ACTA SCIENTIAE

(cc) BY

ISSN: 2178-7727
DOI: 10.17648/acta.scientiae.7262

Attitudes Towards Statistics and Statistical Reasoning of Teachers in Training

Jaime Andrés Gaviria Bedoya [a,b]
Difariney González Gómez [a]
Mónica Marcela Parra-Zapata [b]
Jhony Alexander Villa-Ochoa [b]

[a] Universidad de Antioquia, Facultad Nacional de Salud Pública, Medellín, Colombia
[b] Universidad de Antioquia, Facultad de Educación, Medellín, Colombia

Received for publication 13 Jun. 2022. Accepted after review 4 sep. 2022
Designated editor: Claudia Lisete Oliveira Groenwald

ABSTRACT

Background: Knowing the attitudes towards statistics and the statistical reasoning of teachers in training has been a topic of interest in research in statistical education because they reveal affective and cognitive characteristics that are part of the teaching and learning process of statistics. **Objectives**: To describe attitudes towards statistics and statistical reasoning about descriptive measures and probability of prospective primary education teachers within the framework of a course of their curriculum**. Design**: Case study with quantitative data and analysis. **Instruments and participants:** The Survey of Attitudes Towards Statistics (SATS) and the Statistical Reasoning Assessment (SRA) were applied at the beginning and end of the course. The sample consisted of 18 prospective teachers of a program of primary education of an official university of Colombia. **Data collection and analysis**: For the total scores in the two SATS and SRA applications, we obtained descriptive measurements, performed bivariate analyses, and applied non-parametric methods such as the Wilcoxon test. **Results**: We found a statistical differences in attitudes towards statistics between the pre-test and post-test. No statistical differences were found between the statistical reasoning scores. **Conclusions**: The prospective teachers presented a positive assessment of the usefulness of statistics and the role of statistics in the sociocultural environment and showed erroneous reasoning, mainly about probability.

Keywords: Statistical education; Statistical reasoning; Attitudes towards statistics; Teachers in training; Primary education.

Corresponding author: Jaime Andrés Gaviria Bedoya. E-mail: jaime.gaviria@udea.edu.co

Actitudes hacia la estadística y razonamiento estadístico de profesores y profesoras en formación

RESUMEN

Antecedentes: Conocer las actitudes hacia la estadística y el razonamiento estadístico de los profesores en formación ha sido un tema de interés en la investigación en educación estadística porque revelan características afectivas y cognitivas que hacen parte del proceso de enseñanza y aprendizaje de la estadística. **Objetivos**: Describir las actitudes hacia la estadística y el razonamiento estadístico acerca de las medidas descriptivas y la probabilidad de los futuros profesores de educación infantil en el marco de un curso de su plan de estudios**. Diseño**: Estudio de caso con datos y análisis cuantitativos. **Instrumentos y participantes:** Se aplicó la Escala para Evaluar las Actitudes hacia la Estadística (EAEE) y el instrumento Statistical Reasoning Assessment (SRA) al inicio y al final del curso. La muestra estuvo conformada por 18 futuros profesores de un programa de Educación Infantil de una universidad oficial de Colombia. **Recolección y análisis de datos**: Para las puntuaciones totales en las dos aplicaciones de la EAEE y el SRA se obtuvieron medidas descriptivas, se realizaron análisis bivariados y se aplicaron métodos no paramétricos como la prueba de Wilcoxon. **Resultados**: Se encontró diferencia estadística en las actitudes hacia la estadística entre el pre-test y el post-test. No se encontraron diferencias estadísticas entre las puntuaciones del razonamiento estadístico. **Conclusiones**: Los futuros profesores presentaron una valoración positiva de la utilidad de la estadística y del papel de la estadística en el ámbito sociocultural y mostraron razonamientos erróneos principalmente acerca de la probabilidad.

Palabras clave: educación estadística; razonamiento estadístico; actitudes hacia la estadística; profesores en formación; educación infantil.

Atitudes em relação à estatística e ao raciocínio estatístico de professores e professoras em formação

RESUMO

Antecedentes: A compreensão das atitudes relativamente à estatística e ao raciocínio estatístico dos professores(as) em formação tem sido um tema de interesse na investigação da educação estatística, porque estes aspectos revelam características afetivas e cognitivas que fazem parte do processo de ensino e aprendizagem da estatística. **Objetivos:** Descrever as atitudes em relação à estatística e ao raciocínio estatístico, sobre medidas descritivas e probabilidade, dos futuros(as) professores(as) de educação infantil no quadro de um curso do seu currículo. **Desenho:** Estudo de caso com dados quantitativos e análise. **Instrumentos e participantes:** A Escala de Avaliação das Atitudes Estatísticas (EAEE) e a Avaliação de Raciocínio Estatístico (SRA) foram administradas no início e no fim do curso. A amostra consistiu em 18 futuros(as) professores(as) de um programa de educação infantil numa universidade oficial na Colômbia. **Coleta e análise de dados:** Para as pontuações totais nas duas aplicações do EAEE e do SRA, foram obtidas medidas descritivas, foram realizadas análises bivariadas e aplicados métodos não paramétricos como o teste Wilcoxon. **Resultados:** Foram encontradas diferenças estatísticas nas atitudes em relação à estatística entre o pré-teste e o pós-teste. Não foi encontrada qualquer diferença estatística entre as pontuações de raciocínio estatístico. **Conclusões:** Os futuros professores apresentaram uma avaliação positiva da utilidade das estatísticas e do papel das estatísticas no domínio sociocultural e demonstraram um raciocínio errôneo principalmente sobre a probabilidade.

Palavras-chave: educação estatística; raciocínio estatístico; atitudes em relação à estatística; futuros(as) professores(as) ; educação infantil.

INTRODUCTION

Globally, the democratisation of statistics has broadened and diversified the education, interests, and motivations of those who study it. Statistics are no longer reserved for future scientists but are taught to students of several school levels with very different interests and objectives (Carver et al., 2016). The incorporation of statistical education into school curricula in various countries confirms the growing demand for statistical education and its inclusion in curricula at all school levels (Estrada et al., 2013; González-Gómez, 2014; Lopes, 2008; Ministerio de Educación Nacional (MEN), 1998; Zieffler et al., 2018).

Studies on teacher education in statistics in recent years have attracted a growing number of researchers (Estrada et al., 2004; González-Gómez, 2014; Groth and Meletiou-Mavrotheris, 2018). Since the research focuses on the statistical knowledge of pre-service and in-service teachers, more attention is required to some aspects of didactic-stochastic knowledge, including affective aspects such as attitudes towards statistics (Batanero, 2019).

A relevant component in teaching activity is the teachers' attitude toward statistics, because it is a factor that can influence learning, statistical behaviour inside and outside the classroom, and willingness to attend statistical courses in the future (Estrada et al., 2011). Moreover, teachers' negative attitudes towards statistics could be reflected in a lack of confidence or distaste when teaching statistics (Ruz et al., 2020) and could, in turn, influence students' attitudes, thus creating a vicious circle (Estrada, 2002).

Owing to the limited literature on teachers' attitudes towards statistics and their teaching (Estrada et al., 2011; Groth and Meletiou-Mavrotheris, 2018; Ruz et al., 2020), research on attitudes towards statistics of teachers in primary education is limited (León-Montero et al., 2018; Lopes, 2008). This poses a challenge in identifying the attitudes towards statistics of future primary education teachers.

On the other hand, knowing how the cognitive aspects of teachers in training are is important because, as it has been proposed by Groth and Meletiou-Mavrotheris (2018), cognitive characteristics are related to affective factors such as attitudes and beliefs about the teaching of statistics. In particular, the education of teachers who teach statistics requires different experiences and tools than those used to train teachers who teach mathematics (Garfield et al., 2008). In this sense, statistics teaching based on understanding mathematical algorithms and statistical techniques is insufficient (González-Gómez, 2014); that is why statistics teaching that allows teachers to use statistical reasoning in their professional practice is necessary (Bargagliotti et al., 2020; Batanero, 2019; Franklin et al., 2007; Garfield, 2002).

Some studies have shown that prospective elementary and secondary education teachers lack the statistical reasoning necessary to determine when and why to use statistical concepts such as mean, median, or variance (Karatoprak et al., 2015). It is therefore essential to involve them during their training process in activities that provide opportunities to promote statistical reasoning (Henriques et al., 2019).

According to Estrada et al. (2004), “statistics is a subject that teachers often forget in mandatory education, despite its recognised usefulness and being included in the curricular guidelines” (p. 264). Pre-service and in-service teachers may have a diffuse understanding of statistics or be unaware of the aspects of life in which statistics can be used (Estrada et al., 2011), possibly due to the insufficient training in statistics during their studies, which may negatively impact their teaching practice and attitudes towards statistics (González-Gómez and Zapata-Cardona, 2018). Therefore, this research aims to describe attitudes towards statistics and statistical reasoning about the descriptive measures and probability of prospective primary education teachers within the framework of a course of their curriculum.

Research questions:

What are the attitudes towards statistics of prospective primary education teachers?

What is the statistical reasoning behind descriptive measures and probability of prospective primary education teachers?

The present article is organised into four sections. The first presents the theoretical framework on which statistical reasoning and attitude towards statistics are based. Then, we present the methodology, including the class design, activities, procedures, and statistical methods implemented. Next, the results present the analyses of the two measuring instruments applied and the comparison between the participants' scores. Finally, the discussion and conclusions present the main findings in dialogue with the literature, after which we share the study limitations and make recommendations for future research.

THEORETICAL FRAMEWORK

The framework supporting this study is based on previous research on attitudes towards statistics and statistical reasoning.

Attitudes towards statistics

The literature has shown the importance of the students' attitude towards statistics, especially that of university-level students who specialise in the disciplines of science, technology, engineering, and mathematics (STEM) (Gundlach et al., 2015). However, students from other disciplines, including the social sciences, may not have a positive attitude towards statistics and may feel anxious about those courses, which may negatively affect their academic performance (Dempster & Mccorry, 2009; Tishkovskaya & Lancaster, 2012).

The attitude towards statistics is part of a psychological construct that can be used to understand and predict individuals' reactions and how their behaviour can be influenced (Estrada et al., 2011). Different studies indicate that attitude is a multidimensional construct of both pedagogical components (including the affective, cognitive, and behaviour) and anthropological components (including the social, educational, and instrumental) (Estrada, 2002; Groth & Meletiou-Mavrotheris, 2018; León-Montero et al., 2018).

An understanding generally shared by the international academic community is that the attitude toward statistics is a psychological tendency that arises over time due to emotions and feelings experienced in the context of mathematics and statistics learning (Márquez, 2004). Consequently, and following Estrada et al. (2004), this investigation understands attitude towards statistics as the set of emotions and feelings experienced during the learning period of statistics and probability.

One of the instruments used to measure the attitude towards statistics is the Survey Attitude Towards Statistics -SATS- (Estrada, 2002). This scale has been widely used in research on non-cognitive factors in active and prospective teachers and has consistently demonstrated acceptable psychometric properties (Estrada et al., 2004, 2013). Following a multidimensional vision of attitudes towards statistics, the SATS considers the domains and components (Estrada et al., 2013):

1. Pedagogical components:

1.1. Cognitive: It refers to conceptions and beliefs about statistics.

1.2. Affective or emotional: It includes emotions and feelings that statistics awaken; therefore, they are more subjective.

1.3. Behavioural or tendential: It represents the tendency to act or have a specific intention, for example, how and when statistical intention would be used in a particular way.

2. Anthropological components:

2.1. Social: Related to the perception and assessment of the role of statistics in the sociocultural environment of any citizen.

2.2. Educational: Linked to the interest in statistics and its learning, the vision of its usefulness for the student, their opinion on whether it should be included in the curriculum and the perceived difficulty.

2.3. Instrumental: Usefulness for other curriculum subjects, as a form of reasoning and as a cultural component.

The previous structure of the STAS is aligned with 25 statements: 14 were made positively and 11 negatively. Statements have a rating scale of 1 to 5 points. For positive questions, a score of 5 corresponds to “Strongly agree” and 1 to “Strongly disagree”. The opposite happens with negative questions, a score of 5 corresponds to “Strongly disagree” and 1 to “Strongly agree”. According to Estrada (2002), this is because the calculus of the total score of the instrument requires that all the questions are in the same direction, so that a total score corresponds to a positive attitude towards statistics.

Regarding the validity and reliability of the SATS, Estrada (2002) reported that the content validity of the scale changed from the 36-item version to the final 25-item version, with 75% agreement between the judges and the internal consistency of the α= 0.7741 scale with a sample of 140 teachers (74 pre-service and 66 in-service).

In particular, Pérez et al. (2015) implemented and validated the SATS with a sample of 545 university students from Colombia at the beginning of their first academic semester. The students belonged to nine programmes of the professional schools of exact sciences and engineering, economic sciences and the International School of Administration and Professional Marketing. The authors found that students recognise the importance of statistics in the academic world and everyday life; however, the latter distrust the use, the skills required, and the taste for the discipline they study. These authors estimated the reliability of each scale item using a Cronbach's coefficient of $\alpha>0.8$, indicating that the scale can be considered adequate.

The SATS has been applied in previous studies related to prospective teachers. For example, León-Montero et al. (2018) applied the scale to a random sample of 105 pre-service primary education teachers from a university in Spain. They found that the prospective teachers value statistics positively, considering it useful for their training and the future of their professional practice. However, the authors found that students dislike statistics and avoid using it outside the classroom.

Statistical reasoning

Statistical reasoning is considered one of the important cognitive outcomes in the education of students who take statistical courses (Andrade et al., 2017; Ben-Zvi et al., 2018; Carver et al., 2016; Garfield & Ben-Zvi, 2004). This reasoning is essential because people can develop abstract reasoning processes and identify patterns to make inferences and draw conclusions beyond the data (delMas, 2004).

One of the common understandings found in research and adopted in this paper is to define statistical reasoning as the way people with statistical ideas reason and make sense of statistical information, which involves making interpretations based on data sets, representations, or statistical summaries of data (Garfield, 2003). From this perspective, statistical reasoning links variability, uncertainty, and context to make plausible interpretations and that these make sense with data (Ben-Zvi, 2018; delMas, 2004; Moore, 1998; Pfannkuch & Wild, 1999).

Regarding studies on students' statistical reasoning, when investigating statistical education, it is important and necessary to have validated instruments that account for the students' statistical reasoning at different school levels. From a quantitative perspective, studies on students' statistical reasoning have generally used assessment tools such as objective tests with multiple-choice questions (Zieffler et al., 2008).

One proven instrument for measuring statistical reasoning is the Statistical Reasoning Assessment - SRA (Garfield, 1991). The SRA aims to evaluate statistical reasoning behind a sample of statistical concepts in students of an introductory course of high school or first-year university statistics and was created in response to the need to evaluate the development of a technology-based statistical curriculum (Garfield, 1991). The SRA is one of the first to assess statistical reasoning, and although it has been applied mainly in English-speaking countries, because of the growing interest of the international community, it has been translated and adapted into other languages, such as Spanish and French (Liu & Garfield, 2002).

The SRA consists of 20 multiple-choice questions, with single or multiple answers, and evaluates reasoning on data, data representation, centre measures, variability and position, uncertainty, samples, and association (Garfield, 2003). In addition, the instrument is theoretically based on two types of statistical reasoning: correct and erroneous, reported by researchers from psychology (Tversky & Kahneman, 1982) and statistical education areas (Garfield, 2003) during the seventies and eighties (Lovett, 2001). Besides, each type of reasoning consists of eight rating scales (Table 1).

Thus, for example, as far as erroneous statistical reasoning is concerned, the SRA is based on the research carried out by Tversky & Kahneman, (1982), who revealed some inconsistent ways of thinking about some statistical concepts and ideas. These investigations suggest that even people who can correctly calculate probabilities tend to apply erroneous reasoning when making inferences or judgments about an event subject to uncertainty.

Table 1

SRA Correct and Erroneous Statistical Reasoning Scales. **(**adapted from Garfield, 2003, p. 27)

Type of reasoning	Scales
Correct statistical reasoning	1. Correctly interprets probabilities
	2. Understands how to select an appropriate average
	3. Correctly computes probabilities
	4. Understands independence
	5. Understands sampling variability
	6. Distinguishes between correlation and causation
	7. Correctly interprets 2x2 (two-way) tables
	8. Understands the importance of large samples
Erroneous statistical reasoning	1. Misconceptions involving averages
	2. Outcome orientation
	3. Good samples have to represent a high percentage of the population
	4. Law of small numbers
	5. Misconceptions about representativeness
	6. Correlation implies causation.
	7. Equiprobability bias
	8. Groups can only be compared if they have the same sample size.

Regarding the validity and reliability of the SRA, Garfield (2003) reports the validity of the content with the agreement between a group of experts; the validity related to the criterion –the correlations between the scores of the SRA and the different results of the course– was low; finally, the reliability of the SRA was estimated by means of the test-retest applied to a sample of 32 teachers in training with a time interval of one week. It is reported that the reliability analysis gave a correlation of 0.70 for the total correct statistical reasoning scores and 0.75 for the erroneous reasoning. The translation and validation of the SRA into the Spanish language were carried out with a sample of 113 teachers in training from a university in Spain. They estimated an internal consistency of α=0.70 (Estrada, 2002).

METHODOLOGY

Participants and study context

This work summarises the results of a research with a group of 18 prospective teachers attending a degree in primary education at a public university in Medellín, Colombia, in 2019. Most students were women; the average age was 23.3 years (standard deviation 4.8 years), 53% were in the fifth semester, 66% had only taken one statistics course and most had studied in public schools (88.9%).

The research was developed within the course Mathematical Thinking, a mandatory course of the academic program, held in the second semester of 2019. The course consisted of 64 hours distributed in four-hour weekly sessions and was divided into three modules: random, numerical, and variational thinking. The random thinking module, about a third of the intensity of the course, focused on developing statistics and probability topics. The researchers, in agreement with the professor in charge of the course, planned the design of the random thinking module with a set of tasks and workshops to promote the participants' statistical reasoning.

In the design of the random thinking module, strategies reported in the literature to promote statistical reasoning were considered, such as: working with real data in context, critical reading of articles and using technology for data analysis (Carver et al., 2016). Classroom learning activities were implemented following the Guidelines for Assessment and Instruction in Statistical Education (GAISE) (Carver et al., 2016) and the Statistical Reasoning Learning Environment (SRLE) (Garfield et al., 2008). The implemented activities are described below.

Five four-hour class sessions were scheduled during the module. Each session was divided into three instances: 1) introductory reading related to the fundamentals of descriptive statistics or probability; 2) development of the topic through activities with the support of materials such as paper grids, markers, press clippings and magazines, among others; and 3) summary of the main aspects of the session. The emphasis of the activities developed in the classroom was to promote statistical reasoning on descriptive measures and probability, for which teachers in training were encouraged to use real data, as well as to understand and interpret the results in their context. The schedule of the sessions developed during the module was as follows:

1. General introduction of the research and application of the pre-test of the two instruments used to analyse attitudes towards statistics and statistical reasoning.
2. Socialisation of the experience around the completion of the two instruments with prospective teachers and introductory readings on random thinking in primary education statistics.
3. Study of the main summary measures and statistical graphs indicating how to obtain them in Microsoft Excel and interpretation of the results according to the contexts. Analysis of the advantages and disadvantages of using one measure or another in each specific context.
4. Development of activities developed in class to promote statistical reasoning, such as understanding the existence of different ways of collecting information from surveys or the main aspects of probability through a simulation experiment.
5. Application of the post-test of instruments referring to the attitude towards statistics and statistical reasoning; and development of an activity of analysis of some conceptual maps on the main descriptive measures observed in the previous sessions.

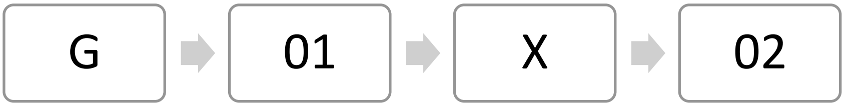
Procedures and instruments

In this research, we conducted a case study with quantitative data and analysis (Gillham, 2000). The sample was non-probabilistic due to the regular characteristics of the course, which ranged from 10 to 25 students enrolled (Creswell, 2014; Szklo & Nieto, 2003) and consisted of 18 pre-service teachers of a primary education program at a public university in Medellín, Colombia.

A pre-test/post-test quasi-experimental design was performed with a single group to provide quantitative information on prospective teachers' attitudes towards statistics and their statistical reasoning (Figure 1). This type of studies is useful when the group is already established and allows the generalisation of results to the target population through small samples (Creswell, 2014). The independent variable consists of the educational intervention performed by the design of the random thinking module. There are two dependent variables: attitudes towards statistics and statistical reasoning.

Figure 1

Outline of the quasi-experimental design with pre-test/post-test.



Note. The notation used in the scheme: G represents the single group, X the treatment (educational intervention), 01 the pre-test and 02 the post-test for the SATS and SRA instruments.

To provide quantitative information on the dependent variables, two instruments were applied at the beginning and the end of the random thinking module: the SATS (Estrada, 2002) for the attitude towards statistics and the SRA (Garfield, 2002, 2003) for the statistical reasoning of teachers in training. The teachers in training anonymously completed the two instruments and signed the informed consent, in which they were notified that their participation in the study was voluntary and that the results obtained would be used for academic purposes. To respect anonymity, each participant was assigned a code to pair the measurements of both instruments at the beginning and the end of the course. Finally, the results of both instruments were tabulated and organised into a Microsoft Excel file.[1]

[1]No prior ethical evaluation was carried out since the results of this study are derived from a didactic innovation project, which did not require the endorsement of an ethics committee. The draft was reviewed by a technical committee for approval. Therefore,

The total scores of the instruments were calculated for each application. The total SATS score corresponds to the result of adding the scores of each statement and represents the attitude towards statistics of the students surveyed. The lowest score that can be obtained is 25, and the highest score is 125, with 75 being the average score, or “indifference” (Estrada, 2002). With the results of the SATS, we performed two analyses, one by statements and the other for the total scores in each application.

As for the SRA, a total score was obtained for each type of reasoning, with a maximum score of 21 points for correct reasoning and 32 for erroneous reasoning (Garfield, 2002, 2003). With the total scores, an analysis was made of the percentage of correct and erroneous answers by groups of questions according to the scales of each type of reasoning.

In the descriptive analysis, the mean, median, and standard deviation of the total scores of both instruments were obtained in the pre-test and post-test. Bivariate analyses were also performed between the total scores in each application and sociodemographic variables, such as age, using Spearman’s correlation. In the statistical inference analysis, we used non-parametric methods since the dependent variables did not follow a normal distribution. Also, because the data structure corresponds to paired samples, we implemented the two-tailed non-parametric Wilcoxon test. To determine if there was a statistical difference between the ranges of the total scores of each dependent variable between the pre-test and the post-test, we obtained the p-value of the test and its respective confidence interval. All statistical analyses were performed with statistical software R (version 4.1.3).

RESULTS AND ANALYSIS

The results were organised in the following sections: analysis of the attitudes towards statistics of teachers in training and analysis of their statistical reasoning.

---

the authors explicitly exempt Acta Scientiae from the consequences derived from this research, including comprehensive assistance and eventual compensation for any damage resulting to/from any of the participants.

Analysis of attitudes towards statistics of teachers in training

We obtained the mean and standard deviation for each of the 25 SATS statements (Table 2). We found that the following questions reached an average score greater than or equal to four in the pre-test and the post-test: 2. Statistics help to understand today's world, 6. In the career that I study, statistics should not be taught, 19. Statistics only work for people in science, 21. Statistics are useless, and 23. If I could eliminate any subject, it would be statistics. All these questions correspond to the cognitive pedagogical component except for question 23. Together, these results show that, in the two applications of the SATS, teachers in training have a positive view of the usefulness of statistics and a good perception and appreciation of the role of statistics in the sociocultural environment.

The questions with the lowest means (less than 3) in the pre-test were: 1, 3 and 8. Question 1. I resent the statistical information that appears in some TV programs, it belongs to the affective-social components; question 3. Through statistics, you can manipulate reality, the cognitive-instrumental component and 8. I find statistics problems easy for the behavioural educational component (Estrada, 2002). Questions 1 and 8 show that, at the beginning of the random thinking module, teachers in training are generally uncomfortable with how statistical information is presented in the media and consider using statistics scarcely outside the university.

Regarding the score in question 3. Through statistics, reality can be manipulated, it was the only one to have an average score of less than three in the two applications of the instrument ($\bar{x}_{pre=}2.4; \bar{x}_{post} = 2.6$). This question refers to the image that students have of statistics. The results suggest that the prospective teachers mistrust the usefulness of statistics in other areas, as they consider that statistical results are easily manipulated. These results are consistent with those obtained by Estrada et al. (2004) and León-Montero et al. (2018), who found that prospective teachers distrust the statistical results published in the mass or popular media. Authors such as (Pérez et al., 2015) point out question 3 as problematic because, when applying the SATS to a sample of 350 university students from a university in Colombia, they found that this question scored the least, and its correlation with the other questions of the instrument was also one of the lowest.

Table 2

Descriptive statistics for the SATS questions

	Pre-test		Post-test	
Item	$\bar{X}$	SD	$\bar{X}$	SD
1	2.9	0.9	3.3	0.8
2	4.4	0.6	4.6	0.5
3	2.4	1.2	2.6	1.5
4	4.1	0.7	3.9	1.1
5	3.1	1.3	3.8	0.9
6	4.6	0.5	4.2	1.3
7	3.6	1.0	3.8	0.6
8	3.0	1.0	3.8	0.8
9	3.2	0.8	3.9	0.8
10	3.7	0.9	3.8	0.7
11	3.8	0.8	3.9	0.9
12	3.7	0.7	4.0	0.6
13	3.2	0.9	3.7	0.7
14	2.8	0.9	3.5	0.9
15	4.0	1.0	4.1	0.9
16	3.6	0.7	3.8	0.7
17	3.2	0.8	3.7	0.8
18	4.2	1.3	3.8	1.0
19	4.8	0.4	4.8	0.5
20	3.2	0.5	3.5	0.7
21	4.8	0.9	4.7	0.6
22	3.1	1.0	3.8	0.7
23	4.3	0.8	4.3	1.0
24	3.7	0.7	4.1	0.7
25	3.5	1.0	3.8	1.0

Note. $\bar{X}$ denotes the arithmetic mean and SD the standard deviation.

Regarding the global analysis of the attitude towards statistics, to determine whether there was a statistical difference between the pre-test and the post-test, we compared the total scores of the attitude towards statistics at the beginning and the end of the course and calculated the size of the effect and its confidence interval using the Hedges' g (Table 3). These results show an increase of almost ten points in the average SATS score. In addition, the size of the effect obtained can be considered average, from which we can conclude that the mean attitudinal scores towards statistics between the pre and the post-test have a standardised difference of 0.6 units and the negative sign indicates that the average obtained in the pre-test is lower than that obtained in the post-test; this evidences an improvement in the attitude towards statistics after completing the random thinking module.

Table 3

Descriptives and size of the effect for total SATS scores.

	$\bar{X}$ (SD)	g	CI 95%
Pre-test	89.6 (10.3)	-0.624	(-0.895,-0.352)
Post-test	97.5 (11.8)		

Note. $\bar{X}$ denotes the arithmetic mean, SD the standard deviation, and g the size of the effect estimated by the Hedges statistic.

Table 4

Wilcoxon non-parametric test for SATS total scores (authors' elaboration)

	Mdn	V	CI 95%
Pre-test	91.5	1*	(-11.5,-5)
Post-test	100		

*p<0.01, Mdn denotes the median, and V is the statistic test of the Wilcoxon test.

Complementarily, the median score between the pre-test was 91.5 points, compared to that of the post-test, which was 100 points. The two-tailed non-parametric Wilcoxon test found a statistical difference between the total scores in the pre-test and the post-test (Table 4). These results suggest an increase in the attitude towards statistics between the pre-test and the post-test, although we cannot affirm that these improvements are due to the redesigning of the module on account of the type of study implemented.

The reliability of the SATS was estimated by the test-retest method with Spearman's correlation of the total scores between the pre-test and the post-test, $r_s = 0.537$,95% CI [0.075,.809]. These results show a moderate degree of stability over time in attitude scores toward statistics. In addition, a statistically significant correlation was found between the total SATS instrument score at the post-test and the participants' ages (Spearman's r=0.55, p=0.022, 95% CI [0.093, 0.815]). The correlation obtained is moderate and indicates some tendency for more adult prospective teachers to obtain higher global scores of attitude towards statistics.

Analysis of the statistical reasoning of teachers in training

For each of the 20 SRA questions, we estimated the number of participants who answered correctly and erroneously, obtaining the corresponding percentage of answers. Then, we grouped the percentages of the questions according to the four themes shown in Table 5 and estimated the average percentage for each theme.

In the reasoning about descriptive statistics and probability, participants had a higher percentage of correct answers in both instrument applications, while the lower percentage was in the reasoning about samples and the reasoning behind the association between variables. In addition, the average percentage on the correct statistical reasoning scale on probabilities decreased slightly from 51.1% in the pre-test to 47.6% in the post-test. As for the scale of erroneous statistical reasoning: there was a decrease in three of the four themes; the reasoning on probability was the only one to increase.

Table 5

The average percentage of correct and erroneous answers by subject in the SRA instrument.

	Pre-test		**Post-test**	
Subject	% Average of correct answers	% Average wrong answers	% Average of correct answers	% Average wrong answers
Descriptive statistics	53.7	50	58,8	35.3
Probability	51.1	40.1	47.6	43.1
Samples	26.4	41.1	30.6	37.6

Association	41.7	72.2	44.1	52.9

With the results of the SRA instrument, the average scores and the size of the effect of correct and incorrect reasoning were also obtained in the two applications (Table 6). These results show that the prospective teachers obtained a higher mean score on the post-test with a mean difference greater than 0.5, which can be interpreted as practically significant (Schau & Emmioğlu, 2012). Furthermore, although the size of the effect between the pre-test and the post-test is small, the negative sign confirms that, on average, the participants in the post-test scored above in the correct statistical reasoning compared to the mean scores in the pre-test.

For the wrong statistical reasoning, the mean score obtained in the pre-test was higher than that obtained in the post-test, which is corroborated by a sign of Hedges' statistic. Although the effect size is also small, a decrease in the mean score was obtained when comparing the total score in the post-test with that in the pre-test. As a lower total score on the incorrect reasoning scale indicates fewer misconceptions, the results found indicate that teachers in training achieved an improvement, on average, in their erroneous reasoning about statistics and probability.

Table 6

Descriptives and size of the effect for the total scores of the correct and wrong statistical reasoning.

Reasoning	**Test**	$\bar{X}$ (**SD**)	**g**	**CI 95%**
Correct	Pre-test	9.22 (2.60)	-0.305	(-6.90,0.08)
	Post-test	10.18 (2.83)		
Wrong	Pre-test	7.56 (2.59)	0.226	(-0.22,0.68)
	Post-test	6.88 (2.26)		

Note. $\bar{X}$ denotes the arithmetic mean, SD the standard deviation and g the effect size estimated by the Hedges' statistic.

Then, the total SRA scores for the two types of reasoning were compared using the non-parametric Wilcoxon test for matched data. No statistical differences were found between the total correct reasoning scores between the pre-test and the post-test and, likewise, for the erroneous reasoning (Table 7).

Table 7

Wilcoxon non-parametric test for SRA instrument total scores (authors' elaboration)

Reasoning	Test	Mdn	V (p)	CI 95%
Correct	Pre-test	9	29 (0.14)	(-2.5,0.5)
	Post-test	10		
Wrong	Pre-test	8	98.5 (0.23)	(-0.99,1.99)
	Post-test	7		

Note. Mdn denotes the median, and V is the statistic test of the Wilcoxon test.

Furthermore, when comparing the total score for correct statistical reasoning in the post-test with age, no significant correlation was found at 5% (p=0.096), and the same happened with the total score for erroneous statistical reasoning and age (p=0.82). This shows that the score obtained in both types of statistical reasoning is independent of the age of the participants.

In order to perform an analysis of the reliability of the SRA instrument, correlations were calculated between the total correct reasoning scores in the pre-test and post-test and, likewise, for the total erroneous reasoning scores. Spearman's correlation coefficient was obtained, $r_s = 0.609$, CI 95% [0.181,0.843] for correct statistical reasoning, and $r_s = 0.388$, CI 95% [-0.114.0.732] for erroneous reasoning. These results show a moderate correlation between the before and after scores for the correct statistical reasoning scale, but a low correlation between the erroneous reasoning scale scores. These results indicate moderate stability in the total scores of the correct statistical reasoning. However, stability over time is low for the total scores on the erroneous statistical reasoning scale.

CONCLUSIONS

The results of this research report an improvement in the scores regarding attitude towards statistics and statistical reasoning of teachers in training of a course of Mathematical Thinking. Several important results derive from this study.

First, we identified an improvement in students' attitude towards statistics after completing the random thinking module, with an average score above 75 points, which could be due to the redesigning of the module and the didactic strategies implemented in the classes. Therefore, the teachers in training judged the usefulness of statistics positively and showed a good perception and appreciation of their role in the sociocultural field. These results are similar to the results of Estrada et al. (2004), who found that pre-service and in-service teachers usually showed a positive attitude towards statistics because the participants had an average overall score above the position of indifference (75 points).

Secondly, there is a statistically significant positive correlation between the total score on the scale of attitudes towards statistics and the age of the teachers in training. Although the correlation is moderate, these results differ from those of León-Montero et al. (2018), where no significant correlation was found. This difference possibly lies in the fact that older students may be more experienced in statistics than younger ones and, as Estrada et al. (2004) pose it: attitude is a construct that is affected by personal and school variables. For example, Ordóñez-Camacho et al. (2019) found that, in general, older participants have a higher level of anxiety towards statistics and show negative emotions about it.

Thirdly, the analysis of the questions of the instrument of attitude towards statistics showed that in the pre-test and the post-test, the teachers in training had assessed the usefulness of statistics positively since they agree that it helps to understand the world. These results are similar to those found in the study carried out by León-Montero et al. (2018), who applied the SATS to a sample of 105 teachers in training with sociodemographic variables similar to those of this study; the variables sex and age presented a similar distribution in both studies. These researchers also found that teachers in training usually had a good perception and appreciation of the role of statistics in the sociocultural field since they considered it useful both for their studies and future teaching practice. This positive assessment can be interpreted as expressions of thought, conceptions, and beliefs that students manifest about statistics and their usefulness in the world; and it reflects the recognition that students show towards this discipline, even if it is not part of their speciality in the professional program.

Regarding statistical reasoning, although no significant differences were found in the total scores for correct reasoning or erroneous reasoning, there was an increase in the total scores. This contrasts with what was reported by Gundlach et al. (2015), who also applied a pre-test-post-test type design with the SRA instrument. Unlike the present study, these authors found significant differences between the pre-test and the post-test. However, we should make clear that: i) these authors only used eight of the 20 questions the instrument has; ii) they excluded the SRA's questions related to probability and association, leaving only questions about descriptive measures; and iii) the time interval between the pre-test and the post-test was much wider, since the SRA was applied at the beginning and at the end of a 16-week academic semester at an American university.

Finally, probability was the only topic in which the average of correct answers decreased and that of erroneous (wrong) answers increased, and where the teachers in training presented difficulties mainly in terms of reasoning about probability. This may be because probability is an area in which students tend to find more difficult (Konold et al., 1993; Shaughnessy, 2007). Thus, for example, investigations carried out with teachers in training for the elementary education level have shown that only 18.5% of the participants could conceive the representativeness of a sample in probabilistic terms (Groth and Bergner, 2005). Moreover, studies have shown that pre-service and in-service teachers have insufficient knowledge of probability (Martins et al., 2012; Ruz et al., 2020, 2021).

Limitations and a general conclusion

The results suggest evaluating the design of the random thinking module in terms of the contents that it should address and the time that is dedicated to this module. The results also suggest re-evaluating the design with respect to the teaching of stochastics to design strategies for its teaching and learning in the context of prospective teachers of a university degree in primary education. A limitation of this study is the short time devoted to the random thinking module and, therefore, to the consideration of very few stochastic topics. Another limitation is the reliability of the SATS in the pre-test and post-test and the SRA, which shows a moderate degree of stability over time in attitude toward statistics and statistical reasoning scores.

In summary, the conclusion is that prospective teachers assess the usefulness of statistics positively and demonstrate a good perception and appreciation of their role in the sociocultural field. However, they show erroneous reasoning mainly about probability, from where it is important to develop strategies in the classroom that brief on activities that promote the statistical reasoning of pre-service and in-service teachers, especially, reasoning on probability.

Implications

The social changes of recent decades require better-informed citizens who can make decisions in situations where uncertainty is present (Alsina, 2017). Therefore, it is important to systematically work on statistical reasoning and attitudes towards stochastics in primary education. To promote statistical reasoning, we need more studies that implement strategies suggested in the literature, such as the critical reading of articles, working with real data in context, and working on projects in which students are involved with the different stages of the research cycle (Carver et al., 2016; Garfield et al., 2008), besides using validated and reliable instruments that allow articulating research and teaching in the classroom (Zieffler et al., 2008). Furthermore, although the investigation of students' attitudes towards statistics is abundant, as Estrada et al. (2011) and Groth and Meletiou-Mavrotheris (2018) state, there is scarce literature on the attitude towards statistics of pre-service and/or in-service teachers. Therefore, it is important to continue with research that accounts for the attitude towards statistics of teachers in training, as these investigations will influence their knowledge of statistics, their teaching practices, and their students' attitudes.

ACKNOWLEDGEMENTS

This research is the result of the project *Desarrollo del Razonamiento Estadístico en Profesores en Formación de la Licenciatura en Pedagogía Infantil* de la convocatoria interna *Innovaciones Didácticas IV* [The development of statistical reasoning in teachers in training of the university degree in child pedagogy of the internal call didactic innovations IV] of 2019. It is financed with resources from the Centro de Investigaciones Educativas y Pedagógicas - CIEP of the University of Antioquia, Medellín, Colombia. We also thank the research groups MATHEMA-FIEM, Demografía y Salud and Aplicaciones Estadísticas y Salud Pública for supporting this work.

CONFLICTS OF INTEREST

The authors declare having no conflict of interest.

DECLARATIONS OF CONTRIBUTIONS BY AUTHORS

JGB, DGG, JAV and MMP carried out the conceptualisation and research. JGB and DGG wrote the draft, developed the methodology, and worked on the data review. JGB, DGG and JAV performed the analyses and writing. All authors participated in the discussion of the results, and reviewed and approved the published version of the manuscript.

DATA AVAILABILITY DECLARATION

The data supporting the results of this study will be made available by the corresponding author, [JAAP], upon reasonable request.

REFERENCES

Alsina, Á. (2017). Contextos y propuestas para la enseñanza de la estadística y la probabilidad en Educación Infantil : un itinerario didáctico. *Épsilon - Revista de Educación Matemática*, *34*(95), 25–48.

Andrade, L., Fernández, F., & Álvarez, I. (2017). Panorama de la investigación en Educación Estadística desde tesis doctorales 2000-2014. *TED: Tecné, Episteme y Didaxis*, *14*(41), 87–107. https://doi.org/https://doi.org/10.17227/01203916.6039

Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Perez, L., & Spangler, D. A. (2020). *Pre-K–12 Guidelines for Assessment and Instruction in Statistics Education II (GAISE II): A Framework for Statistics and Data Science Education*. American Statistical Association. https://www.amstat.org/asa/files/pdfs/GAISE/GAISEIIPreK-12_Full.pdf

Batanero, C. (2019). Treinta años de investigación en educación estocástica: Reflexiones y desafíos [Congreso]. *Actas Del Tercer Congreso Internacional Virtual de Educación Estocástica*, 15.

https://www.ugr.es/~fqm126/civeest/ponencias/batanero_ing.pdf

Ben-Zvi, D. (2018). Three paradigms to develop students' statistical reasoning [Conference]. *The International Conference on Teaching Statistics*. http://iase-web.org/icots/10/proceedings/pdfs/ICOTS10_2E1.pdf

Ben-Zvi, D., Makar, K., & Garfield, J. (Eds.). (2018). *International handbook of research in statistics education*. Springer. https://doi.org/10.1007/978-3-319-66195-7

Carver, R., Everson, M., Gabrosek, J., Horton, N., Lock, R., Mocko, M., Rossman, A., Holmes Rowell, G., Velleman, P., Witmer, J., & Wood, B. (2016). *Guidelines for Assessment and Instruction in Statistics Education: College Report*. https://www.amstat.org/docs/default-source/amstat-documents/gaisecollege_full.pdf

Creswell, J. (2014). *Research Design: Qualitative, quantitative and mixed methods approaches* (4°). SAGE. https://doi.org/10.1192/bjp.112.483.211-a

delMas, R. (2004). A Comparison of Mathematical and Statistical Reasoning. In D. Ben-Zvi & J. Garfield (Eds.). *The Challenge of Developing Statistical Literacy, Reasoning and Thinking* (pp. 79–95). Springer. https://doi.org/10.1007/1-4020-2278-6_4

Dempster, M., & Mccorry, N. K. (2009). The role of previous experience and attitudes toward statistics in statistics assessment outcomes among undergraduate psychology students. *Journal of Statistics Education*, *17*(2). https://doi.org/10.1080/10691898.2009.11889515

Estrada, A. (2002). *Analisis de las actitudes y conocimiento estadísticos elementales en la formación del profesorado [*published doctoral dissertation, Universitat autònoma de Barcelona]. https://ddd.uab.cat/record/38525

Estrada, A., Batanero, C., & Fortuny, J. M. (2004). Un estudio comparado de las actitudes hacia la estadística en profesores en formación y en ejercicio. *Enseñanza de Las Ciencias: Revista de Investigación y Experiencias Didácticas*, *22*(2), 263–273. https://doi.org/10.5565/rev/ensciencias.3888

Estrada, A., Batanero, C., & Lancaster, S. (2011). Teachers' attitudes towards statistics. In C. Batanero, G. Burrill, & C. Reading. (Eds.). *Teaching Statistics in School MathematicsChallenges for Teaching and Teacher*

Education A Joint ICMIIASE Study (pp. 163–171). Springer. https://doi.org/10.1007/978-94-007-1131-0_18

Estrada, A., Bazán, J., & Aparicio, A. (2013). Evaluación de las propiedades psicométricas de una escala de actitudes hacia la estadística en profesores. *Avances de Investigación En Educación Matemática*, *3*, 5–23. https://doi.org/10.35763/aiem.v0i3.61

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report*. https://www.amstat.org/docs/default-source/amstat-documents/gaiseprek-12_full.pdf

Garfield, J. (1991). Evaluating student's understanding of statistics: Development of the satistical Reasoning assessment. *Proceedings of the Thirteenth Annual Meeting North American Chapter of the International Group for the Psychology of Mathematics Education*, *2*, 1–7. https://files.eric.ed.gov/fulltext/ED352274.pdf

Garfield, J. (2002). The challenge of developing statistical reasoning. *Journal of Statistics Education*, *10*(3). https://doi.org/10.1007/1-4020-2278-6

Garfield, J. (2003). Assessing Statistical Reasoning. *Statistics Education Research Journal*, *2*(1), 22–38.

Garfield, J., & Ben-Zvi, D. (2004). *The challenge of developing statistical literacy, reasoning and thinking* (D. Ben-Zvi & J. Garfield. (Eds.). Springer. https://doi.org/10.1007/1-4020-2278-6

Garfield, J., Ben-Zvi, D., Chance, B., Roseth, C., & Zieffler, A. (2008). *Developing students' statistical reasoning: Connecting research and teaching practice*. Springer. https://doi.org/10.1007/978-1-4020-8383-9

Gillham, B. (2000). *Case study: research methods*. Continuum.

González-Gómez, D. (2014). *Constitución de la identidad del profesor que enseña estadística* [published doctoral dissertation, Universidad de Antioquia]. http://funes.uniandes.edu.co/12045/1/Gonzalez2014Constitución.pdf

González-Gómez, D., & Zapata-Cardona, L. (2018). Professional development programs for statistics teachers: the role of reflection [Conference]. *International Conference on Teaching Statistics*.

https://iase-web.org/icots/10/proceedings/pdfs/ICOTS10_4A2.pdf?1531528851

Groth, R., & Bergner, J. (2005). Pre-service elementary school teachers' metaphors for the concept of statistical sample. *Statistics Education Research Journal*, *4*(2), 27–42.

Groth, R., & Meletiou-Mavrotheris, M. (2018). Research on Statistics Teachers' Cognitive and Affective Characteristics. In D. Ben-Zvi, K. Makar, & J. Garfield. (Eds.). *International Handbook of Research in Statistics Education* (pp. 327–355). Springer. https://doi.org/10.1007/978-3-319-66195-7_10

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), 1–33. https://doi.org/10.1080/10691898.2015.11889723

Henriques, A., Oliveira, H., & Baptista, M. (2019). Prospective teachers' interdisciplinary learning scenario to promote students' statistical reasoning [Congress]. *Eleventh Congress of the European Society for Research in Mathematics Education*. https://hal.archives-ouvertes.fr/hal-02429057/file/TWG05_15.pdf

Karatoprak, R., Karagöz Akar, G., & Börkan, B. (2015). Prospective Elementary and Secondary School Mathematics Teachers' Statistical Reasoning. *International Electronic Journal of Elementary Education*, *7*(2), 107–124.

Konold, C., Pollatsek, A., Well, A., Lohmeier, J., & Lipson, A. (1993). Inconsistencies in Students' Reasoning about Probability. *Journal for Research in Mathematics Education*, *24*, 392–414. https://doi.org/10.2307/749150

León-Montero, C., Casas, J. C., Madrid, M. J. M. J., Jiménez,-Fanjul, N., & Maz-Machado, A. (2018). Actitudes hacia la estadística en futuros maestros de educación infantil. *Yupana*, *10*(16), 8–15. https://doi.org/10.14409/yu.v0i10.7693

Liu, H.-J., & Garfield, J. (2002). Sex Differences in Statistical Reasoning. *Bulletin of Educational Psychology*, *34* (1), 123–136.

Lopes, C. E. (2008). O ensino da estatística e da probabilidade na educação básica e a formação dos professores. *Cadernos CEDES*, *28*(74), 57–73. https://doi.org/10.1590/s0101-32622008000100005

Lovett, M. (2001). A collaborative convergence on studying reasoning processes: A case study in statistics. In D. Klahr & S. Carver. (Eds.). *Cognition and Instruction: 25 Years of Progress* (pp. 347–384). https://learnlab.org/research/wiki/images/5/52/Lovett01CandI.pdf

Márquez, J. C. (2004). Una revisión de las evidencias de fiabilidad y validez de los cuestionarios de actitudes y ansiedad hacia la estadística [A review of the evidence of reliability and validity of the attitudes and anxiety questionnaires towards statistics]. *Statistics Education Research Journal*, *3*(1), 5–28.

Martins, J. A., Nascimento, M. M., & Estrada, A. (2012). Looking back over their shoulders: A qualitative analysis of portuguese teachers' attitudes towards statistics. *Statistics Education Research Journal*, *11*(2), 26–44. https://doi.org/10.52041/serj.v11i2.327

Ministerio de Educación Nacional [MEN]. (1998). *Lineamientos curriculares Matemáticas*. https://www.mineducacion.gov.co/1621/articles-89869_archivo_pdf9.pdf

Moore, D. (1998). Statistics among the Liberal Arts. *Journal of the American Statistical Association*, *93*(444), 1253–1259. https://doi.org/10.1080/01621459.1998.10473786

Ordóñez-Camacho, X., Romero-Martínez, S., & Ruiz, D. M. (2019). Actitudes hacia la Estadística en Alumnos de Educación: Análisis de Perfiles. *Revista de Educacion*, *2019*(385), 165–192. https://doi.org/10.4438/1988-592X-RE-2019-385-421

Pérez, L. L., Aparicio, A. P., Bazán, J. G., & Abdounur, O. J. (2015). Actitudes hacia la estadística de estudiantes universitarios de Colombia. *Educación Matemática*, *27*(3), 111–149.

Pfannkuch, M., & Wild, C. J. (1999). Statistical Thinking in Empirical Enquiry. *International Statistical Review*, *67*(3), 223–265.

Ruz, F., Molina-Portillo, E., & Contreras, J. M. (2021). Exploring knowledge about probability content in prospective mathematics teachers. *Boletín de Estadística e Investigación Operativa (BEIO)*, *37*(2), 130–147.

Ruz, F., Molina-Portillo, E., Vásquez, C., & Contreras, J. M. (2020). Attitudes towards probability and its teaching in prospective mathematics teachers from Chile and Spain. *Acta Scientiae*, *22*(2), 48–66. https://doi.org/10.17648/acta.scientiae.5489

Schau, C., & Emmioğlu, E. (2012). Do introductory statistics courses in the United States improve students' attitudes? *Statistics Education Research Journal*, *11*(2), 86–94.

Shaughnessy, M. (2007). Research on statistics learning and reasoning. In F. Lester. (Ed.). *Second handbook of research on mathematics teaching and learning* (pp. 957–1009). NCTM.

Szklo, M., & Nieto, J. (2003). *Epidemiologia Intermedia: Conceptos y Aplicaciones.* Díaz Santos.

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), 1–56. https://doi.org/10.1080/10691898.2012.11889641

Tversky, A., & Kahneman, D. (1982). Judgment under uncertainty: Heuristics and biases. In D. Kahneman, P. Slovic, & A. Tversky. (Eds.). *Judgment under uncertainty* (pp. 3–20). Cambridge University Press. https://doi.org/10.1017/CBO9780511809477.002

Zieffler, A., Chang, B., Holleque, K., Garfield, J., Dupuis, D., & Alt, S. (2008). What Does Research Suggest About the Teaching and Learning of Introductory Statistics at the College Level? A Review of the Literature. *Journal of Statistics Education*, *16*(2), 26. https://doi.org/10.1080/10691898.2008.11889566

Zieffler, A., Garfield, J., & Fry, E. (2018). What Is Statistics Education? In D. Ben-Zvi, K. Makar, & J. Garfield. (Eds.). *International Handbook of Research in Statistics Education* (pp. 37–70). Springer. https://doi.org/10.1007/978-3-319-66195-7_2