had average or below-average mathematical ability, which meets the purpose of this study. Only 11 (5.6%) respondents scored above 75% on the GSAT Math and were excluded from the sample. In addition, questionnaires with consistent answers or too many missing values were also eliminated; therefore, 181 valid questionnaires were retained and analyzed.

Table 2. Number of the sample.

Frequency/Period	Jan. 2021	June 2021	Jan. 2022	June 2022	Total
Collect	67	47	51	32	197
Exclude Math >=75th	4	1	3	3	11
Other (Missing)	3	1	1	0	5
Valid sample	60	45	47	29	181

Since the number of valid samples in this study does not reach 100 per semester, which is a small sample size, this study uses the Partial Least Squares (PLS) method recommended by Hair Jr, Sarstedt, Hopkins, and Kuppelwieser (2014) to test the validity of the structural equation modeling. SmartPLS 4.0 software was used for analysis.

4.1. Data Analysis

As shown in Table 3, 38.7% of the valid samples are men and 61.3% are women. The math level of most of the valid samples was around or below the GSAT average; some students (15.5%) could not remember or did not report math levels.

The distribution of gender and math achievement in the valid sample did not differ significantly between study years, as the chi-square test did not reach significance.

Table 3. Valid sample size by gender and math ability.

Variable	Jan. 2021	June 2021	Jan. 2022	June 2022	Total	X^2-test
Gender						
Male	20(33.3%)	13(28.9%)	22(46.8%)	15(51.7%)	70(38.7%)	X^2=5.932
Female	40(66.7%)	33(71.1%)	25(53.2%)	14(48.3%)	111(61.3%)	df=3
Total	60 (100%)	45 (100%)	47 (100%)	29 (100%)	181 (100%)	P=0.115
Math ability (percentile on GSAT in math)						
$75^{th}>=50^{th}$	10(16.7%)	12(26.7%)	18(38.3%)	7(24.1%)	47(26.0%)	
$50^{th}>=25^{th}$	25(41.7%)	17(37.8%)	15(31.9%)	12(41.4%)	69(38.1%)	
$25^{th}>=12^{th}$	12(20.0%)	5(11.1%)	7(14.9%)	2(06.9%)	26(14.4%)	X^2=15.576
$12^{th}>$	2(03.3%)	4(08.9%)	2(04.3%)	3(10.3%)	11(06.1%)	df=15
N.A.	11(18.3%)	7(15.6%)	5(10.6%)	5(17.2%)	28(15.5%)	P=0.411
Total	60 (100%)	45 (100%)	47 (100%)	29 (100%)	181 (100%)	

Table 4 shows the statistics of each construct in the research model. The construct of the open-book exam has the highest mean (4.259), and the construct of unfair assessment has the lowest mean (2.376). All values for Cronbach's α and composite reliability for each construct were above 0.7, meeting the recommended measurement reliability thresholds. Convergent validity is satisfied because all item factor loadings are above 0.7 and reach statistically significant (Hair Jr et al., 2014) and all values of average variance extracted (AVE) of all factors are higher than 0.5 (Fornell & Larcker, 1981). All Variance Inflation Factors (VIFs) are below 10, which indicates that multicollinearity of items is not a severe problem in the research model (Kline, 2011).

Table 4. Constructs and statistics.

Construct	Item	F. loading	VIF	Mean	S.D.	Cronbach α	C.R.	AVE
Statistics anxiety (SA)	SA1	0.774***	2.310	2.974	0.742	0.870	0.872	0.794
	SA2	0.972***	3.454					
	SA3	0.764***	2.237					
Attitude towards statistics (ATS)	ATS1	0.890***	3.181	3.437	0.728	0.898	0.899	0.831
	ATS2	0.855***	2.634					
	ATS3	0.850***	2.669					
Self-efficacy (SE)	SE1	0.923***	2.423	3.329	0.710	0.899	0.901	0.833
	SE2	0.867***	3.513					
	SE3	0.811***	2.955					
Statistical software teaching (SST)	SST1	0.960***	6.658	4.008	0.744	0.959	0.960	0.961
	SST2	0.960***	6.658					
Supplementary materials (SM)	SM1	0.847***	2.476	3.975	0.707	0.871	0.871	0.886
	SM2	0.912***	2.476					
Formative assessment (FA)	FA1	0.833***	1.433	3.794	0.680	0.709	0.714	0.774
	FA2	0.660***	1.433					
Test content for learning (TCL)	ECL1	0.859***	2.975	3.799	0.642	0.915	0.918	0.854
	ECL2	0.899***	3.855					
	ECL3	0.896***	3.084					
Open-book exams (OBE)	OBE1	0.892***	2.881	4.259	0.746	0.894	0.896	0.904
	OBE2	0.906***	2.881					
Unfair exams (UE)	UE1	0.936***	3.928	2.376	0.885	0.927	0.928	0.932
	UE2	0.922***	3.928					
Course satisfaction (CS)	CS1	0.870***	2.596	3.983	0.732	0.904	0.940	0.839
	CS2	0.855***	2.943					
	CS3	0.889***	3.449					
Perceived learning outcomes (PLO)	PLO1	0.769***	2.270	3.816	0.737	0.885	0.929	0.813
	PLO2	0.817***	2.732					
	PLO3	0.933***	2.692					

Note: F. indicates Factor; S.D. indicates standard deviation; C.R. indicates composite reliability. ***indicates p<0.001.

Discriminant validity was evaluated using the criterion that the square root of the AVE for each latent variable should be greater than the correlation coefficients between that latent variable and other latent variables in the measurement model (Fornell & Larcker, 1981). Table 5 shows that the model meets the discriminant validity criterion.

Table 5. Correlation between constructs.

Construct	SA	ATS	SE	SST	SM	FA	TCL	OBE	UE	CS	PLO
SA	(0.891)										
ATS	-0.152	(0.912)									
SE	-0.266	0.608	(0.912)								
SST	-0.057	0.464	0.412	(0.980)							
SM	0.024	0.416	0.345	0.635	(0.941)						
FA	-0.062	0.516	0.390	0.561	0.482	(0.880)					
TCL	-0.045	0.494	0.306	0.573	0.457	0.627	(0.924)				
OBE	-0.070	0.336	0.213	0.570	0.490	0.532	0.476	(0.951)			
UE	0.067	-0.063	0.027	-0.282	-0.204	-0.260	-0.288	-0.357	(0.965)		
CS	-0.211	0.524	0.327	0.581	0.499	-0.269	0.627	0.507	-0.357	(0.916)	
PLO	-0.188	0.592	0.436	0.579	0.486	0.592	0.610	0.506	-0.305	0.789	(0.902)

Note: The values in the diagonal row are square roots of the average variance extracted. SA: Statistics anxiety; ATS: Attitude towards statistics; SE: Self-efficacy; SST: Statistical software teaching; SM: Supplementary materials; FA: Formative assessment; TCL: Test content for learning; OBE: Open-book exams; UE: Unfair exams; CS: Course satisfaction; PLO: Perceived learning outcomes

Table 6 shows that the mean values of students' statistics attitude, statistical software teaching, and supplementary materials in the second semester are significantly higher than those in the first semester; the other means are not significantly different between the two semesters.

Table 6. Significance tests for means of a two-semester survey.

Construct	1st semester	2nd semester	Difference	t	p
Attitude towards statistics	3.322	3.603	-0.281	-2.585	0.011
Statistics anxiety	2.925	3.054	-0.129	-1.150	0.252
Self-efficacy	3.293	3.378	-0.085	-0.794	0.428
Perceived learning outcomes	3.788	3.856	-0.068	-0.607	0.545
Course satisfaction	3.922	4.072	-0.150	-1.358	0.179
Statistical software teaching	3.883	4.192	-0.309	-2.767	0.006
Supplementary materials	3.841	4.169	-0.328	-3.129	0.002
Formative assessment	3.734	3.872	-0.138	-1.339	0.182
Test content for learning	3.769	3.840	-0.070	-0.721	0.472
Open-book exam	4.196	4.351	-0.155	-1.429	0.155
Unfair exams	2.378	2.372	0.006	0.051	0.959

The test results of the research model are shown in Table 7. Students' statistics self-efficacy, perceived learning outcomes, and course satisfaction will significantly and positively affect their attitudes toward statistics. Only students' statistics self-efficacy will significantly and negatively affect their statistics anxiety. Students' perceptions of the course's learning outcomes will significantly and positively affect their statistics self-efficacy. The teaching strategies of the teaching of statistical software, the placement of formative assessments, and the design of test content that facilitates learning can all significantly and positively affect students' perceptions of course learning outcomes. The teaching strategies of the placement of formative assessments and the design of test contents that facilitate learning will significantly and positively affect students' satisfaction with the course. Students' perception of unfairness on open-book exams will significantly and negatively affect their satisfaction with the course. Students' favorable opinion of open-book exams will significantly negatively impact their perception of the unfairness of open-book exams. In summary, H2, H3, H4b, H4c, H5b, H7a, H8a, H8b, H9a, H9b, H11, and H12 are supported, and the rest of the hypotheses are not supported in the structural model.

Table 7. Test results of the research model.

Hypothesis	β	t	p	R^2
Statistics anxiety→Attitude towards statistics	0.062	0.916	0.360	0.543
Statistics self-efficacy→Attitude toward statistics	0.421	5.692	0.000	
Perceived learning outcomes→Attitude toward statistics	0.255	2.495	0.013	
Course satisfaction→Attitude toward statistics	0.161	1.816	0.069	
Statistics self-efficacy→Statistics anxiety	-0.230	2.484	0.013	0.094
Perceived learning outcomes→Statistics anxiety	0.063	0.470	0.638	
Course satisfaction→Statistics anxiety	-0.175	1.550	0.121	
Perceived learning outcomes→Statistics self-efficacy	0.461	3.933	0.000	0.218
Course satisfaction→Statistics self-efficacy	-0.065	0.577	0.564	
Supplementary materials→Perceived learning outcomes	0.087	1.242	0.214	0.503
Statistical software teaching→Perceived learning outcomes	0.173	2.084	0.037	

Hypothesis	β	t	p	R^2
Formative assessment→Perceived learning outcomes	0.212	2.720	0.007	
Test content for learning→Perceived learning outcomes	0.276	3.514	0.000	
Open-book exam→Perceived learning outcomes	0.114	1.449	0.147	
Supplementary materials→Course satisfaction	0.108	1.544	0.123	0.534
Statistical software teaching→Course satisfaction	0.156	1.845	0.065	
Formative assessment→Course satisfaction	0.196	2.350	0.019	
Test content for learning→Course satisfaction	0.291	0.3.397	0.001	
Open-book exam→Course satisfaction	0.076	0.997	0.319	
Unfair exam→Course satisfaction	-0.134	2.258	0.024	
Open-book exam→Unfair exams	-0.329	5.080	0.000	0.147

Figure 2 shows the path coefficients through the mediator of perceived learning outcomes. These teaching strategies, such as using statistical software, conducting formative assessments, and designing test content that facilitates learning, can significantly improve students' perceptions of learning outcomes. High perceived learning outcomes can improve students' statistics self-efficacy and attitudes.

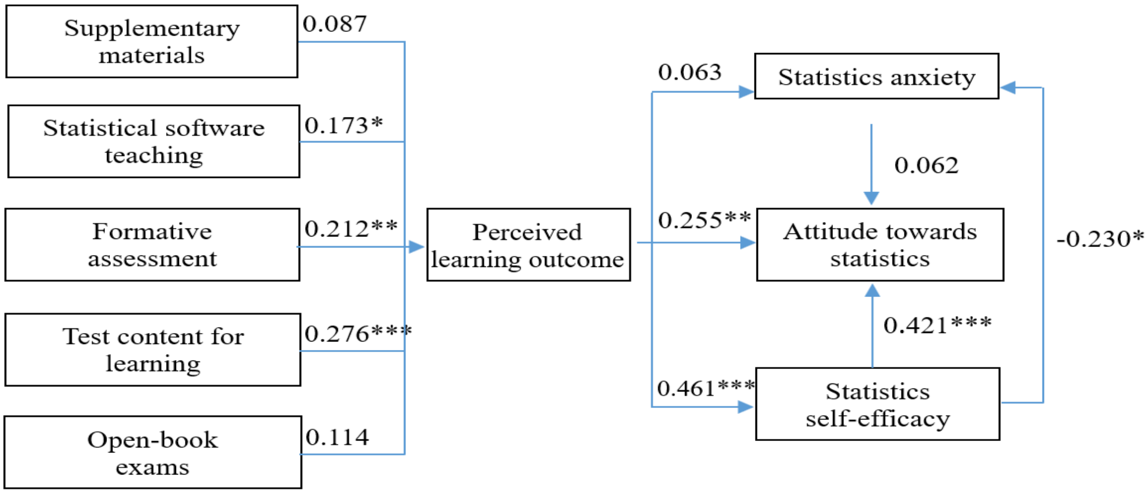

Figure 2. The path coefficients through the mediator of perceived learning outcomes.

Note: * indicates p ≤ 0.10; ** indicates p ≤ 0.01; *** indicates p ≤ 0.001.

Figure 3 shows the path coefficients through the mediator of course satisfaction. These teaching strategies of formative assessments and test content for learning can significantly improve students' course satisfaction. High course satisfaction can improve students' attitudes toward statistics.

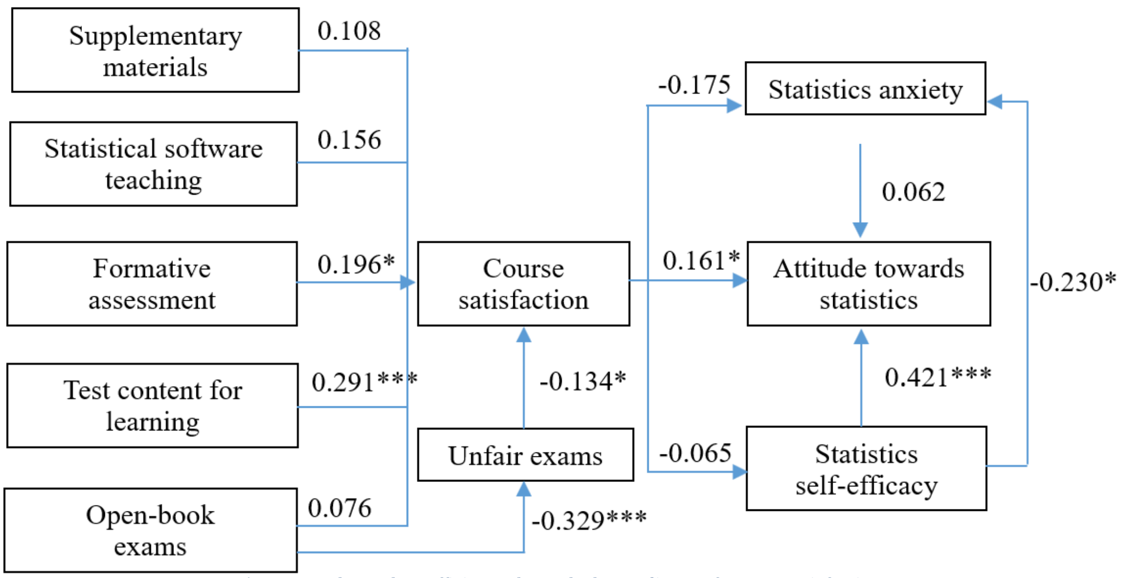

Figure 3. The path coefficients through the mediator of course satisfaction.

Note: * indicates p ≤ 0.10; *** indicates p ≤ 0.001.

5. Discussions

Students' attitudes toward statistics have been shown to predict their academic performance (Chiesi & Primi, 2009; Emmioğlu & Capa-Aydin, 2012; Zimmerman & Austin, 2018) and belong to the dispositional component of statistical literacy (Gal, 2002) so how to improve students' attitudes towards statistics are crucial in the teaching of statistics (Schau, 2003).

According to the independent sample mean t-test results in Table 6, the mean value of the students' attitudes toward statistics in the second semester was significantly higher than that in the first semester, indicating that the teaching strategies adopted in this study significantly improved statistics attitudes. According to the path analysis results in Table 7, the antecedents that affect the improvement of statistics attitudes can be the perceived learning outcomes and course satisfaction of statistics courses.

Figure 2 shows that for students with weaker mathematical abilities, teachers use statistical software for teaching (Excel in this study), high-frequency small-scale assignments for formative assessment, and design test questions that facilitate learning, can effectively improve students' perceived learning outcomes. Figure 3 shows that formative assessment and test questions facilitating learning can improve course satisfaction.

The above results suggest that students are most concerned about grades. Therefore, in teaching strategies, the evaluation methods and content that affect grades are the key factors that affect the perception of learning outcomes and course satisfaction.

Although the impact of statistical software teaching on course satisfaction is insignificant, Table 7 shows that its p-value is 0.065, which is close to the threshold (0.05). Table 4 shows that the mean value of statistical software teaching perception is 4.008, indicating that most students in this course agreed with the strategy for statistical software teaching to help them learn statistics. Due to the small sample size of this study, the impact of statistical software teaching on course satisfaction is still worthy of follow-up research to verify its effect.

In addition, although students almost agreed with the strategy of supplementary materials to help them learn statistics (mean=3.975), Table 7 shows that it had no significant effect on improving perceived learning outcomes and course satisfaction, respectively. The reasons may be that statistics contain mathematical notation, mathematical derivations, and formulas and require calculations, so conceptual interpretation, therefore, has less impact on statistical learning outcomes for students with lower mathematical abilities than other teaching strategies or because there is still room for improvement in the design and discussion of supplementary materials. By contrast, encouraging students to carry out hands-on learning activities such as EXCEL software operation and daily homework exercises can help students understand statistics more directly and effectively.

The results of the path analysis in Table 7 also show that the design of test questions is essential for statistics teaching, and the questions that facilitate learning can effectively improve students' perceived learning outcomes and course satisfaction. Therefore, teachers not only need to spend time and effort designing teaching materials but also pay attention to the design of test questions. This study suggests that for students with weaker mathematical abilities, the test questions should be designed to help students learn rather than increase the difficulty and discrimination of test questions to evaluate students' levels of statistical ability.

In addition, Table 4 shows that the mean value of the contribution of open-book examinations to learning statistics was 4.259, which was significantly higher than that of other teaching strategies. However, the path analysis results in Table 7 show that students' favorability for open-book exams has no significant impact on their perceived learning outcomes and course satisfaction. The reason may be that open-book exams also bring fairness problems, and the sense of unfairness in open-book exams will reduce course satisfaction.

However, Figure 3 shows that open-book exam favorability will reduce test unfairness perceptions, suggesting that students who prefer open-book tests are less likely to perceive open-book tests as unfair. Thus, how open-book exam affects course satisfaction will be influenced by the sense of fairness. Open-book exams will negatively impact course satisfaction for students who perceive open-book exams as unfair. Therefore, when conducting open-book exams, teachers should minimize students' perception of unfairness in the exam.

Table 4 also shows that the mean value of the perception of the exam unfairness is 2.376, the lowest among all the questionnaire constructs. This suggests that most students think the open-book exam will not bring unfairness. However, the take-home exam is not easy to supervise, so there is no guarantee that the students will complete the answers themselves. This study suggests that teachers use formative assessments and other evaluation methods to cooperate with open-book examinations to reduce the unfairness caused by open-book examinations. The multi-assessment method is more suitable for classes whose teaching purpose is to improve students' statistical literacy rather than to evaluate and grade their professional statistical ability.

This study also found that statistics anxiety was significantly negatively correlated with statistics attitudes; however, when the variables of statistics self-efficacy and course satisfaction were considered simultaneously, statistics anxiety had no significant effect. The results showed that statistics self-efficacy and course satisfaction better explained the variation in statistics attitudes. This finding reminds teachers that even when teaching classes where students are weaker in math, they can still improve students' attitudes toward statistics by increasing students' self-efficacy and course satisfaction.

In addition, none of the teaching strategies used in this study can directly reduce students' statistics anxiety. However, they can only indirectly reduce statistical anxiety by enhancing statistics self-efficacy. The reason may be that statistics anxiety stems from the frustration of learning mathematics, a long-term accumulated mental state, so it is not easy to reduce statistics anxiety.

Finally, Table 6 shows that the average favorability of statistical software teaching and supplementary materials in the second semester is significantly higher than in the first semester. The reason may be that the teaching scope in the second semester includes the central limit theorem, estimation and confidence interval, hypothesis testing, variance analysis, correlation analysis, and linear regression, which are difficult to understand and require calculations, so students are more grateful for the pedagogical scaffold of statistical software and supplementary materials.

6. Conclusions and Limitations

Statistics is usually a foundational subject in natural science and social science in higher education, so how to improve the effectiveness of statistics teaching has always been a concern to universities. Especially in the era of the popularization of higher education, an increased number of students entering colleges are weak in mathematics. Therefore, they are prone to anxiety about statistics that require mathematical symbols and formulas.

Past studies have shown that statistics anxiety, statistics self-efficacy, and statistics course satisfaction are crucial factors affecting students' learning achievement in statistics. The empirical results of this study found that for college students with weak mathematics ability, their perceptions of learning outcomes and course satisfaction of a statistics course can improve their attitudes towards statistics that can predict statistics academic achievement; perceived learning outcomes can also improve statistics self-efficacy. Although there was a significant negative relationship between statistics anxiety and statistics attitudes, this significant effect was attenuated to insignificance by perceived learning outcomes and course satisfaction. Moreover, statistics anxiety is not easy to reduce; it can be improved by enhancing statistics self-efficacy.

This study refers to the literature on statistical pedagogy and believes that for students with weaker mathematical abilities, teaching scaffolding can improve the teaching effect of statistics. After analyzing empirical results, integrating statistical software into teaching, conducting small-scope multi-frequency formative assessment, and designing test contents that facilitate learning will help improve students' perceived learning outcomes, thereby improving students' statistics self-efficacy and attitudes.

Formative assessment and test content that facilitates learning can also increase student course satisfaction and thus improve statistical attitudes. The sense of unfairness will intervene in the positive impact of open-book exams on course satisfaction; the open-book exams will improve course satisfaction only through students' low test unfairness perceptions.

The main contribution of this study is to suggest that the statistics class, whose primary purpose is to cultivate statistical literacy, can integrate statistics courses into the teaching of statistics software, adopt formative assessment, and design examination content to help students learn. This study believes that teaching students who are weak in mathematics to learn statistical software will not only help them learn statistics but the software operation skills they learn can also be used in practical work.

This study also recommends using open-book examinations, but at the same time, multiple assessment methods must be used to reduce the unfairness of evaluation.

A limitation of this study is that the samples were drawn only from the statistics class the authors taught at the university. In addition to the small sample size, the representativeness needs to be improved. The research results are easily affected by variables such as countries, regions, institutions, departments, and teachers. Future extended research can be carried out to enhance the sample's representativeness to verify the research model.

Due to the small number of samples obtained in this study, the partial least squares research method was used to verify the proposed structural equation model. Future research can use different analytical methods to verify this study's hypothesis and research model.

References

Abbiati, N. N., Fabrizio, M. D. C., Lopez, M. V., Pérez, A., Plencovich, M. C., & Cueto, G. (2021). Attitudes related to students' performance in Statistics in university programs in Argentina. *Statistics Education Research Journal, 20*(2), 8-8. https://doi.org/10.52041/serj.v20i2.356

Agarwal, P. K., Karpicke, J. D., Kang, S. H., Roediger III, H. L., & McDermott, K. B. (2008). Examining the testing effect with open-and closed-book tests. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 22*(7), 861-876. https://doi.org/10.1002/acp.1391

American Association for the Advancement of Science. (2011). *Vision and change in undergraduate biology education: A call to action*. Washington, DC: American Association for the Advancement of Science.

Anaya, L., Evangelopoulos, N., & Lawani, U. (2010). *Open-book vs. closed-book testing: An experimental comparison*. Paper presented at the 2010 Annual Conference & Exposition (Louisville, KY), American Society for Engineering Education (ASEE).

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215. https://doi.org/10.1037//0033-295x.84.2.191

Bengtsson, L. (2019). Take-home exams in higher education: A systematic review. *Education Sciences, 9*(4), 267. https://doi.org/10.3390/educsci9040267

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education, 32*(3), 347-364.

Block, R. M. (2012). A discussion of the effect of open-book and closed-book exams on student achievement in an introductory statistics course. *Primus*, *22*(3), 228-238. https://doi.org/10.1080/10511970.2011.565402

Bolen, R. M. (2006). Utilizing web-based databases to introduce social work content in research statistics courses. *Social Work Education*, *25*(1), 17-27. https://doi.org/10.1080/02615470500477805

Boniface, D. (1985). Candidates' use of notes and textbooks during an open-book examination. *Educational Research*, *27*(3), 201-209. https://doi.org/10.1080/0013188850270307

Calderwood, K. A. (2002). Incorporating multiple epistemologies into teaching statistics to social work students. *Journal of Teaching in Social Work*, *22*(1-2), 17-32. https://doi.org/10.1300/J067v22n01_03

Cheng, S., Ferris, M., & Perolio, J. (2018). An innovative classroom approach for developing critical thinkers in the introductory statistics course. *The American Statistician*, *72*(4), 354-358. https://doi.org/10.1080/00031305.2017.1305293

Chiesi, F., & Primi, C. (2009). Assessing statistics attitudes among college students: Psychometric properties of the Italian version of the Survey of Attitudes toward Statistics (SATS). *Learning and Individual Differences*, *19*(2), 309-313. https://doi.org/10.1016/j.lindif.2008.10.008

Cladera, M., Rejón-Guardia, F., Vich-I-Martorell, G. À., & Juaneda, C. (2021). Evolution of undergraduate tourism students' attitudes towards Statistics: Pre and post evaluation. *Journal of Teaching in Travel & Tourism*, *21*(2), 120-141. https://doi.org/10.1080/15313220.2020.1828014

Conners, F. A., Mccown, S. M., & Roskos-Ewoldsen, B. (1998). Unique challenges in teaching undergraduate statistics. *Teaching of Psychology*, *25*(1), 40–42. https://doi.org/10.1207/s15328023top2501_12

DeVaney, T. A. (2010). Anxiety and attitude of graduate students in on-campus vs. online statistics courses. *Journal of Statistics Education*, *18*(1), 1-15. https://doi.org/10.1080/10691898.2010.11889472

Doménech-Betoret, F., Abellán-Roselló, L., & Gómez-Artiga, A. (2017). Self-efficacy, satisfaction, and academic achievement: The mediator role of students' expectancy-value beliefs. *Frontiers in Psychology*, *18*(8), 1193. https://doi.org/10.3389/fpsyg.2017.01193

Durning, S. J., Dong, T., Ratcliffe, T., Schuwirth, L., Artino Jr, A. R., Boulet, J. R., & Eva, K. (2016). Comparing open-book and closed-book examinations: A systematic review. *Academic Medicine*, *91*(4), 583-599. https://doi.org/10.1097/acm.0000000000000977

Earl, L. (2003). *Assessment as learning: Using classroom assessment to maximise student learning*. Thousand Oaks, CA: Corwin Press.

Emmioğlu, E., & Capa-Aydin, Y. (2012). Attitudes and achievement in statistics: A meta-analysis study. *Statistics Education Research Journal*, *11*(2), 95-102. https://doi.org/10.52041/serj.v11i2.332

Finney, S. J., & Schraw, G. (2003). Self-efficacy beliefs in college statistics courses. *Contemporary Educational Psychology*, *28*(2), 161-186. https://doi.org/10.1016/S0361-476X(02)00015-2

Fornell, C., & Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: Algebra and statistics. *Journal of Marketing Research*, *18*(3), 382-388. https://doi.org/10.2307/3150980

GAISE. (2016). *Guidelines for assessment and instruction in statistics education college report 2016*. Retrieved from GAISE College Report ASA Revision Committee:

Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, *70*(1), 1-25.

Garfield, J. (2011). *Statistical literacy, reasoning, and thinking in: Lovric, M. (Eds.), international encyclopedia of statistical science*. Berlin, Heidelberg: Springer.

Garfield, J., & Ben-Zvi, D. (2007). How students learn statistics revisited: A current review of research on teaching and learning statistics. *International Statistical Review*, *75*(3), 372-396. https://doi.org/10.1111/j.1751-5823.2007.00029.x

Hair Jr, F. J., Sarstedt, M., Hopkins, L., & G. Kuppelwieser, V. (2014). Partial least squares structural equation modeling (PLS-SEM) An emerging tool in business research. *European Business Review*, *26*(2), 106-121. https://doi.org/10.1108/EBR-10-2013-0128

Handelsman, J., Ebert-May, D., Beichner, R., Bruns, P., Chang, A., DeHaan, R., . . . Wood, W. B. (2004). Education: Scientific teaching. *Science*, *304*(5670), 521-522.

Hanoch, Y., & Vitouch, O. (2004). When less is more: Information, emotional arousal and the ecological reframing of the Yerkes-Dodson law. *Theory & Psychology*, *14*(4), 427-452. https://doi.org/10.1177/0959354304044926

Harrington, C. F., & Schibik, T. J. (2004). Methods for maximizing student engagement in the introductory business statistics course: A review. *Journal of American Academy of Business*, *4*(1/2), 360-364.

Haughton, J., & Kelly, A. (2015). Student performance in an introductory business statistics course: Does delivery mode matter? *Journal of Education for Business*, *90*(1), 31-43. https://doi.org/10.1080/08832323.2014.968518

Heijne-Penninga, M., Kuks, J. B., Hofman, W. A., & Cohen-Schotanus, J. (2010). Influences of deep learning, need for cognition and preparation time on open-and closed-book test performance. *Medical Education*, *44*(9), 884-891. https://doi.org/10.1111/j.1365-2923.2010.03732.x

Hilton, S. C., Schau, C., & Olsen, J. A. (2004). Survey of attitudes toward statistics: Factor structure invariance by gender and by administration time. *Structural Equation Modeling*, *11*(1), 92-109. https://doi.org/10.1207/S15328007SEM1101_7

Ioannidou, M. K. (1997). Testing and life-long learning: Open-book and closed-book examination in a university course. *Studies in Educational Evaluation*, *23*(2), 131-139. https://doi.org/10.1016/S0191-491X(97)00008-4

Johanns, B., Dinkens, A., & Moore, J. (2017). A systematic review comparing open-book and closed-book examinations: Evaluating effects on development of critical thinking skills. *Nurse Education in Practice*, *27*, 89-94. https://doi.org/10.1016/j.nepr.2017.08.018

Johannssen, A., Chukhrova, N., Schmal, F., & Stabenow, K. (2021). Statistical literacy—Misuse of statistics and its consequences. *Journal of Statistics and Data Science Education*, *29*(1), 54-62. https://doi.org/10.1080/10691898.2020.1860727

Johnson, C. M., Green, K. A., Galbraith, B. J., & Anelli, C. M. (2015). Assessing and refining group take-home exams as authentic, effective learning experiences. *Journal of College Science Teaching*, *44*(5), 61-71. https://doi.org/10.2505/4/jcst15_044_05_61

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York: Guilford Press.

Krathwohl, D. R. (2002). A revision of bloom's taxonomy: An overview. *Theory Into Practice*, *41*(4), 212-218.

Lee, J. K., & Calandra, B. (2004). Can embedded annotations help high school students perform problem solving tasks using a Web-based historical document? *Journal of Research on Technology in Education*, *37*(1), 65-84. https://doi.org/10.1080/15391523.2004.10782426

Macher, D., Paechter, M., Papousek, I., & Ruggeri, K. (2012). Statistics anxiety, trait anxiety, learning behavior, and academic performance. *European Journal of Psychology of Education*, *27*, 483-498. https://doi.org/10.1007/s10212-011-0090-5

Makar, K., & Rubin, A. (2009). A framework for thinking about informal statistical inference. *Statistics Education Research Journal*, *8*(1), 82-105.

Malik, S. (2015). Undergraduates' statistics anxiety: A phenomenological study. *The Qualitative Report*, *20*(2), 120-133.

Marson, S. M. (2007). Three empirical strategies for teaching statistics. *Journal of Teaching in Social Work*, *27*(3-4), 199-213. https://doi.org/10.1300/J067v27n03_13

McGrath, A. L., Ferns, A., Greiner, L., Wanamaker, K., & Brown, S. (2015). Reducing anxiety and increasing self-efficacy within an advanced graduate psychology statistics course. *Canadian Journal for the Scholarship of Teaching and Learning*, *6*(1), 5. https://doi.org/10.5206/cjsotl-rcacea.2015.1.5

Mitra, S. (2023). How are students learning in a business statistics course? Evidence from both direct and indirect assessment. *INFORMS Transactions on Education*, *23*(2), 95-103. https://doi.org/10.1287/ited.2022.0270

Murtagh, L., & Webster, M. (2010). Scaffolding teaching, learning and assessment in higher education. *Teacher Education Advancement Network Journal*, *1*(2), 1-20.

Oakleaf, M. (2008). Dangers and opportunities: A conceptual map of information literacy assessment approaches. *Portal: Libraries and the Academy*, *8*(3), 233-253. https://doi.org/10.1353/pla.0.0011

Onwuegbuzie, A. J. (2000). Statistics anxiety and the role of self-perceptions. *The Journal of Educational Research*, *93*(5), 323-330. https://doi.org/10.1080/00220670009598724

Onwuegbuzie, A. J. (2004). Academic procrastination and statistics anxiety. *Assessment & Evaluation in Higher Education*, *29*(1), 3-19. https://doi.org/10.1080/0260293042000160384

Onwuegbuzie, A. J., Da Ros, D., & Ryan, J. M. (1997). The components of statistics anxiety: A phenomenological study. *Focus on Learning Problems in mathematics*, *19*(4), 11-35.

Onwuegbuzie, A. J., & Leech, N. L. (2003). A framework for assessing the validity of research: Adapting and applying research validation methods. *Quality & Quantity*, *37*(2), 167-193. https://doi.org/10.1023/A:1024188206473

Pan, W., & Tang, M. (2005). Students' perceptions on factors of statistics anxiety and instructional strategies. *Journal of Instructional Psychology*, *32*(3), 205–214.

Pea, R. D. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *Journal of the Learning Sciences*, *13*(3), 423-451. https://doi.org/10.1207/s15327809jls1303_6

Salcedo, A. (2014). Statistics test questions: Content and trends. *Statistics Education Research Journal*, *13*(2), 202-217. https://doi.org/10.52041/serj.v13i2.291

Sato, B. K., He, W., Warschauer, M., & Kadandale, P. (2015). The grass isn't always greener: Perceptions of and performance on open-note exams. *CBE—Life Sciences Education*, *14*(2), ar11. https://doi.org/10.1187/cbe.14-08-0121

Saye, J. W., & Brush, T. (2002). Scaffolding critical reasoning about history and social issues in multimedia-supported learning environments. *Educational Technology Research and Development*, *50*(3), 77-96. https://doi.org/10.1007/bf02505026

Schau, C. (2003). *Students' attitudes: The "other" important outcome in statistics education Parts of this paper were presented at the Joint Statistics Meetings, San Francisco*. Retrieved from http://evaluationandstatistics.com/JSM2003.pdf

Schau, C., & Emmioğlu, E. (2012). Do introductory statistics courses in the United States improve students' attitudes? . *Statistics Education Research Journal*, *11*(2), 86-94. https://doi.org/10.52041/serj.v11i2.331

Schau, C., Stevens, J., Dauphinee, T. L., & Del Vecchio, A. (1995). The development and validation of the survey of attitudes toward statistics. *Educational and Psychological Measurement*, *55*(5), 868-875. https://doi.org/10.1177/0013164495055005022

Sharma, S. (2017). Definitions and models of statistical literacy: A literature review. *Open Review of Educational Research*, *4*(1), 118-133. https://doi.org/10.1080/23265507.2017.1354313

Sherwood, C., & Kwak, D. W. (2017). New insights into an old problem–enhancing student learning outcomes in an introductory statistics course. *Applied Economics*, *49*(56), 5698-5708. https://doi.org/10.1080/00036846.2017.1332750

Shin, S., Brush, T. A., & Glazewski, K. D. (2020). Examining the hard, peer, and teacher scaffolding framework in inquiry-based technology-enhanced learning environments: impact on academic achievement and group performance. *Educational Technology Research and Development*, *68*, 2423-2447. https://doi.org/10.1007/s11423-020-09763-8

Simkin, M. G. (2005). An experimental study of the effectiveness of collaborative testing in an entry-level computer programming class. *Journal of Information Systems Education*, *16*(3), 273-280.

Sizemore, O. J., & Lewandowski, G. W. (2009). Learning might not equal liking: Research methods course changes knowledge but not attitudes. *Teaching of Psychology*, *36*(2), 90-95. https://doi.org/10.1080/00986280902739727

Soesmanto, T., & Bonner, S. (2019). Dual mode delivery in an introductory statistics course: Design and evaluation. *Journal of Statistics Education*, *27*(2), 90-98. https://doi.org/10.1080/10691898.2019.1608874

Stiggins, R. J. (2006). Assessment for learning: A key to student motivation and learning. *Phi Delta Kappa Edge*, *2*(2), 1-19.

Suanpang, P., Petocz, P., & Kalceff, W. (2004). Student attitudes to learning business statistics: Comparison of online and traditional methods. *Journal of Educational Technology & Society*, *7*(3), 9-20.

Teodorczuk, A., Fraser, J., & Rogers, G. D. (2018). Open book exams: A potential solution to the "full curriculum"? *Medical Teacher*, *40*(5), 529-530. https://doi.org/10.1080/0142159X.2017.1412412

Theophilides, C., & Dionysiou, O. (1996). The major functions of the open-book examination at the university level: A factor analytic study. *Studies in Educational Evaluation*, *22*(2), 157-170. https://doi.org/10.1016/0191-491X(96)00009-0

Theophilides, C., & Koutselini, M. (2000). Study behavior in the closed-book and the open-book examination: A comparative analysis. *Educational Research and Evaluation*, *6*(4), 379-393. https://doi.org/10.1076/edre.6.4.379.6932

Vaessen, B. E., Prins, F. J., & Jeuring, J. (2017). University students' perceptions of the quality of teacher-provided online feedback in blended learning. *The Internet and Higher Education*, *33*, 41-53. https://doi.org/10.1016/j.iheduc.2017.01.002

Waters, L., Martelli, T. A., Zakrajsek, T., & Popovich, P. M. (1988). Attitudes toward statistics: An evaluation of multiple measures. *Educational and Psychological Measurement*, *48*(2), 513-516. https://doi.org/10.1177/0013164488482026

Wathen, S., & Rhew, N. D. (2019). Using real-life major league baseball data in an introductory statistics course. *Decision Sciences Journal of Innovative Education*, *17*(3), 194-213. https://doi.org/10.1111/dsji.12185

Williams, A. S. (2010). Statistics anxiety and instructor immediacy. *Journal of Statistics Education*, *18*(2), 1-18. https://doi.org/10.1080/10691898.2010.11889495

Williams, J. B., & Wong, A. (2009). The efficacy of final examinations: A comparative study of closed-book, invigilated exams and open-book, open-web exams. *British Journal of Educational Technology*, *40*(2), 227-236. https://doi.org/10.1111/j.1467-8535.2008.00929.x

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, *17*(2), 89-100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

Zackariasson, M. (2020). Encouraging student independence: Perspectives on scaffolding in higher education supervision. *Journal of Applied Research in Higher Education*, *12*(3), 495-505. https://doi.org/10.1108/JARHE-01-2019-0012

Zare, H., Rastegar, A., & Hosseini, S. M. D. (2011). The relation among achievement goals and academic achievement in statistics: The mediating role of statistics anxiety and statistics self-efficacy. *Procedia-Social and Behavioral Sciences*, *30*, 1166-1172. https://doi.org/10.1016/j.sbspro.2011.10.227

Zeidner, M. (1991). Statistics and mathematics anxiety in social science students: Some interesting parallels. *British Journal of Educational Psychology*, *61*(3), 319-328. https://doi.org/10.1111/j.2044-8279.1991.tb00989.x

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. https://doi.org/10.52041/serj.v17i2.159