APPROACHING STATISTICS AS A DISTINCT SUBJECT AND ITS IMPACT ON REASONING AND CONCEPTUAL ERRORS

José Carlos Casas del Rosal 1
Oneida Muñoz Ñungo 2
Alexander Maz Machado 3
Astrid Cuida Gómez 4

ABSTRACT

Objective: This study aims to determine the level of statistical reasoning and misconceptions expressed by ninth-grade students in two schools. In one of the schools, statistics is taught as a specific subject, while in the other, it is included within the mathematics curriculum.

Theoretical Framework: Students are taught statistics throughout their primary education, so by secondary school, they are expected to have developed a good level of statistical reasoning. However, classroom experience suggests this may not be the case. Additionally, this topic often receives less attention in favor of other subjects included in the mathematics curriculum.

Method: The methodology adopted for this research includes descriptive and inferential analyses to compare the rate of correct reasoning responses and conceptual errors. Data collection was carried out using the Statistical Reasoning Assessment (SRA) test.

Results and Discussion: The results showed different behaviors between the two groups. This suggests that teaching statistics separately from mathematics can influence the learning of this subject.

Research Implications: The advantages offered by separating statistics from the rest of mathematics suggest that this approach should be considered in the future teaching of the subject.

Originality/Value: This study contributes to the literature by demonstrating that the separation of statistics from mathematics teaching should be considered. The relevance and value of this research are evidenced by the need to improve students' statistical reasoning abilities.

Keywords: Statistics Reasoning, Correct Reasoning, Misconceptions, Statistics Education, Secondary Education.

ABORDAR A ESTATÍSTICA COMO UM TEMA DISTINTO E O SEU IMPACTO NO RACIOCÍNIO E NOS ERROS CONCEPTUAIS

RESUMO

Objetivo: Este estudo tem como objetivo determinar o grau de raciocínio estatístico e conceitos errôneos expressados por alunos do nono ano em duas escolas. Em uma delas, é oferecida uma disciplina específica de estatística e, na outra, a estatística é incluída na disciplina de matemática.

Referencial Teórico: Na maioria dos países, os alunos aprendem estatística ao longo de toda a sua educação primária, de modo que no ensino secundário se supõe que tenham desenvolvido um bom nível de raciocínio...
Approaching Statistics as a Distinct Subject and Its Impact on Reasoning and Conceptual Errors

Método: A metodologia adotada para esta pesquisa compreende análises descritivas e inferenciais que permitem comparar a taxa de respostas de raciocínios corretos e erros conceituais. A coleta de dados foi realizada por meio do teste de Avaliação do Raciocínio Estatístico (SRA).

Resultados e Discussão: Os resultados mostraram um comportamento diferente entre os dois grupos. Isso sugere que trabalhar com estatística separadamente da matemática pode influenciar na aprendizagem desta disciplina.

Implicações da Pesquisa: As vantagens oferecidas pela separação da estatística do restante da matemática sugerem que isso deve ser considerado no futuro ensino da disciplina.

Originalidade/Valor: Este estudo contribui para a literatura ao evidenciar que deve ser considerada a separação da estatística no ensino da matemática. A relevância e o valor desta pesquisa se evidenciam na necessidade de melhorar a capacidade de raciocínio estatístico dos alunos.


EL ESTUDIO DE LA ESTADÍSTICA COMO UNA DISCIPLINA INDEPENDIENTE Y SU IMPACTO EN EL RAZONAMIENTO Y LOS ERRORES CONCEPTUALES

RESUMEN

Objetivo: Este estudio tiene como objetivo determinar el grado de razonamiento estadístico y conceptos erróneos expresados por estudiantes de noveno grado en dos escuelas. En uno de ellos se imparte una materia específica de estadística y en el otro colegio la estadística se incluye en la materia de matemáticas.

Marco Teórico: En la mayoría de los países, a los estudiantes se les enseña estadística durante toda su educación primaria, por lo que en la educación secundaria se supone que han desarrollado un buen nivel de razonamiento estadístico. Sin embargo, la experiencia en el aula sugiere que este puede no ser el caso. Además, a este tema se le suele dedicar menos tiempo en favor de otros temas incluidos en el plan de estudios de matemáticas.

Método: La metodología adoptada para esta investigación comprende análisis descriptivos e inferenciales que permiten comparar la tasa de respuesta de razonamientos correctos y errores conceptuales. La recolección de datos se realizó mediante la prueba de Evaluación del Razonamiento Estadístico (SRA).

Resultados y Discusión: Los resultados mostraron un comportamiento diferente entre los dos grupos. Esto sugiere que trabajar con estadística por separado de las matemáticas puede influir en el aprendizaje de esta materia.

Implicaciones de la investigación: Las ventajas que ofrece la separación de la estadística del resto de las matemáticas sugieren que debe considerarse en la docencia futura de esta materia.

Originalidad/Valor: Este estudio contribuye a la literatura al evidenciar que debe considerarse separar la estadística en la docencia de las matemáticas. La relevancia y valor de esta investigación se evidencian en la necesidad de mejorar la capacidad de razonamiento estadístico de los estudiantes.

Palabras clave: Razonamiento Estadístico, Razonamiento Correcto, Error Conceptual, Educación Estadística, Educación Secundaria.

RGSA adota a Licença de Atribuição CC BY do Creative Commons (https://creativecommons.org/licenses/by/4.0/).
1 INTRODUCTION

In the nineties of the last century, it was typical for studies related to statistics to be found only in secondary education. Additionally, they were focused on incorporating statistical skills. Consequently, most students began their university studies with minimal formal experience in statistical reasoning, probability laws, and probabilistic reasoning (Derry, Levin, Osana, & Jones, 1998).

Today, there is a growing awareness of the need and usefulness of statistics and probability in many aspects of everyday life. Many situations require statistical reasoning by people with little training. They must be able to apply it when facing the information presented in the media or in the development of their professional work (Batanero Bernabéu, 2019). This is reflected in the inclusion of statistics and probability in school curricula in several countries, which, unfortunately, does not guarantee correct teaching and learning.

A factor that can also negatively influence the acquisition of statistical reasoning is that its inclusion is carried out within the mathematics subject in most cases (Weiland, 2019; Zieffler, Garfield, & Fry, 2018). This causes, in many cases, less time to be invested in this block of content compared to other blocks of the mathematics curriculum. It is left until the end of the academic year or omitted due to lack of time (Vásquez Ortiz & Alsina, 2023).

Statistics education is an interdisciplinary field that focuses attention on the teaching and learning processes of statistics (Zieffler et al., 2018). Not surprisingly, statistics and probability have also become an essential part of the mathematics curriculum during primary education. For example, it is established that the main reason for including the mathematical study of random phenomena in primary and secondary education is the strong presence of random and probability situations in our environment (Batanero & Godino, 2001). It is not surprising that, since the publication of the principles and standards for school mathematics (Principles, 2000) and its recommendations on starting the teaching of statistics from an early age (A. Bargagliotti, 2020; Franklin et al., 2007), most countries have begun to include this subject at primary levels.

In the case of the school curriculum in Colombia, where the present study was carried out, the educational system is divided into two cycles. The first is basic education, which consists of five years of primary education and four years of secondary education. That is, from first to ninth grade. The second is secondary education, which lasts two years: tenth and eleventh grade. The Congress of the Republic of Colombia, in 1994, through Law 115 General of Education ("Law 115 of February 8, 1994 by which the general law of education is issued,"
Approaching Statistics as a Distinct Subject and Its Impact on Reasoning and Conceptual Errors

(1994), established in the educational curriculum that Students must learn statistics.

Furthermore, this should be taught from first to eleventh grade, developing competencies related to descriptive and inferential statistics and the interpretation of tables and graphs. This is established in Art. 25 of the Law. That same year, the Ministry of National Education (MEN) issued Decree 1860, which partially regulates Law 115 of 1994, in general pedagogical and organizational aspects ("Decree 1860 of August 3, 1994, which partially regulates Law 115 of 1994, in pedagogical and organizational aspects," 1994). Later, in 2006, the Basic Standards of Competence in Mathematics were published, where it was established that to be competent in mathematics, there must be a well-organized curriculum focused on the development of five types of thinking, among which are random thinking or probabilistic, also called probabilistic or stochastic (Nacional, 2006).

As regulated in the aforementioned legislation, the Colombian curriculum addresses, up to ninth grade, competencies in sampling, representativeness, data collection, univariate and bivariate descriptive analysis, inference, randomness, calculation of probabilities in simple and compound experiments, and the independence of events (National, 2006).

This study aims to analyze whether teaching statistics separately from mathematics can influence the improvement of statistical reasoning and the reduction of misconceptions in ninth grade students in Colombia.

2 THEORETICAL FRAMEWORK

2.1 STATISTICAL REASONING

However, at an educational and cognitive level, statistical reasoning goes beyond knowledge of statistical concepts and therefore differs from statistical literacy. This is because statistical reasoning can be defined as:

the way people reason with statistical ideas and understand statistical information. Statistical reasoning may involve connecting one concept to another (for example, center and dispersion) or may combine ideas about data and chance. Reasoning means understanding and being able to explain statistical processes and being able to fully interpret statistical results (J. Garfield & Gal, 1999).

Although no consensus has been reached on statistical reasoning (Sabbag, Garfield, & Zieffler, 2018), the above definition has been chosen because it is widely used in the scientific
literature and the instrument used in this research is based on it. Statistical literacy focuses primarily on the social construction of statistics, the understanding of texts, and the meaning and implications of statistical information in the context to which it belongs (Watson, 1997).

Unlike statistical literacy, which is insufficient, statistical reasoning is of great importance in decision-making, since it allows decisions to be made more quickly, intuitively and reliably (Kahneman, Lovallo, & Sibony, 2011).

Statistical reasoning combines ideas about data and chance. Furthermore, these two aspects allow interpreting inferences and statistical results. This reasoning requires understanding essential concepts such as distribution, center, dispersion, association, uncertainty, randomness and sampling (J. Garfield, 2002). Furthermore, statistical reasoning can accompany deductive and inductive reasoning in research situations in which hypotheses are formulated and tested about experiments or surveys designed to answer specific questions (Lavigne & Lajoie, 2007).

Likewise, in Garfield's work, five levels of statistical reasoning are identified: idiosyncratic, verbal, transitional, procedural, and integrated process (J. Garfield, 2002). In the first, students vaguely know the concepts and use them without understanding them, so they often do so incorrectly. In the second they know the concepts but cannot apply them to real situations. In the third, students can identify the concepts involved, but do not know how to relate them correctly. At the fourth level, they can perform the process correctly, but they do not understand how the response is generated. In the latter, students have a complete mastery that allows them to understand the problem in its entirety.

Identifying statistical reasoning as it occurs in the social context of the classroom requires an extension of studies on the development of students' statistical reasoning. It requires the development of reasoning about critical statistical concepts (e.g., modeling and sampling), including their psychological, social, pedagogical and epistemological dimensions (Ben-Zvi, 2008).

As mentioned in their work by Lavigne and Lajoie (Lavigne & Lajoie, 2007), initially, works on statistical reasoning, such as that of Kahneman and Tversky (1982), focused mainly on the first part of the definition (reasoning with ideas or statistical rules) by examining how rules develop in children and the extent to which adults use these rules to make decisions in well-defined hypothetical problems.

Subsequently, research on this topic expanded the focus to the second part of the definition (making inferences from data), studying the development of children's ability to reason with data and how statistical rules are applied in different learning situations. learning
Approaching Statistics as a Distinct Subject and Its Impact on Reasoning and Conceptual Errors

(Lavigne & Lajoie, 2007). Some research has also noted that personal knowledge and beliefs play a critical role in children’s statistical reasoning (Jacobs, 1999; Schwartz & Goldman, 1996).

Through various investigations, six types of content-specific reasoning have been identified: reasoning about samples, reasoning about data, reasoning about data representations, reasoning about statistical measures, reasoning about uncertainty, and reasoning about association (J. Garfield & Gal, 1999).

As Lavigne and Lajoie point out, these types of statistical reasoning coincide with certain activities proposed in the NCTM (2000) standards (Lavigne & Lajoie, 2007). In these proposals, students are asked to pose questions that involve collecting, organizing, analyzing and representing data to make inferences and predictions.

2.2 STATISTICAL REASONING

Garfield's Statistical Reasoning Assessment (SRA) instrument is an instrument widely used in the literature to evaluate statistical reasoning. It is a multiple choice test composed of 20 items in which one or more answers can be considered conclusions of correct reasoning and others that are distractors that sometimes evaluate some type of misconception. The selection of responses by the respondents allows the evaluation of eight types of correct reasoning and eight misconceptions. These can be consulted in Table 1. According to the Colombian mathematics curriculum, it is assumed that a student who has completed ninth grade has the necessary skills to respond to the contents involved in the SRA.

<table>
<thead>
<tr>
<th>Correct Reasoning Skills</th>
<th>Misconceptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR1. Correctly interpret probabilities</td>
<td>MC1. Misconceptions related to averages</td>
</tr>
<tr>
<td>CR2. Understand how to select an appropriate average</td>
<td>MC2. Misconception of result orientation</td>
</tr>
<tr>
<td>CR3. Correctly calculate the probability</td>
<td>MC3. Good samples should represent a high percentage of the population</td>
</tr>
<tr>
<td>CR4. Understand independence</td>
<td>MC4. Law of small numbers</td>
</tr>
<tr>
<td>CR5. Understand sampling variability</td>
<td>MC5. Misconception of representativeness</td>
</tr>
<tr>
<td>CR6. Distinguish between correlation and causation</td>
<td>MC6. Correlation implies causation</td>
</tr>
<tr>
<td>CR7. Correctly interpret contingency tables</td>
<td>MC7. Equiprobability bias</td>
</tr>
<tr>
<td>CR8. Understand the importance of large samples</td>
<td>MC8. Groups can only be compared if they are the same size</td>
</tr>
</tbody>
</table>
To evaluate statistical reasoning, Saidi and Siew recently conducted a study with 320 high school science students in Malaysia (Saidi & Siew, 2022). Attitude and anxiety towards statistics were analyzed, as well as the level of literacy. They found positive results in the degree of literacy at a general level and in the description, organization and representation of data. However, statistical reasoning was found to be fundamentally deficient in questions about giving inferences, reasons or justifications for their answers.

Another study compared the level of statistical reasoning of 1360 students in five secondary schools in China (Wang, Wang, & Chen, 2009) with the results obtained by Tempelaar's work in 2004 for Dutch students (Tempelaar, 2004). For this purpose, they applied the SRA test. Chinese students scored lower than Dutch students on all eight correct conceptions. For the misconceptions, the performance of the Chinese students was not as good as that of the Dutch, but the general trends were the same, obtaining good results in both cases.

At the university level, the statistical reasoning of Communication Sciences students was studied (Jauhari, Ariany, Fardillah, & Ayu, 2021). To do this, they carried out a qualitative study in which 100 students were classified within one of the five levels of reasoning proposed by (J. Garfield, 2002). The extreme levels were assigned to only 8 students (four each). Instead, 48 were assigned to level 3, corresponding to transitional reasoning. At this level, students can identify problems using appropriate symbols or terms, determine research hypotheses, and know the concepts used in problem solving, but do not understand the use of the concepts.

On the other hand, the statistical misconceptions of future teachers were also studied using the SRA (Gorham Blanco & Chamberlin, 2019). A total of 134 students participated in the study. They found that the most common misconceptions were comparing groups of the same size, equiprobability bias, and correlation implies causation. However, they did not find, in general, significant differences by gender.

Work on the statistical reasoning of high school students in Colombia is very scarce. In one of them, a study on statistical reasoning was carried out with 40 tenth and eleventh grade students from a school in a rural area of the department of Antioquia who designed statistics applications for mobile devices. To do this, they used a design with an experimental group (those who had worked on the implementation of the application) and a control group, and carried out pre- and post-tests. The SRA test was applied to evaluate statistical reasoning. They found significant differences in favor of the experimental group in the categories of conceptions of statistical reasoning and the types of misconceptions. In addition, the didactic sequence improved concepts such as the correct interpretation of probabilities and the law of small numbers.
To analyze the influence of influential factors on statistical reasoning, two key factors that can significantly affect statistical reasoning skills are statistical training (Fong, Krantz, & Nisbett, 1986) and gender (Martin, Hughes, & Fugelsang, 2017). Regarding gender, several investigations have found significant differences in the skills of men and women. In these cases, women tend to obtain lower levels than men (J. B. Garfield, 2003; Martin et al., 2017; Yusuf & Sukestiyarno, 2022).

2.3 PROBLEM STATEMENT

A study on statistical reasoning in ninth grade students in Colombia was carried out because there are no previous studies that analyze the need to study statistics separately from the rest of mathematics. Statistics has specific characteristics that differentiate it from other branches of mathematics, in addition to the importance it has in everyday life. The chosen degree is justified because primary education ends in this course, the students have received statistical training in all courses, and the skills acquired up to that point, according to the Colombian curriculum, are sufficient to adequately respond to the test used in this research. It is expected that at the end of their Basic Secondary Education, they will have developed an adequate level of this type of reasoning.

Furthermore, in most cases, the treatment of statistics is carried out within the subject of mathematics, which can become an obstacle to not devoting enough time to statistics in favor of other mathematical topics. Given the autonomy of Colombian educational centers, some centers have chosen to separate statistics from the subject of mathematics. This decision is motivated by its differentiated characteristics with respect to the rest of the blocks and the little time that is usually dedicated to the topic of statistics when it is taught at the end of the topic (Naya, Ríos and Zapata, 2012).

This research aims to analyze the different statistical reasoning skills and the errors that can be attributed to specific misconceptions made by ninth grade students from two Colombian schools with similar characteristics. That is, students from both schools have the same textbooks, the same time distribution, the same content and objectives in statistics, similar resources and the same methodology based on classes followed by problem solving in class. In this way, an adequate schedule and study load is ensured that can significantly influence students' statistical instruction and, therefore, statistical reasoning. Furthermore, given that previous results obtained by other studies indicate a possible effect of gender in this type of reasoning, it is necessary to control this factor to prevent possible differences between men and
women from masking the natural effect that the separation of statistics may have, on the topic of mathematics.

For all these reasons, the question posed in this research is the following: Does the separation of statistics from the rest of mathematics positively influence the statistical reasoning of boys and girls in their last year of primary education? Furthermore, given that gender is a ubiquitous factor in mathematics assessment studies with a high influence in many cases, can gender have a significant influence in this case?

3 METHODOLOGY

3.1 PARTICIPANTS

The test was applied to a sample of ninth grade students from two schools in Ibagué, Colombia (n=183), at the end of the academic year with the same teacher. The gender distribution was as follows: 94 men and 89 women. The ages ranged between 14 and 17 years. School A (men = 45, women = 50) includes the teaching of statistics within the mathematics subject, while School B (men = 49, women = 39) teaches statistics as a specific subject with its timetable. The selection of students was intentional, and participation was anonymous and voluntary. The test was administered during the first week of November 2020 by the authors of the study, at a time when the academic year is about to end, when the confinement measures imposed by COVID-19 were relaxed. The students had one hour to answer the test.

The objective in selecting the study sample was to seek maximum homogeneity in the characteristics of students, teachers and schools so that the main discriminating factor would be the separation of statistics from the rest of mathematics. Therefore, no significant differences were observed in the mathematics attitude and performance of students from both schools at the beginning of the course. The socioeconomic level of the students is lower middle class in both schools, the two centers are similar, and the teacher followed similar methodologies in both classrooms. Furthermore, once the teachers related to the teaching of statistics in the previous courses of the two schools were consulted, it can be stated that they had received training in statistics in all the previous grades as indicated by the official curriculum and, in all cases, the statistical contents established by law were addressed.

According to the Colombian curriculum, acquiring mathematical skills must focus on five types of mathematical thinking: numerical, spatial, metric, variational and random, between which time must be shared equally. The latter is the name used for statistics. According
to this distribution, statistics should occupy 20% of the teaching load. The weekly dedication
to mathematics is five hours, equivalent to a study load of one hour per week for each type of
thinking. Thus, in School B, where statistics was taught independently, students took one hour
of statistics class per week during the last year. Therefore, this time was like the expected time
that would have been dedicated to this topic within the mathematics subject at School A. That
is, at School A there are 5-hour mathematics classes and at School B there are 5-hour
mathematics classes. 4-hour mathematics classes and 1-hour statistics classes per week. The
methodology of an article describes the procedures used to conduct the research, including the
type of study, sample selection, data collection and analysis methods, ethical considerations,
and limitations of the study. Its detailed and transparent description is essential to guarantee the
replicability and reliability of the results, in addition to providing a solid basis for the
interpretation and generalization of the findings.

3.2 INSTRUMENT

A previously mentioned and repeatedly validated instrument for different samples, the
Statistical Reasoning Assessment (SRA), was used. It is a questionnaire of basic statistical
knowledge (J. B. Garfield, 2003). It was translated into Spanish by a team of experts consisting
of a bilingual statistician with Spanish as his native language and a bilingual person with
English as his native language. It consists of twenty questions/problems and is frequently used
in various research on statistical knowledge and reasoning internationally (Martin et al., 2017;
Muñoz-Ñungo, Maz-Machado, & Pedrosa-Jesús, 2020; Wang et al., 2009).

As mentioned above, the scale assesses 16 characteristics of statistical reasoning
grouped into two categories: the correct reasoning scale (8 characteristics) and the
misconceptions scale (8 characteristics). The numerical score is obtained from the selection of
answers given by the students in each of the 20 questions of the questionnaire. Each
characteristic is analyzed through responses to one or more questions in the questionnaire. If
students select an item option that represents a particular characteristic, one is assigned to that
variable. Thus, the value for each characteristic is the average response rate for each set of items
included in that characteristic.

The scores of all characteristics are expressed between 0 and 1. Therefore, when
considering the total score of each dimension, the difference in the number of items involved
in each characteristic does not alter this value. In this way, all features have equal weight in
estimating the statistical reasoning value of misconceptions.
It should be noted that a high score on the statistical reasoning scale indicates better performance, while a low score on the misconceptions scale is indicative of a positive evaluation.

3.3 STATISTICAL ANALYSIS

First, inferential techniques of correct reasoning and misconceptions were compared for the two schools. This comparison was then made between men and women. Finally, multiple comparison tests were carried out to analyze the influence of the two factors simultaneously.

The first analysis compared the average scores of the analyzed characteristics of correct reasoning and misconceptions of students in each school. Hypothesis tests were also carried out to infer these comparisons according to the normality results previously carried out. Thus, the t test was applied, and the Cohen statistic (effect size) (Cohen, 1988) was calculated for the characteristics. The Mann-Whitney test was applied, and the effect size was calculated using the Tomczak and Tomczak statistic (small is considered 0.2, medium is 0.5 and large is 0.8) (Tomczak & Tomczak, 2014). These results are shown below, first for both schools and then to compare the results by gender. Finally, both factors were analyzed simultaneously.

4 RESULTS AND DISCUSSIONS

4.1 STATISTICS AS A SEPARATE SUBJECT: CORRECT REASONING SKILLS AND MISCONCEPTIONS

As seen in Table 2, the scores indicate low levels of statistical reasoning in the students of both schools (the general variables of correct reasoning skills and misconceptions are between 0 and 8 points, while the individual variables are between 0 and 1). They also show the types of reasoning that present the greatest difficulties for students. For example, there is a low proportion of correct answers to CR2 questions related to the selection of an appropriate measure (0.20 and 0.25). In addition, difficulty is observed in understanding the variability caused by sampling (CR5), mainly in school B, and difficulty in interpreting the double-entry tables (CR7) in school A.
Table 2

Comparison of scaled scores for SRA for two schools.

<table>
<thead>
<tr>
<th></th>
<th>School A (n=95)</th>
<th>School B (n=88)</th>
<th>p value</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Med</td>
<td>Desv</td>
<td>Med</td>
<td>Desv</td>
</tr>
<tr>
<td>Correct Reasoning Skills</td>
<td>2.07</td>
<td>0.969</td>
<td>2.52</td>
<td>1.083</td>
</tr>
<tr>
<td>Correct interpret probabilities</td>
<td>0.28</td>
<td>0.046</td>
<td>0.32</td>
<td>0.050</td>
</tr>
<tr>
<td>Understand how to select an appropriate average</td>
<td>0.20</td>
<td>0.041</td>
<td>0.25</td>
<td>0.046</td>
</tr>
<tr>
<td>Correctly calculate the probability</td>
<td>0.29</td>
<td>0.047</td>
<td>0.32</td>
<td>0.050</td>
</tr>
<tr>
<td>Understand independence</td>
<td>0.29</td>
<td>0.047</td>
<td>0.43</td>
<td>0.053</td>
</tr>
<tr>
<td>Understand sampling variability</td>
<td>0.28</td>
<td>0.046</td>
<td>0.18</td>
<td>0.041</td>
</tr>
<tr>
<td>Distinguish between correlation and causation</td>
<td>0.29</td>
<td>0.047</td>
<td>0.43</td>
<td>0.053</td>
</tr>
<tr>
<td>Correctly interpret double entry tables</td>
<td>0.17</td>
<td>0.039</td>
<td>0.31</td>
<td>0.049</td>
</tr>
<tr>
<td>Understand the importance of large samples</td>
<td>0.27</td>
<td>0.046</td>
<td>0.27</td>
<td>0.047</td>
</tr>
<tr>
<td>Misconceptions</td>
<td>2.78</td>
<td>0.986</td>
<td>3.30</td>
<td>1.066</td>
</tr>
<tr>
<td>Misconceptions involving averages</td>
<td>0.47</td>
<td>0.051</td>
<td>0.50</td>
<td>0.053</td>
</tr>
<tr>
<td>Orientation to results</td>
<td>0.33</td>
<td>0.048</td>
<td>0.32</td>
<td>0.050</td>
</tr>
<tr>
<td>Good samples should represent a high percentage of the population</td>
<td>0.49</td>
<td>0.051</td>
<td>0.35</td>
<td>0.051</td>
</tr>
<tr>
<td>Law of small numbers</td>
<td>0.20</td>
<td>0.041</td>
<td>0.35</td>
<td>0.051</td>
</tr>
<tr>
<td>Misconception of representativeness</td>
<td>0.35</td>
<td>0.049</td>
<td>0.39</td>
<td>0.052</td>
</tr>
<tr>
<td>Correlation implies causation</td>
<td>0.35</td>
<td>0.049</td>
<td>0.56</td>
<td>0.053</td>
</tr>
<tr>
<td>Equiprobability bias</td>
<td>0.34</td>
<td>0.049</td>
<td>0.44</td>
<td>0.053</td>
</tr>
<tr>
<td>Groups can only be compared if they are the same size</td>
<td>0.25</td>
<td>0.044</td>
<td>0.39</td>
<td>0.052</td>
</tr>
</tbody>
</table>

On the other hand, the most frequent misconceptions in both schools are those related to averages (MC1). The misconception about the size of the sample and its representativeness (MC3) predominates in school A. However, the belief that correlation necessarily implies causation (MC6) is the most common misconception in school B, where statistics as a separate subject from mathematics.

Regarding the inferential analysis, significant differences are observed between the schools globally, both for statistical reasoning (p = 0.004) and for conceptual errors (p = 0.001), with a medium effect size for conceptual errors (0.507 is greater than 0.5) and medium-low for statistical reasoning (0.432). In both cases the scores are higher in the school where statistics is taught separately. A high degree of statistical reasoning is observed in this school, but also in conceptual errors.

When analyzing each characteristic individually, the differences are significant in the item on understanding independence (CR4) for statistical reasoning, with a small effect size (0.206) and misconceptions related to the law of small numbers (MC4), with also small effect size (0.212). Furthermore, the idea that implies correlation-causation (MC6). In the case of misconceptions, the scores are higher in school B, where statistics is taught as a separate subject.
4.2 GÉNERO: HABILIDADES DE RAZONAMIENTO CORRECTO Y CONCEPTOS ERRÓNEOS

Secondly, statistical reasoning and misconceptions were analyzed based on the gender of the students surveyed. The inclusion of this analysis was based on the results obtained by previous studies that concluded that gender could be a determining factor in these constructs (Garfield, 2003; Martin et al., 2017; Yusuf & Sukestiyarno, 2022). The results are shown in Table 3.

**Table 3**

*Comparison of scaled scores for SRA by gender*

<table>
<thead>
<tr>
<th></th>
<th>Male (n=94)</th>
<th>Female (n=89)</th>
<th>p value</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct Reasoning Skills</td>
<td>Med</td>
<td>Desv</td>
<td>Med</td>
<td>Desv</td>
</tr>
<tr>
<td>1. Correctly interpret probabilities</td>
<td>0.33</td>
<td>0.048 0.28</td>
<td>0.48</td>
<td>0.204 0.240</td>
</tr>
<tr>
<td>2. Understand how to select an appropriate average</td>
<td>0.21</td>
<td>0.042 0.24</td>
<td>0.45</td>
<td>0.342 0.123</td>
</tr>
<tr>
<td>3. Calculate the probability correctly</td>
<td>0.29</td>
<td>0.047 0.32</td>
<td>0.49</td>
<td>0.244 0.164</td>
</tr>
<tr>
<td>4. Understand independence</td>
<td>0.33</td>
<td>0.048 0.39</td>
<td>0.52</td>
<td>0.255 0.161</td>
</tr>
<tr>
<td>5. Understand sampling variability</td>
<td>0.23</td>
<td>0.043 0.23</td>
<td>0.45</td>
<td>0.839 0.026</td>
</tr>
<tr>
<td>6. Distinguish between correlation and causation</td>
<td>0.36</td>
<td>0.050 0.36</td>
<td>0.51</td>
<td>0.976 0.004</td>
</tr>
<tr>
<td>7. Interpret double-entry tables correctly</td>
<td>0.21</td>
<td>0.042 0.26</td>
<td>0.46</td>
<td>0.468 0.079</td>
</tr>
<tr>
<td>8. Understand the importance of large samples</td>
<td>0.26</td>
<td>0.045 0.28</td>
<td>0.48</td>
<td>0.451 0.099</td>
</tr>
<tr>
<td>Misconceptions</td>
<td>Med</td>
<td>Desv</td>
<td>Med</td>
<td>Desv</td>
</tr>
<tr>
<td>1. Misconceptions involving averages</td>
<td>0.48</td>
<td>0.052 0.49</td>
<td>0.53</td>
<td>0.533 0.086</td>
</tr>
<tr>
<td>2. Results orientation</td>
<td>0.30</td>
<td>0.047 0.35</td>
<td>0.51</td>
<td>0.110 0.228</td>
</tr>
<tr>
<td>3. Good samples must represent a high percentage of the population</td>
<td>0.41</td>
<td>0.051 0.43</td>
<td>0.52</td>
<td>0.863 0.024</td>
</tr>
<tr>
<td>4. Law of small numbers</td>
<td>0.31</td>
<td>0.048 0.24</td>
<td>0.45</td>
<td>0.090 0.224</td>
</tr>
<tr>
<td>5. Misconception of representativeness</td>
<td>0.41</td>
<td>0.051 0.32</td>
<td>0.49</td>
<td>0.029 0.298</td>
</tr>
<tr>
<td>6. Correlation implies causation</td>
<td>0.45</td>
<td>0.051 0.45</td>
<td>0.53</td>
<td>0.972 0.005</td>
</tr>
<tr>
<td>7. Equiprobability bias</td>
<td>0.35</td>
<td>0.049 0.43</td>
<td>0.52</td>
<td>0.033 0.305</td>
</tr>
<tr>
<td>8. Groups can only be compared if they are the same size</td>
<td>0.34</td>
<td>0.049 0.29</td>
<td>0.48</td>
<td>0.484 0.083</td>
</tr>
</tbody>
</table>

At a general level, no influence of gender is observed either on correct reasoning skills or on misconceptions. Significant differences were observed in the law of small numbers (MC4, d = 0.224, small), prevalence of misconception of representativeness (MC5, d = 0.298, small) and equiprobability bias (MC7, d = 0.305, small). The former shows a higher prevalence among men than the latter. However, in both cases, the effect size is small.

Therefore, we can affirm that in the population studied, gender does not seem to be a determining factor in statistical reasoning. This result, contrary to previous studies, does not identify differences in statistical reasoning skills between boys and girls, which makes it possible to identify the teaching of statistics as a subject separate from mathematics as the factor...
that can influence these skills of both. analyzed. However, to complete the study, it is necessary to perform an analysis that covers both factors to identify if the influence of gender could be masked in the results of the main factor.

4.3 BOTH FACTORS: CORRECT REASONING SKILLS AND MISCONCEPTIONS

For the reasons mentioned above, statistics, as a separate subject, and gender, in statistical reasoning and misconceptions, were simultaneously investigated. For this purpose, the total scores of both constructs were used. To this end, two-way ANOVA models were estimated after verifying the hypotheses of normality and homoscedasticity, which are necessary for their application.

Two-way analysis of variance for total correct reasoning scores by school and gender indicates a significant effect of school (p-value = 0.003). No significant differences are observed regarding gender (p-value = 0.288). This is consistent with what was previously observed for preservice teachers (Gorham & Chamberlin, 2019) and contrary to what was obtained by other studies (e.g., Garfield for undergraduate students; Martin et al. for undergraduate and graduate students; and Yusuf and Sukestyarno for teachers). In training). Furthermore, no interaction was observed between both factors (p value = 0.549). This allows us to conclude the limited influence of students’ gender on the results of the study.

After obtaining these results, the mean correct reasoning scores were graphed by gender and school. As shown in Figure 1, students at School B, where statistics is a separate subject, have better results in statistical reasoning than those at School A, although the results are low for students at both schools. On the contrary, women obtained slightly better results than men in school A. In school B, no differences are observed. Both the effect of gender and the interaction between school and gender are irrelevant.
On the other hand, the results of the two-way ANOVA test for the misconceptions score by school and gender indicate significant differences due to school (p = 0.001), so students from school A have scores significantly lower than those of school B. Again, gender differences are not significant (p = 0.976). Neither is the interaction effect between school and gender (p = 0.542).

As shown in Figure 2, women's scores at both schools are similar to those of their male counterparts. These latter results differ from those Garfield obtained for students in Taiwan and the United States, in which female samples scored significantly higher on misconceptions than their male counterparts.
5 CONCLUSION

The study has allowed us to understand the development of statistical reasoning in ninth grade students, who share a context of similar curricular content, teaching and learning methodologies, and objectives but approach statistics differently. In one of the Colombian schools, school A, statistics is part of the mathematics subject, which occurs more frequently in educational systems around the world. In the other school, school B, statistics is an independent subject, although, according to current legislation, with the same time dedicated as in the previous school.

However, a previous survey carried out in numerous educational centers, as well as teaching practice itself, indicate that the time dedicated to this subject is much less than the corresponding time because it is taught at the end of the mathematics subject, with hardly any time to address it and when students are tired after months of teaching with intensive schedules (Naya et al., 2012). This practical refusal is not exclusive to Colombian educational centers (Vásquez Ortiz & Alsina, 2023) and may be motivated by the relatively recent incorporation of statistics in the secondary school curriculum (about 30 years) and the poor preparation of teachers. (Weiland, 2019; Zieffler et al., 2018).

On the one hand, the sample of students analyzed, the study does not reveal that there is a gender gap in statistical reasoning. Nor does there exist the general existence of the misconceptions analyzed. This contradicts what was obtained by other studies (J. B. Garfield,
2003; Martin et al., 2017; Yusuf & Sukestiyarno, 2022), although it is in line with the results obtained in the work of Gorham and Chamberlin (Gorham Blanco & Chamberlin, 2019). However, significant differences are observed when delving deeper into each of the misconceptions. Specifically, differences occur in misconceptions related to the representativeness of the sample. Men seem to exhibit this misconception to a greater extent. On the other hand, the error related to equiprobability bias is more frequent among women.

On the other hand, it is observed that students (both men and women) in school B, where statistics is taught independently, present a better level of correct statistical reasoning (CR) skills than in the school where it is taught as part of the mathematics subject (school A). This difference is significant, although moderately high.

However, this difference is reversed when analyzing misconceptions. Students from school A perform better in terms of misconceptions (MC) than those from school B. In addition to being significant, this difference is medium in size. Therefore, it seems that learning statistics separately from mathematics facilitates and improves statistical reasoning. On the contrary, when studying the subject of statistics in mathematics, students acquire statistical knowledge to reduce misconceptions. The reduced importance could justify mathematics teachers' attribution to statistics in many cases, as this could significantly reduce the time spent on it in favor of other parts of the subject. This reduced dedication of time forces them to teach the subject from a more conceptual and less procedural point of view. Additionally, this could translate into fewer misconceptions.

On the contrary, when statistics is taught separately, the time dedicated to it is predetermined by the academic calendar of the subject. This allows the subject to be approached in a more leisurely and, therefore, more reasoned and less procedural manner, which leads to better statistical reasoning.

Finally, it should be noted that the authors were surprised by the low results obtained in the statistical reasoning scores, with an average slightly higher than 2 points on a scale from 0 to 8. Furthermore, almost half of the students have misconceptions about the averages, which is also surprising since the concept of the average (or average) is taught from the fourth and fifth grades of primary education, according to the mathematics curriculum in Colombia.

Therefore, from the study carried out, it can be concluded that separating statistics from the mathematics subject could have a positive impact on the statistical reasoning of Colombian high school students, since schools have the independence granted by educational legislation. Colombian to address statistics in this way.
These results lead to the need to replicate the study with larger samples from different cities and countries to determine if it is generalized or concerns these schools or those of this city. As has been mentioned repeatedly in the literature, statistics is a subject that has characteristics that differentiate it from the rest of mathematics, mainly in the use of context and the permanent existence of uncertainty. For these reasons and those explained in this work, it is necessary to determine the need to address statistics separately from mathematics or highlight the importance of this subject in the academic and social development of students.

REFERENCES


Ley 115 de Febrero 8 de 1994 por la cual se expide la ley general de educación, (1994).


