EXPLORING THE PERCEPTIONS OF CHILEAN MATHEMATICS PRE-SERVICE TEACHERS ABOUT A STATISTICAL SIGNIFICANCE LESSON SEQUENCE

Felipe Ruz, Francisca Ubilla, and Valentina Giaconi
Pontificia Universidad Católica de Valparaíso
Pontificia Universidad Católica de Chile
Universidad de O'Higgins
felipe.ruz.a@pucv.cl

Statistical knowledge is culturally significant, making its integration into school curricula crucial. Educators, especially in statistical inference, face challenges. This study explores Chilean preservice mathematics teachers' perceptions of a lesson sequence on statistical significance, focusing on validity (theoretical consistency) and utility (usability for teaching). Employing a Design-Based Research approach, the study involved two focal groups from different Chilean universities, each undergoing a four-session sequence. Participants found the sequence theoretically coherent and effective, progressing from informal to formal inference with robust computational support. Weekly assessments and technological tools were key to its utility. The findings highlight important aspects for educators, providing insights for future improvements and identifying areas needing reinforcement.

INTRODUCTION

Stochastics, encompassing both statistics and probability, has become an essential competence in the information age (Batanero, 2019; Ben-Zvi et al., 2018). This relevance is reflected in the inclusion of statistical education in the school curricula of numerous countries. Despite these curricular efforts, however, many mathematics teachers feel insufficiently prepared to teach this discipline, considering it marginal and outside their primary competence (Batanero, 2009; Batanero et al., 2011; Groth & Meletiou-Mavrotheris, 2018).

An example of this reality is Chile, where statistical education has been incorporated throughout the entire school trajectory (MINEDUC, 2009, 2012, 2015, 2021), but teachers have been shown to have gaps in their competency to teach statistics, even upon recently graduating from their initial training (CPEIP, 2021, 2022, 2023; Ruz et al., 2021). This inadequate preparation creates a significant gap between educational needs and teaching competencies.

The deficiencies in teacher preparation are particularly acute in the area of statistical inference, a crucial component of stochastic knowledge. Statistical inference, and particularly the teaching of statistical significance, presents unique challenges due to its abstract and technical nature (Rossman, 2008). This situation motivates the need to rethink initial teacher training, focusing on the development of practical skills and the use of advanced technologies for teaching these concepts (Ruz et al., 2021; Pfannkuch & Ben-Zvi, 2011). Specifically, it is crucial to equip pre-service teachers with tools/resources and methods that facilitate the effective and accessible understanding and teaching of statistical inference.

An innovative methodology to address these challenges is the Simulation-Based Introduction to Inference (SBI). This approach employs resampling and simulation techniques to teach inferential concepts in an intuitive and practical manner (Lock et al., 2017; Rossman & Chance, 2014; Tintle et al., 2016). Based on Cobb's (2007) three R's model - Randomize, Repeat, and Reject, which allows students to experiment directly with data and inference processes, this approach facilitates a deeper and more practical understanding of statistical significance. This methodology not only improves students' conceptual understanding but also provides future teachers with a solid example of good pedagogical practices in teaching statistical inference.

This study aims to evaluate the quality of a teaching sequence on statistical significance based on the perceptions of future Chilean mathematics teachers in terms of validity (theoretical coherence) and utility (usability for teaching). The study participants included two focus groups (group 1 = eight students; group 2 = six students) of pre-service teachers from two different Chilean universities. Through this research, the goal was not only to improve initial teacher training in stochastics but also to provide a replicable model of good pedagogical practices in teaching statistical significance.

In: Kaplan, J. & Luebke, K. (Ed.) (2024). *Connecting data and people for inclusive statistics and data science education. Proceedings of the Roundtable conference of the International Association for Statistics Education (IASE), July 2024, Auckland, New Zealand.*

METHODOLOGY

This study followed a qualitative interpretative research paradigm, utilizing the Design-Based Research (DBR) methodology. DBR is characterized by an iterative and cyclical approach, structured in three main stages: design, implementation, and retrospective evaluation (DBR Collective, 2003). This paper presents the iteration of a complete cycle of these stages, highlighting the importance of designing robust and practical educational interventions that can be continuously evaluated and refined (DBR Collective, 2003; Easterday et al., 2018).

Moreover, within the framework of DBR, Tena and Couso (2023) emphasized that a sequence can be considered of high quality when it balances the interaction of three aspects: validity (internal and theoretical coherence), utility (practicality and productivity), and reliability of material resources. Thus, in this work, we were particularly interested in analyzing theoretical coherence and utility. Theoretical coherence refers to the ability of the designed sequence activities and tasks to achieve the proposed learning objectives theoretically. In contrast, utility relates to the practical effectiveness and the productivity of the tools used in the implementation of the sequence to achieve the expected learning outcomes (Tena & Couso, 2023).

Context and Participants

For data collection, two focus groups were conducted independently at two Chilean universities and led by researchers different from those who implemented the sessions, following the same application protocol (common guide questions). The objective of the focus groups was to understand the perceptions of pre-service teachers regarding the theoretical coherence and utility of a lesson sequence focused on learning statistical significance.

The first focus group (FG1) included eight future teachers who were taking an elective course on new methodologies for teaching statistical inference, all of whom had previously completed two disciplinary courses in their initial training. The second focus group (FG2) consisted of six future teachers who were taking the third mandatory statistics course of their initial training, focused on statistical inference. In both contexts, the same classes and materials were used, implemented by one of the study's authors. The focus groups lasted approximately 60 minutes and were audio-recorded, the transcripts of which were the main source of information on participants' perceptions.

Organization of the Sequence to Evaluate its Quality

In this paper, we aim to present the results regarding the evaluation of the quality of a sequence, which we describe below. The teaching sequence was implemented within courses spanning a total of 16 weeks, with one weekly face-to-face class of two and a half hours. This study focused on the first four sessions of this course, corresponding to the unit on statistical significance in hypothesis testing for proportion. The learning objectives associated with each session of this unit included

1. Understanding and applying a strategy to evaluate statistical significance using a progressive and simulation-based computational approach.
2. Gradually understanding statistical significance tests and the notion of p-value in tests for proportions.
3. Applying an alternative tool to measure the strength of evidence, such as standardized statistics.
4. Theoretically formalizing hypothesis tests for proportions.

Each session was organized following the six stages of the spiral process proposed by Tintle et al. (2016) to conduct statistical investigations, always starting with a real investigation. These stages are (1) Posing a research question in a real context; (2) Designing the study and collecting data; (3) Exploring the data; (4) Drawing inferences; (5) Formulating conclusions; and (6) Retrospective review and revisiting the other stages. The Rossman/Chance applets collection (https://www.rossmanchance.com/applets/) was used as the main computational tool. Additionally, at the end of each session, a project was given to delve deeper into the topics covered, but its analysis is not the focus of this work.

Therefore, evaluating student perceptions in the first cycle of DBR (design, implementation, and evaluation) is crucial to identify areas for improvement in the didactic sequence, a topic we will delve into in the following sections.

RESULTS

The results are organized from Tena and Couso (2023) framework, which evaluates teaching sequences based on validity (theoretical coherence) and utility (practicality and productivity). These criteria informed the focus group protocol, with questions designed to elicit perceptions of how the sequence supported its intended learning objectives and practical implementation. The four themes under theoretical coherence and the five themes under utility emerged inductively from participant responses during thematic analysis, guided by this framework. While other potential themes were considered, these were prioritized based on their alignment with the study's objectives and the richness of participant reflections. Quotes were selected to represent typical and illustrative responses across the two focus groups, highlighting shared perceptions and occasional divergent views.

Theoretical Coherence of the Sequence

Theoretical clarity. Participants from both focus groups demonstrated a clear identification of the statistical topics covered in the sequence. When asked to name the topics addressed, various terms emerged, indicating a broad disciplinary scope of the sequence. Key terms identified included statistical significance, p-value, simulation, null hypothesis, alternative hypothesis, and standard deviation. This variety of terms suggests a diverse and varied understanding of the content among the participants. When expressing these topics, however, it is noteworthy that FG1 delved into content, while FG2 referred to skills or procedural content.

Coherence and Progression. Participants unanimously described the sequence as coherent and well-structured, logically advancing from basic to more complex concepts. This progressive nature was highlighted as a key factor that facilitated understanding. One participant commented, "*I think it was adequate because we have been working from basic concepts to more complex concepts to verify whether the analysis is correct or not*" (FG1). This indicates the sequence's design adequately and progressively structured the learning, aligning with the intended educational objectives.

Relation Between Planned/Implemented and Perceived Objectives. While participants identified the general topics, there was some ambiguity regarding the specific learning objectives perceived for each session. Examples of this variety included, "*understanding when simulation and theory are applied, both computationally and theoretically, comparing evidence and determining which was more accurate*" (FG1) or "*analyzing how to approach a problem, defining which graph to use, and therefore, what strategy to employ*" (FG2). Despite this, upon being given the list of learning objectives for each session, there was a consensus that the sequence's progression was considered adequate for achieving the educational objectives. Moreover, some participants felt the gradual increase in complexity helped them understand both theoretical and practical aspects of statistical significance.

Computational Simulation. A relevant aspect of the sequence was the use of computational tools, perceived as a significant aid in understanding the concepts. This aligns with the theoretical perspective assumed (Tintle et al., 2016), which places great value on computational simulation to gradually introduce more complex content. Participants mentioned these tools not only facilitated learning but also made the classes more dynamic and interactive. A student expressed, "*I feel it is very beneficial to use [the applets] because they allow generating very large samples that we couldn't do in real life, making it easier to understand certain things*" (FG1).

The sequence was perceived as logically structured, progressing in a way that supported understanding statistical significance. Although there was some lack of precision in recalling the planned/implemented objectives, the general perception was that the sequence was effective in achieving its educational goals. This feedback underscores the importance of clear and incremental learning steps in complex topics such as statistical inference.

Utility of the Sequence: Practicality and Productivity

Pedagogical Practicality. The implementation of the sequence was well received by the participants, who appreciated its didactic and innovative approach. The design of the sequence, emphasizing contextualized learning instead of merely algorithmic processes, was especially valued. A participant highlighted, "*It is very useful that this course is much more pedagogical because the statistics* *seen, at least, are very mathematical, pure mathematics, so here having a much more didactic and innovative approach helps a lot to understand the concepts*" (FG2). This approach, in the words of the pre-service teachers, ensured the tasks were consistent with the expected learning outcomes.

Practicality for Achieving Learning. Participants generally agreed the sequence was effective in helping them achieve the expected learning objectives. The use of progressive weekly tasks and computational tools for simulation were noted as significant contributors to this success. One participant commented on the usefulness of gradual learning: "*I highlight the last session… because we saw how both ways, like the more computational when we talked about the simulation and on the other hand, the theoretical that had to do with the formula came together and were useful to understand the concepts*" (FG2). This highlights that this progressive approach allowed for a positive understanding of both basic and advanced statistical concepts.

Productivity of the Materials. The didactic materials, particularly the weekly projects and the simulator or applets, were highlighted as productive tools. These resources were considered crucial for achieving the learning objectives, although some participants found the weekly projects too extensive and somewhat repetitive. The applet, while valuable, was noted as a tool requiring specific knowledge for effective use, indicating a need for better preparation or simpler alternatives for some students. For instance, a participant observed, "[the applets have] *many buttons and many data points and many things like that… so, although useful, sometimes I felt I needed more guidance to understand how to use it correctly"* (FG1). Students, however, emphasized the flexibility of the teachers in charge of each course and their ability to discuss the present limitations, to optimize the effectiveness of the resources and tasks presented.

Areas for Improvement. Despite the generally positive feedback, some challenges were identified. Participants mentioned obstacles such as limited class time, insufficient prior knowledge, and the pace of feedback on weekly assessments. Specific difficulties included understanding significance and the p-value in the second session and applying alternative tools to measure the strength of evidence in the third session. A student indicated, "*I imagine the teacher assumes that we all understand the topic when it's really not like that, and that can cause difficulties when doing an exercise*" (FG2).

Transferability to School Contexts. Participants felt the structure and methods learned through the sequence could be transferred effectively to school settings, where in Chile, since 2022, understanding statistical significance has been incorporated into the final school grade. Thus, tools like the applet and Kahoot were praised for their practicality in classroom environments. The need for adjustments to fit specific school contexts, however, was emphasized. Participants also highlighted potential issues with the availability of computational resources and the need to carefully consider sensitive topics. A student commented, "*I think it is possible, especially the applet, to try to bring an experiment like this to students* [in schools], *an experiment with coins could be a good introduction for them to learn to identify the graphs*" (FG2).

The sequence was considered positively as practical and productive by the participants, that is, it is useful. Its pedagogical approach, the structure of progressive tasks, and the use of computational tools were key factors in its success. However, challenges such as pace, prior knowledge, and the complexity of the tools need to be addressed. The transferability of the sequence to school contexts appears promising, but adaptations are required to ensure broader applicability and effectiveness.

CONCLUSION

The implementation of a lesson sequence centered on simulation for teaching statistical significance has generated positive results in terms of both theoretical coherence and practical utility. The pre-service teacher participants highlighted various strengths and areas for improvement, providing a comprehensive view of the sequence's effectiveness and potential adaptability in diverse educational contexts.

Discussion

The results support the theoretical coherence and alignment with the intended goals of the didactic sequence in its design and implementation. The progressive structure of the sequence, advancing from basic to more complex concepts, was widely appreciated by the participants. This facilitated a deeper and gradual understanding of the topics, in line with literature that highlights the importance of well-structured teaching for understanding statistical inference (Batanero, 2019; Ben-Zvi et al., 2018). The transition from an informal to a formal approach, supported by computational tools, reflects an effective application of simulation-based methodology, as proposed by authors like Tintle et al. (2016) and Rossman and Chance (2014).

Simultaneously, areas requiring attention were identified, such as the need for greater clarity and time to address certain concepts. This underscores the importance of adjusting the teaching pace and ensuring that all students can adequately follow the content. This finding is consistent with previous studies that highlight the need for additional support in teaching complex statistical concepts such as statistical significance and the p-value to teachers (Batanero et al., 2011; Groth & Meletiou-Mavrotheris, 2018).

Regarding practical utility, the teaching materials, especially the weekly projects and computational tools or applets, were perceived as highly productive. These resources not only facilitated learning but also made classes more dynamic and interactive. This aligns with studies that emphasize the effectiveness of technological tools in teaching statistics (Lock et al., 2017; Pfannkuch & Ben-Zvi, 2011). There was, however, also a noted need for better preparation for the effective use of these tools, indicating that some students required more guidance to handle them properly.

Future projections and recommendations

To improve the lesson sequence, it is essential to consider several key aspects. First, more time and clarity in explaining complex concepts is necessary. This could include additional review sessions and supplementary resources that allow students to reinforce their understanding. Furthermore, training in computational tools should be strengthened. Offering specific workshops on using applets and other technological resources before their implementation in the sequence will ensure that all students are equipped to use them effectively.

Adapting the sequence to specific school contexts is also crucial. The availability of computational resources and the preparation of activities that can be implemented with more accessible tools should be considered. This adaptation will make the sequence relevant and applicable in various educational realities.

Additionally, establishing professional development programs that train practicing teachers in implementing this type of teaching sequence is fundamental. These programs should focus on simulation-based strategies and the use of technological tools in teaching statistics. Continuous teacher training is essential to ensure effective teaching and to keep up with best pedagogical practices in teaching inference.

Finally, the replication and adaptability of the sequence in different educational contexts are promising projections. Therefore, evaluating its effectiveness and making necessary adjustments contributes to more robust and effective teacher training in the field of statistical inference, expanding its applicability nationally and internationally. This perspective will not only improve the initial training of teachers in stochastic but also provide a replicable model of good pedagogical practices in teaching statistical inference, reflecting ongoing development work.

ACKNOWLEDGEMENTS

This research was carried out thanks to funding from ANID-Chile grants: FONDECYT 3220122, FOVI 220056, the Basal Funding for centres of excellence (CMM) FB210005, and the Programa de Inserción a la Academia 2024, Vicerrectoría Académica y Prorrectoría, Pontificia Universidad Católica de Chile.

REFERENCES

Batanero, C. (2019). Thirty years of stochastics education research: Reflections and challenges. In J. M. Contreras, M. Gea, M. López-Martín & E. Molina-Portillo (Eds.), *Actas del Tercer congreso*

Internacional Virtual de Educación Estadística (pp. 1-15). Universidad de Granada. http://hdl.handle.net/10481/55011

Batanero, C. (2009). Retos para la formación estadística de profesores. *II Encontro de Probabilidade e Estatística na Scola*, 1-23.

Batanero, C., Burrill, G., & Reading, C. (Eds.). (2011). *Teaching statistics in school mathematics: Challenges for teaching and teacher education. A joint ICMI/IASE study*. Springer. https://doi.org/10.1007/978-94-007-1131-0

Ben-Zvi, D., Makar, K., & Garfield, J. (Eds.). (2018). *International Handbook of Research in Statistics Education*. Springer. https://link.springer.com/book/10.1007/978-3-319-66195-7

Centro de Perfeccionamiento Experimentación e Investigaciones Pedagógicas. (2021). *Resultados nacionales evaluación nacional diagnóstica de la formación inicial docente 2020*. CPEIP.

Centro de Perfeccionamiento Experimentación e Investigaciones Pedagógicas. (2022). *Resultados nacionales evaluación nacional diagnóstica de la formación inicial docente 2021*. CPEIP.

Centro de Perfeccionamiento Experimentación e Investigaciones Pedagógicas. (2023). *Resultados nacionales evaluación nacional diagnóstica de la formación inicial docente 2022*. CPEIP.

Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational researcher, 32*(1), 5-8. https://doi.org/10.3102/0013189X032001005

Easterday, M., Rees, D., & Gerber, E. (2018). The logic of design research. *Learning: Research and Practice, 4*(2), 131-160. https://doi.org/10.1080/23735082.2017.1286367

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

Lock, R., Frazer, P., Lock, K., Lock, E., & Lock, D. (2017). *Statistics: Unlocking the power of data*. John Wiley & Sons.

Ministerio de Educación Chile. (2009). *CURRICULUM. Objetivos fundamentales y contenidos mínimos obligatorios de la educación básica y media*. MINEDUC.

Ministerio de Educación Chile. (2012). *Bases curriculares educación básica*. MINEDUC.

Ministerio de Educación Chile. (2015). *Bases curriculares 7° a 2° medio*. MINEDUC.

Ministerio de Educación Chile. (2021). *Programa de estudio de 3o o 4o medio, Formación diferenciada matemáticas, Probabilidad y estadísticas descriptiva e inferencial.* MINEDUC.

Pfannkuch, M. & Ben-Zvi, D. (2011). Developing teachers' statistical thinking. In C. Batanero, G. Burrill & C. Reading (Eds.), *Teaching statistics in school mathematics: Challenges for teaching and teacher education. A joint ICMIIASE study* (pp. 323-333). Springer. https://doi.org/10.1007/978-94-007-1131-0_31

Rossman, A. (2008). Reasoning about informal statistical inference: One statistician's view. *Statistics Education Research Journal*, *7*(2), 5-19. https://doi.org/10.52041/serj.v7i2.467

Rossman, A., & Chance, B. (2014). Using simulation-based inference for learning introductory statistics. *Wiley Interdisciplinary Reviews: Computational Statistics*, *6*(4), 211-221. https://doi.org/10.1002/wics.1302

Ruz, F., Chance, B., Medina, E., & Contreras, J.M. (2021). Content knowledge and attitudes towards stochastics and its teaching in pre-service Chilean mathematics teachers. *Statistics Education Research Journal, 20*(1), paper 5. https://doi.org/10.52041/serj.v20i1.100

Tena, È., & Couso, D. (2023). ¿Cómo sé que mi secuencia didáctica es de calidad? Propuesta de un marco de evaluación desde la perspectivade Investigación Basada en Diseño. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, *20*(2), 280100-280115. https://doi.org/10.25267/Rev_Eureka_ensen_divulg_cienc.2023.v20.i2.2801

Tintle, N., Chance, B., Cobb, G., Rossman, A., Roy, S., Swanson, T., & VanderStoep, J. (2016). *Introduction to Statistical Investigations*. John Wiley & Sons.