Innovating mobile learning in statistics: Profiling university students' needs and expectations

Felipe Ruz[1], Francisca Ubilla[2], Patricio Videla[1], Juan Zamora[1] and Javiera Herrera[3]
[1]Pontifical Catholic University of Valparaíso, Chile
[2]Pontifical Catholic University of Chile, Chile
[3]Central University of Chile, Chile
felipe.ruz.a@pucv.cl

This study explores university students' profiles and expectations for the use of a mobile application to support statistics learning. Using a mixed-methods approach, 440 undergraduates participated in a survey combining quantitative Likert-scale responses and over 200 qualitative open-ended answers. A factor analysis revealed four key dimensions shaping students' perceptions toward mobile-assisted learning, while cluster analysis identified three distinctive user profiles with varying levels of engagement. Topic modeling of qualitative responses revealed critical preferences and concerns, such as the need for real-data applications, interactive tools, gamification features, and intuitive design. Concerns included excessive advertising, poor interface usability, and lack of personalized feedback. These findings offer a framework for developing adaptable, student-centered mobile applications in statistics education, aligned with learner diversity and pedagogical goals.

INTRODUCTION

In recent years, the integration of digital technologies into educational environments has reshaped teaching and learning practices in higher education. This transformation has highlighted the need for students to develop advanced competencies in data comprehension and statistical reasoning. These skills are essential across disciplines and professional domains, as they enable individuals to make evidence-based decisions grounded in the analysis and interpretation of data (Carver et al., 2016).

The teaching of statistics plays a central role in this context. Traditionally approached as a mechanical discipline focused on calculations and procedural knowledge, statistics education has progressively shifted toward models that emphasize conceptual understanding and real-world applicability. The Guidelines for Assessment and Instruction in Statistics Education II (GAISE II) advocate for learning experiences that connect statistical content to authentic data, promote reasoning under uncertainty, and integrate technological tools in meaningful ways (Bargagliotti et al., 2020). These recommendations align with current demands for a more engaging, data-literate, and critically thinking student population.

Technology, when thoughtfully implemented, offers powerful opportunities to support this educational shift. In particular, mobile applications have gained increasing attention as tools for enhancing students' engagement and understanding of statistics. These applications can offer interactive environments that combine data visualization, problem-solving guidance, and multimedia resources, facilitating deeper engagement with statistical content (Ben-Zvi et al., 2017). The use of mobile tools can also support asynchronous learning, allowing students to explore concepts at their own pace and revisit complex ideas as needed (Hepworth et al., 2019).

A growing body of research has investigated the features that make educational applications effective for statistics learning. Herrera (2023), in a systematic review of mobile apps in higher education statistics courses, identified six essential characteristics: interactivity, data visualization capabilities, supportive learning materials, use of real and varied data, personalized feedback, and accessible technical design. These features are considered fundamental for fostering meaningful learning experiences and addressing the challenges students often face in acquiring statistical knowledge.

Other studies have echoed these findings. For example, Chance et al. (2004) found that students who used applications that incorporated simulations, and visual representations improved their understanding of sampling distributions. Similarly, González et al. (2018) emphasized the importance of allowing students to manipulate data and generate their own representations, as this empowers them to construct statistical knowledge more actively. Moreover, platforms that integrate instructional support—such as explanatory videos or guided problem-solving steps—have shown to improve student comprehension and confidence (Kolpashnikova & Bartolic, 2019).

In: Birk, L., Loth, G., Jotzo, L., Binder, K., & Frischemeier, D. (Eds.) (2026). *Statistics and Data Science Education in STEAM. Proceedings of the Satellite Conference of the International Association for Statistical Education (IASE), September/October 2025, Münster, Germany*. © ISI/IASE.

However, the successful adoption of these tools is not solely a matter of technical capacity. Several barriers can limit their pedagogical impact, particularly when students' perspectives and preferences are overlooked. Prior research highlights that factors such as user interface complexity, cognitive overload, and lack of alignment with curricular goals may undermine student motivation and willingness to engage with educational applications (Li, 2021; Zetterqvist, 2017). In this regard, the principles of user-centered design (UCD) and design thinking advocate for the early involvement of learners in the development process. These approaches emphasize the identification of user profiles and the creation of adaptable solutions that respond to real user needs (Herrera, 2023).

On the other hand, university students are not a homogeneous group. Their engagement with statistics and educational technology varies according to their academic background, prior experiences, and dispositions toward the discipline. As noted by Zieffler et al. (2008), students' understanding of statistics is shaped not only by instructional design but also by their disposition toward the subject and the nature of classroom interactions. Some students may appreciate applications that promote real-world problem-solving and gamification, while others may prefer structured explanations and minimal distractions. Understanding this diversity is key for designing inclusive educational tools.

Despite this, most existing studies on educational apps in statistics focus on post-intervention evaluation or technical performance. Few adopt a proactive approach that seeks to understand students' needs and expectations prior to the design or implementation of mobile applications. This gap limits our capacity to create truly responsive learning tools and to ensure their relevance across academic contexts. Without a solid understanding of students' perspectives—regarding what they expect from a mobile learning environment, what features they value, and what barriers they anticipate—it is difficult to develop tools that are both pedagogically sound and genuinely useful. Capturing this diversity of expectations is essential for guiding the creation of flexible, student-centered digital tools that respond effectively to learners' realities.

Accordingly, the present study addresses the following research problem: What are the expectations, needs, and usage profiles of university students regarding mobile applications for learning statistics, and how can these insights inform the design of inclusive, user-centered educational technologies? To respond to this problem, the study pursues two specific objectives: (1) To identify and characterize university student profiles based on their experiences, attitudes, and expectations toward mobile-supported statistics learning, and (2) To explore students' perceived requirements and usability preferences for a mobile statistics learning application.

Together, these aims contribute to a better understanding of how mobile technologies can be designed and adapted to support meaningful and equitable learning experiences in statistics education.

METHOD

This study adopted a mixed-methods research design with an exploratory purpose, integrating both quantitative and qualitative components to gain a comprehensive understanding of university students' expectations, needs, and usage profiles regarding mobile applications for learning statistics. The quantitative phase aimed to identify patterns in students' perceptions through a structured survey, analyzed using factor and cluster analysis techniques. In parallel, the qualitative phase sought to explore students' views and usability preferences in depth by analyzing their responses to open-ended questions through topic modeling.

Participants

The study was conducted with a non-probabilistic sample of 440 undergraduate students from a public, non-state university in central Chile. Participants were enrolled in second- or third-year statistics courses as part of their academic programs. The sample was composed of 182 women (41.4%), 252 men (57.3%), and 6 individuals (1.3%) who identified as non-binary or preferred not to disclose their gender. The mean age of participants was 21.25 years. In terms of academic distribution, 89 students (20.2%) came from the Faculty of Basic Sciences, 131 (29.8%) from the Faculty of Engineering, and 220 (50.0%) from the Faculty of Economics. Participation was voluntary, and all students signed informed consent forms through a digital and anonymized platform. The study protocol was reviewed and approved by the university's bioethics committee under approval code BIOPUCV-H 489-2022.

Instruments

A three-part online survey was designed specifically for this study. The first section collected demographic and academic background information, including age, gender, academic program, and year of university entry. The second section included ten Likert-scale items grouped into four theoretical components: (1) design and utility of the application, (2) problem-based learning with real and professional contexts, (3) willingness toward statistics, and (4) classroom participation. These items were rated on a 5-point scale ranging from 1 ("Strongly disagree") to 5 ("Strongly agree").

The third section consisted of three open-ended questions, each designed to explore different aspects of students' perceptions about the use of a mobile application in statistics learning. These questions asked participants to describe: (1) the reasons they would choose not to use such an application; (2) the features they would value most in an educational app; and (3) their views on integrating game-based learning into statistics instruction. Responding to these open-ended questions was optional and resulted in over 200 usable responses per item.

Data Analysis

Quantitative data from the Likert-scale items were first analyzed using exploratory factor analysis (EFA), applying principal component extraction and varimax rotation. The purpose of the EFA was to assess the internal structure of the instrument and confirm whether the theoretical components were empirically supported. Bartlett's test of sphericity and the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy were used to assess suitability for factor analysis. Subsequently, k-means cluster analysis was employed to group students into distinct profiles based on their responses. The optimal number of clusters was determined using hierarchical agglomerative techniques as a preliminary step.

For the qualitative component, the open-ended responses were analyzed using Latent Dirichlet Allocation (LDA), a probabilistic topic modeling technique that detects latent themes within a corpus of text (Blei et al., 2003). LDA was applied separately to each of the three open-ended questions, allowing the identification of dominant topics and associated vocabulary patterns. This analysis provided insight into students' expectations regarding app features, usability concerns, and pedagogical preferences.

RESULTS

Student Profiles Based on Experiences and Expectations

The quantitative component of this study aimed to identify clusters of students who shared similar perceptions and attitudes toward the use of mobile applications for learning statistics. To achieve this, responses to the ten Likert-scale items were first analyzed through exploratory factor analysis (EFA), which confirmed the presence of four coherent dimensions: perceived usefulness and design of the application, problem-based learning preferences, willingness toward the subject, and classroom participation. These dimensions aligned well with the theoretical constructs underpinning the instrument, supporting the validity of the survey structure.

Following this, a cluster analysis was conducted using the k-means method to group students into distinct profiles. Three clusters emerged, each characterized by differing combinations of attitudes, motivational orientations, and learning expectations. The distribution of students across clusters was relatively balanced: Cluster 1 included 161 students (36.6%), Cluster 2 had 155 students (35.2%), and Cluster 3 comprised 124 students (28.2%).

Cluster 1 grouped students who consistently expressed strong support for the integration of mobile applications into statistics learning. Students in this group scored very highly on the perceived usefulness of a mobile app (mean = 4.78), strongly valuing features such as interactive design, data visualization, and the potential to increase motivation. They also showed a strong preference for learning with real and professional data contexts (mean = 4.50), along with moderate to high willingness toward statistics (mean = 3.32) and high classroom participation (mean = 3.75). This group had a balanced gender distribution and representation across the three academic faculties.

Cluster 2 included students who also saw value in using a mobile application, especially in terms of practical utility and design, but who were less enthusiastic about active participation in class or the use of real-world data. Students in this group recognized the potential value of mobile applications,

reporting a similarly high mean score for perceived usefulness (mean = 4.75). However, they were less enthusiastic about other dimensions. Their valuation of problem-based learning with real contexts was considerably lower than Cluster 1 (mean = 3.72), and they exhibited notably low willingness toward statistics (mean = 2.44) as well as minimal classroom participation (mean = 2.49), and the influence of the instructor on their motivation appeared to be less pronounced compared to Cluster 1. The composition of this group included a higher proportion of women, especially from the faculties of Economics and Sciences.

Cluster 3, in contrast, reflected a more skeptical stance. Students in this group were less convinced of the potential benefits of a mobile application for improving learning or engagement. Students in this cluster reported the lowest mean score for perceived usefulness (mean = 3.53), expressing doubts about the benefits of using mobile applications in their learning process. Their valuation of real-world data for learning was also low (mean = 3.54), and while their willingness toward statistics was somewhat better than Cluster 2 (mean = 2.88), their classroom participation remained moderate (mean = 3.19). This group had a strong male representation, particularly from the Faculty of Engineering.

When comparing these profiles across the four factors (or dimensions) derived from the EFA, statistically significant differences were observed. Clusters 1 and 2 scored similarly high on the perceived usefulness of the application but diverged in their engagement with statistics and classroom dynamics. Cluster 3 consistently scored lower across all dimensions, highlighting the presence of a student subgroup with more resistant or disengaged attitudes toward both the discipline and its digital mediation.

These findings underscore the heterogeneity of the student population in terms of how they envision the role of mobile applications in their learning. They also point to the need for flexible, adaptable solutions that can cater to a range of learning dispositions—from highly engaged students seeking deeper interaction with the content, to those who may require more structured support to overcome resistance or lack of motivation.

Students' Requirements and Usability Expectations

The qualitative component of the study was designed to explore students' perspectives on the practical features, design considerations, and pedagogical expectations associated with mobile applications for statistics learning. To this end, students' responses to three open-ended questions were analyzed using Latent Dirichlet Allocation (LDA), which enabled the identification of dominant themes across the corpus of qualitative data.

The first open-ended question asked students to describe what factors would discourage them from using a mobile application for learning statistics. The most frequently cited concerns included excessive complexity in the interface, the presence of intrusive advertisements, and a lack of intuitive design. Many respondents emphasized that if the application were difficult to navigate or overloaded with visual elements, they would be less likely to adopt it. Other students mentioned that overly simplistic exercises or a lack of conceptual depth would also diminish their interest in using such a tool. These concerns suggest that usability and content quality are critical deterrents when poorly implemented.

The second question focused on the features that students would value most in a future application. Three primary themes emerged: the inclusion of step-by-step problem-solving guidance, the ability to manipulate and save tables and graphs, and the presence of explanatory audiovisual materials. Students expressed a clear preference for applications that provide a structured and supportive environment for exploring statistical problems. They emphasized that practical tools enabling data handling, visual exploration, and conceptual explanation—especially through videos—could significantly enhance their understanding. Some also highlighted the importance of being able to use the application both on mobile and desktop platforms to improve accessibility and visual clarity.

The third question explored students' opinions on incorporating game-based learning approaches into the application. Responses revealed an overall positive attitude, particularly when such strategies were linked to real-world scenarios or used to reinforce core statistical concepts. Students appreciated the potential of interactive games to increase motivation, participation, and retention of knowledge. They mentioned examples such as quizzes, simulations, and competitive elements like

Kahoot, which could create a more engaging learning environment. However, some cautioned that games should not replace the core instructional content but rather serve as a complement to reinforce learning.

Taken together, these results highlight the importance of designing educational applications that are not only technically robust but also pedagogically sensitive and adaptable to a wide range of user profiles. By integrating feedback from diverse students and acknowledging their differentiated expectations, developers and educators can co-construct tools that are more likely to succeed in improving engagement and learning outcomes in statistics education.

CONCLUSION

The present study aimed to explore university students' learning profiles and usability expectations regarding the use of mobile applications in statistics education. In pursuit of this goal, two specific objectives were formulated: (1) to identify and characterize student profiles based on their experiences, attitudes, and expectations toward mobile-supported statistics learning; and (2) to explore their perceived requirements and usability preferences for such applications.

In relation to the first objective, the identification of three student clusters revealed meaningful differences in how students perceive the role of mobile technologies in their learning process. While Clusters 1 and 2 expressed similarly high expectations about the usefulness of mobile applications, they differed sharply in their willingness to engage with statistics and their levels of participation in class. Cluster 3, on the other hand, was more skeptical overall, indicating lower receptivity across all measured dimensions. These findings confirm that although technology may be broadly accepted, it does not uniformly influence student motivation or behavior—a result that aligns with prior observations in the literature (Zieffler et al., 2008).

The observed heterogeneity reinforces the need to avoid designing educational tools based on assumptions of homogeneity in the learner population. Rather than treating university students as a uniform group, it is crucial to recognize differentiated profiles that reflect varied levels of interest in statistics, confidence, and engagement with technology. These insights support the foundational principles of user-centered design (Herrera, 2023), which emphasize empathy with end users and the creation of solutions that are responsive to specific user profiles.

As for the second objective, the analysis of students' responses to open-ended questions—processed using Latent Dirichlet Allocation—offered a nuanced view of their expectations for mobile learning environments. Students identified a range of essential features that they would like to see in such tools. These included intuitive navigation, clear visual design, access to real data and examples, and scaffolded guidance such as step-by-step problem-solving tutorials. Many also mentioned the value of incorporating audiovisual content, such as explanatory videos, to support their understanding of key concepts.

Overall, the LDA results revealed a rich set of student expectations that align closely with the key features identified in the literature: interactivity, usability, clarity, and contextual relevance (Herrera, 2023). Importantly, students also introduced perspectives that extend beyond the technical and into the pedagogical—such as the need for differentiated levels of difficulty and mechanisms for tracking progress or receiving feedback. These insights suggest that students are not passive users but critical evaluators of educational technologies, capable of articulating coherent and actionable design expectations.

Additionally, it is particularly notable that students with lower disposition toward statistics—those associated with Clusters 2 and 3—did not reject the idea of using mobile tools. Rather, they emphasized that these tools must be simple, purposeful, and clearly connected to their learning goals. This supports earlier claims that the effectiveness of educational technologies depends not only on their technical functionality but also on how well they are aligned with student needs and contexts (Li, 2021; Zetterqvist, 2017). In this sense, applications that seek to enhance statistics education must incorporate flexible pathways that can accommodate both motivated and hesitant learners.

Taken together, these findings found that the development of mobile applications for statistics learning must be rooted in a deep understanding of learner diversity. Applications should not only deliver content but also support different levels of statistical maturity, motivation, and engagement. By

identifying and analyzing student profiles, this study contributes empirical evidence that can guide the design of inclusive and adaptive learning technologies in the field of statistics education.

On the other hand, the present research has certain limitations that must be acknowledged. First, the sample was drawn from a single Chilean university, which may limit the generalizability of the findings to other institutional or cultural contexts. Future studies would benefit from cross-institutional or international comparisons to test the robustness of the identified profiles. Second, the open-ended responses were voluntary, and not all participants completed them. This may introduce a degree of self-selection bias, as students more comfortable with expressing their views in writing may be overrepresented in the qualitative results. Complementary methods, such as interviews or focus groups, could help deepen the interpretation of these textual data and capture voices less represented in survey responses. A third limitation relates to the absence of an experimental or longitudinal component. Future research could build on these findings by developing and piloting prototypes tailored to specific user profiles, evaluating their effect on student learning outcomes, motivation, and engagement. Finally, while the use of LDA topic modeling allowed for the systematic extraction of themes from a large dataset, it does not replace the interpretive richness that manual coding or thematic analysis might provide. Future work could triangulate these approaches for deeper insight.

Despite these limitations, the study offers a robust empirical contribution to the growing field of technology-enhanced statistics education. It provides both a typology of learner profiles and a prioritized inventory of usability and design expectations, grounded in students' own perspectives. In conclusion, as the integration of digital tools in statistics education continues to expand, the importance of understanding the diverse experiences and expectations of students becomes ever more pressing. Research of this kind can inform the development of tools that are not only technically innovative but also pedagogically meaningful and socially inclusive. In doing so, it reaffirms the value of empirical, student-centered inquiry in advancing more equitable and effective forms of technology-supported learning.

ACKNOWLEDGEMENTS

This study was supported by the Chilean National Agency for Research and Development (ANID) through the FONDEF IDeA I+D project ID24I10075, and the Programa de Inserción a la Academia 2024, Vicerrectoría Académica y Prorrectoría, Pontificia Universidad Católica de Chile. We extend our sincere thanks to the participating students for their valuable contributions, as well as to the academic units that facilitated access to their courses.

REFERENCES

Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Perez, L., & Spangler, D. (2020). *Pre-K-12 Guidelines for Assessment and Instruction in Statistics Education II (GAISE II)*. American Statistical Association.

Ben-Zvi, D., Gravemeijer, K., & Ainley, J. (2017). Design of statistics learning environments. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 473–502). Springer.

Blei, D., Ng, A., & Jordan, M. (2003). Latent Dirichlet Allocation. *Journal of machine Learning research, 3*, 993–1022.

Carver, R., Everson, M., Gabrosek, J., Horton, N., Lock, R., Mocko, M., Rossmann, A., Roswell, G. H., Vellemann, P., Witmer, J. & Wood, B. (2016). *Guidelines for assessment and instruction in statistics education, College report.* American Statistical Association.

Chance, B., delMas, R., & Garfield, J. (2004). Reasoning about sampling distributions. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 295–323). Dordrecht: Springer.

González, J. A., López, M., Cobo, E., & Cortés, J. (2018). Assessing shiny apps through student feedback: Recommendations from a qualitative study. *Computer Applications in Engineering Education, 26*(5), 1813–1824. https://doi.org/10.1002/cae.21932

Hepworth, K., Ivey, C., Canon, C., & Holmes, H. A. (2019). Embedding online, design-focused data visualization instruction in an upper-division undergraduate atmospheric science course. *Journal of Geoscience Education, 68*(2), 168–183. https://doi.org/10.1080/10899995.2019.1656022

Herrera, J. (2023). *Herramientas y aplicaciones para la enseñanza de la estadística en educación superior: Una revisión sistemática* [Unpublished master's thesis]. Pontificia Universidad Católica de Valparaíso.

Kolpashnikova, K., & Bartolic, S. (2019). Digital divide in quantitative methods: The effects of computer-assisted instruction and students' attitudes on knowledge acquisition. *Journal of Computer Assisted Learning, 35*(2), 208–217. https://doi.org/10.1111/jcal.12322

Li, Q. (2021). Computer technology in the course of probability and mathematical statistics. In *Application of intelligent systems in multi-modal information analytics: 2021 international conference on multi-modal information analytics* (Vol. 2, pp. 531–538). Springer.

Zetterqvist, L. (2017). Applied problems and use of technology in an aligned way in basic courses in probability and statistics for engineering students–a way to enhance understanding and increase motivation. *Teaching Mathematics and its Applications: An International Journal of the IMA, 36*(2), 108–122. https://doi.org/10.1093/teamat/hrx004

Zieffler, A., Garfield, J., Alt, S., Dupuis, D., Holleque, K., & Chang, B. (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), 122. https://doi.org/10.1080/10691898.2008.11889566