Testing a digital serious game on statistics learning with future school and high school teachers

Iris Celorrio-Aguilera,[1,2][a], Manuel Freire-Morán[1][b] and Alba García-Barrera[3][c]

[1]*Dept. of Software Engineering and AI. Complutense University of Madrid, UCM, Spain*

[2]*Dept. of Enterprise Sciences and Technology, Open University of Madrid, UDIMA, Spain*

[3]*Dept. of Health and Education, Open University of Madrid, UDIMA, Spain*

icelor01@ucm.es, iris.celorrio@udima.es, manuel.freire@fdi.ucm.es, alba.garcia@udima.es

Keywords: Serious games, Game learning analytics, Statistics, Player experience.

Abstract: This paper presents the pedagogical design and evaluation of a digital serious game, *Biased Distributions*, aimed at teaching central tendency measures to students enrolled in undergraduate and master's education programs, in an interactive and engaging way. The game design is grounded in constructivist learning theory and incorporates elements of experiential training in order to address the challenges often faced by education students in grasping statistical concepts. We describe the design of the game, the methodology applied in the experiment and the results of two rounds of testing with a total of n=56 participants within a higher education distance-learning institution. We evaluate increase in knowledge, player experience, and actual game usage collected via game learning analytics. The results show a moderate increase in knowledge, high user engagement and reveal several usability problems that can be addressed in future game versions, highlighting both the promise of our approach and the importance of validating games to diagnose and improve their effectiveness.

1 INTRODUCTION

Mathematics learning is often perceived by college students as boring, complex, or impractical for real life situations (Dele-Ajayi et al., 2016; Kislenko et al., 2007). Educators also find it more difficult, unexciting or worrying compared to other curricular areas (Macnab and Payne, 2003).

In particular, when learning statistics, which is considered a particular branch of mathematics, students face problems such as conceptual misunderstanding (Saidi and Siew, 2019), possibly due to not using suitable visualizations or not being able to interact with the contents (Pirker and Gütl, 2015). Although statistical concepts are fundamental for understanding data analysis and interpretation, students often struggle with its understanding and application. College students also report other concerns and challenges when studying introductory statistics, such as lack of competence, low intrinsic motivation, or low perceived usefulness of statistics for the real-world (Sutter et al., 2024); as well as the need to deal with formulas which they feel are complicated and unfamiliar (Zakaria et al., 2012; Mohamed et al., 2012).

Serious games (SGs), loosely defined as games with a primary goal other than that of entertainment, have proven to both motivate and boost student learning (M.Connolly et al., 2012), and represent a resource that is increasingly being used for pedagogical purposes, including math fluency and thinking (Fraga-Varela et al., 2021; Paciarotti et al., 2021). While SGs can be carried out with physical materials, digital games, implemented using video games, can be easily reused long-term , while allowing user interactions to be logged and analyzed to better understand how those games are being played, as part of Game Learning Analytics (GLA) as described in (Freire et al., 2023).

We have designed and tested a digital SG that facilitates the learning of statistical concepts, specifically distributions and central tendency measures —mean, median, and mode— targeted towards primary and secondary school teachers. In this work, we explore its design and effectiveness as learning tool The following sections discuss similar approaches to teaching statistics; present the pedagogical design of our SG; and describe and discuss the results of the experiments we have carried out with several groups

[a] https://orcid.org/0000-0002-2189-2154

[b] https://orcid.org/0000-0003-4596-3823

[c] https://orcid.org/0000-0003-1993-1406

of students (totalling n=56 participants) to test the effects of the game in both learning and motivation.

2 RELATED WORK

We have found several SGs that address the field of introductory statistics. To teach statistical concepts to children of primary school, we can find *The Electric Company Prankster Planet* (Meletiou-Mavrotheris et al., 2019), which introduces data collection concepts to 6-10 year old children through quests. The quests involve interpreting and representing statistical graphs, which children must analyze to overcome challenges, and they employ different gamification elements such as avatar creation, a rewards system or a collaborative play option in order to engage students. Results in two 40-minute sessions with 18 students showed that the game was well-received, increasing student motivation.

In the context of higher education, *The Playground Game* (Westera et al., 2014) teaches research methods and statistics through practical cases. It is aimed at students of Social Science and Nursery. Players must decide the most suitable location for a children's playground in a fictitious town, taking into account different statistical variables, justifying the game's name. The gamification techniques employed in the game include narrative, inquiry-based problems, decision-making, questions, score and audio cueing. Participants (N=103, with 4 separate conditions, including a control group) highly appreciated the integration of statistical learning contents in a game environment; and across all conditions, significant positive effects in test scores were observed. Notably, this game includes game learning analytics support, which allows gameplay data to be analyzed to better understand how players learn and interact with the game.

At a similar level, the *Deadly Distribution* SG (Wronowski et al., 2020) teaches concepts such as the sample, mean, standard deviation, margins of error, confidence intervals, and the Central Limit Theorem to 4th-year college students. Based on epidemiology models, players analyze a fictional location to eradicate diseases. The SG employed narrative and levels with increasing problem complexity. The game demonstrated statistically significant improvements in students' affective behaviors (engagement and interest in the subject), according to experiments with 218 students, with a smaller effect on knowledge.

Finally, the *High or Low* game (Siepermann, 2016) uses a card game with two learning modes, teaching and practice, for students of Business Sciences to learn about fundamental principles of statistics and risk management. The main gamification techniques employed in the game are luck, risk, and decision-making, in combination of helping elements such as hints or recommendations. In an experiment with 42 participants, players proved able to estimate risk correctly, with an interesting twist where the game proposed possibly-manipulated recommendations, forcing students to double-check before relying on them at face value. The game's effectiveness at increasing grades was not tested, nor was its affective effect on players.

After analyzing these games, we found that none of them exactly matched our needs and the way we wanted to convey its contents, since we wanted to cover a more specific topic of statistical frequency distributions, in a very interactive and visual way, in the language of the course - Spanish - as well as provide insights to Educational Sciences students into how serious games could help to teach such concepts.The *Biased Distributions* SG covers conversion of observations to distributions and the measures of central tendency in those distributions (mean, median and mode), with the goal of not only help future primary and secondary education teachers (and currently students of Educational Sciences) to assimilate these concepts in a more entertaining, interactive and visual way, but also to guide them when teaching the same concepts to their students later on, as both statistical distributions and measures of central tendency are often part of elementary and secondary school mathematics curricula.

3 DESIGN OF *Biased Distributions*

Our design is based on constructivist theory, and specifically on building knowledge through experience (Bada and Olusegun, 2015): immediately after being introduced to explanations of each idea or concept, players can put them into practice. We have opted for simplicity in the interface, to allow players to concentrate on the requested activities without losing focus.

The game is composed of two parts: an initial guided tutorial to learn both game mechanics and the concepts of statistical distributions and calculation of central moments; and a game phase in which these concepts are applied across three game levels of increasing difficulty, as illustrated in Figure 1. To make abstract concepts as visually affordable as possible, addressing concerns of (Pirker and Gütl, 2015), we designed the game so that distributions are test-scores of simulated individuals, each of which is represented

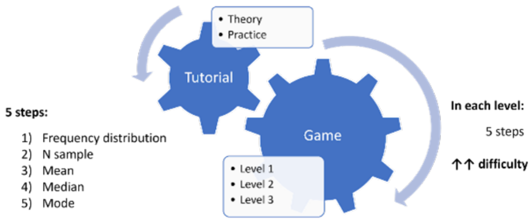

Figure 1: Diagram with the general structure of the SG. The tutorial has shorter and more controlled steps than the main game, which has a greater degree of freedom, and iterates through the core concepts with increasing difficulty.

with a small avatar. An initial step is to drag each avatar icon from the observation table to its corresponding histogram bar, building the histogram incrementally. Players can switch back and forth between their current exercise and the underlying theory, by clicking the "Formulas" button (see bottom-left corners of Figure 2). Additionally, players can interact with a small on-screen calculator (also visible in Figure 2), where operands can be inserted via keyboard or by interacting with the visualization: clicking anywhere on a displayed bar inputs its height (frequency) to the current calculation; clicking on the x-axis of the bar inputs its x-coordinate (observation value). One of the tutorial's goals is to make sure that players are exposed to all mechanics, from histogram-building to mean, median, and mode calculations. Later on, exercises are repeated with identical mechanics but increasingly complex distributions, progressively increasing the difficulty of the game.

The game's pedagogical framework integrates three key components:

- Scaffolded learning progression: Players advance through levels that gradually increase in complexity, from basic concept introduction to applied problem-solving.
- Interactive data visualization: Dynamic and manipulable graphs allow students to explore the effects of data changes on central tendency measures in real-time.
- Contextual learning scenarios: The game presents educational situations where the knowledge to be learnt is highly relevant.

Implementation

From a technical standpoint, the game was developed using Unity3D, which allows creation of full-featured cross-platform games with moderate programming knowledge, and is free for small developers. The game is composed of several types of scenes, roughly corresponding to mini-games such as building a histogram or calculating a distribution's mean with the aid of a histogram and a calculator (as seen in Figure 2.

The contents of the game's questions, and the scene types where those questions are displayed, are specified in a teacher-editable JSON file, which is read during game initialization. In file, it is possible to specify either static datasets or datasets that are to be randomly generated within certain parameters. We did not test the extent to which this file can be easily edited by real users, as our primary goal was to evaluate the game itself. However, having a degree of customization available to those that deploy a game seems to be a very important element for future reuse. Such game customization allows game authors to adjust the game's difficulty, flow, and length to the specific population where it will be played. We then packaged the game for three different platforms according to the operating system used by students: Windows, Linux or MacOS X; for all versions, the games could log how they were being played using xAPI-SG statements (see Results), allowing game learning analytics techniques to be used to understand how they were being played.

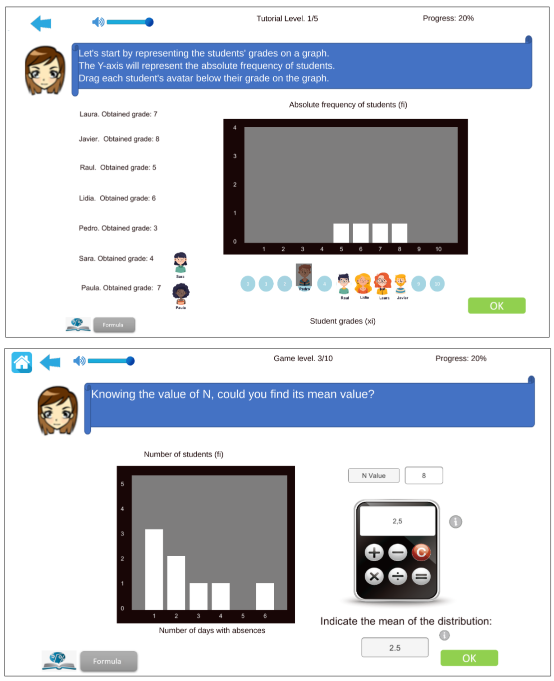

Figure 2: Screenshots from the game, translated to English. (top) Players can drag observations to build a histogram. (bottom) Players computes the distribution's mean by clicking on bars in the histogram, which updates the calculation displayed in the calculator.

4 EVALUATION

We evaluated *Biased Distributions* using a mixed-methods approach on students from 2 distinct programs: 17 participants were enrolled in a "Mathematical Fundamentals" course from an undergraduate degree to become primary education teachers; while 39 were enrolled in a master's degree subject called "Teaching Innovation and Introduction to Educational Research in the Mathematics Specialty", with 56 participants in total. Experiments were conducted during academic year 2022-23, with 2 groups of master's students and 2 groups of graduate students. Data collection was based on a series of pre- and post- knowledge experimental tests, surveys, and game learning analytics to study how students actually interacted with the game.

Participants were all students in the First Author's remote-learning university. Participants were chosen not only because they were expected to learn the concepts seen in the game, but also because, as future educators, they could provide useful feedback on how these concepts could be taught; and would, in many cases, have to teach them to their future students. The experiments were therefore designed to help detect problems in the game design, determine whether learning was occurring, and to verify the extent to which the SG motivated students and helps them to understand more abstract contents through visual representation and interaction. Since all participants were remote, most of the tasks in our experimental design (except for Step 4: playing the game itself) were mediated through the university's institutional Learning Management System. In order, participants were asked to:

1. Provide or deny informed consent to use their data, anonymized, for research purposes.
2. Answer a survey of prior knowledge. Participants rated, using 7-point Likert scales, their level of knowledge regarding mathematics and statistics in general and statistical distributions and central tendency measures in particular.
3. Fill a pre-test knowledge questionnaire.
4. Download and play the game.
5. Fill a post-test questionnaire, with different questions as the pre-test in task 3, but otherwise intended to be equivalent.
6. Answer a self-assessment knowledge survey.
7. Post-activity satisfaction questionnaire.
8. Fill a final Technology Acceptance Model survey (Marangunić and Granić, 2015).

Participants that completed all tasks received 0.5 extra points on the course grade (5% of the total grade). As part of step 4, participants were requested to submit at least one gameplay trace as proof-of-playing, but had the option to send more than one. Despite the grade incentive, only 56 out of 146 students completed until Task 6 and 51 out of 146 students completed all tasks. A breakdown of participation by degree can be seen in Table 1. This is likely due to the scheduling of the experiment, which was very close to the last weeks of the course, when students were busy with exams.

Table 1: Research sample, where n = Number of students completing up to Task 6, n_{all} = Number of students completing all Tasks, N = Number of total students in the course.

Level	n	n_{all}	N	n/N
Education Degree	17	14	50	34.00%
Master's Program	39	37	96	40.63%
Total	56	51	146	38.36%

Game Analytics

We collected and analyzed the in-game interactions of players generated during activity 4, storing them as xAPI-SG statements (Serrano-Laguna et al., 2017). This allows us to gain insights into how the game was actually played by each player, gaining additional value by taking into account all other activities by that same player. This analysis requires two key ingredients: first, the game must store and allow retrieval of player interactions for analysis; for this, we included an analytics library into the game's Unity code which can both store and, if so configured, submit user interactions to a server in the xAPI-SG format. Second, participant activities, including the interaction statements, must be linked together with a suitable identifier, allowing interactions from a single player, and other player activities such as questionnaire answers, to be linked for analysis. We used unique numerical identifiers for this purpose.

4.1 Results

Comparing pre- and post-knowledge tests of the total sample of students (activities 3 and 5), 46.43% of students improved their grades after playing the game, 30.36% worsened, and 23.21% remained at the same grade. The group average varied from 7.71 points (on a 0-10 scale) to 8.07 points, with an average score increase of 0.36 points in the sample group (improvement in percentage of 3.6%). As can be seen in Figure 3, players from the undergraduate groups generally improved in their scores after playing the game (by 2 and 0.61 points in each group), while master's students did not change their scores significantly (less than 0.1 point difference). A paired sample t-test confirms that the increment is statistically significant (with a p-value of 0.045, testing whether the post-test has higher scores than the pre-test) for the 1st undergraduate group, but no statistically significant difference (in either sense) is found for the other groups. We did expect master's students to have more knowledge on the subject, as confirmed by their higher grades, and therefore to gain less by playing the game.

Comparing the results obtained from the pre- and post-surveys (activities 2 and 7), students perceived, on average, that their general knowledge of mathematics had increased (from a mean of 4.86 to 5.29 points on a 7-point scale), as well as their level of knowledge in the subject (4.58 vs. 5.29). Regarding their level of interest in the subject, it decreased slightly (5.83 vs. 5.71 points), although they value that serious educational games can be an interesting resource for learning subject concepts, and strongly consider the game to be a novel resource for learning subject concepts (6.43). Furthermore, as a key aspect of the research, the majority of students reported an improvement in their general knowledge of statistical distributions (3.92 vs. 5.57 points) and on how to calculate mean, median and mode for such distributions (+2.04, +2.08 and +1.89 points on a 7-point scale, respectively).

Figure 4 illustrates a first approach to gameplay data analysis, showing distributions for the number of xAPI statements for each player according to their group (which roughly corresponds to the amount of in-game interactions), and the amount of time spent playing. The last plot in Figure 4 shows that most students managed to finish the game (which required 20 correct answers), and a significant number chose to play far beyond: the average across all groups is 25, with many players, mostly in master's groups, significantly exceeding this number.

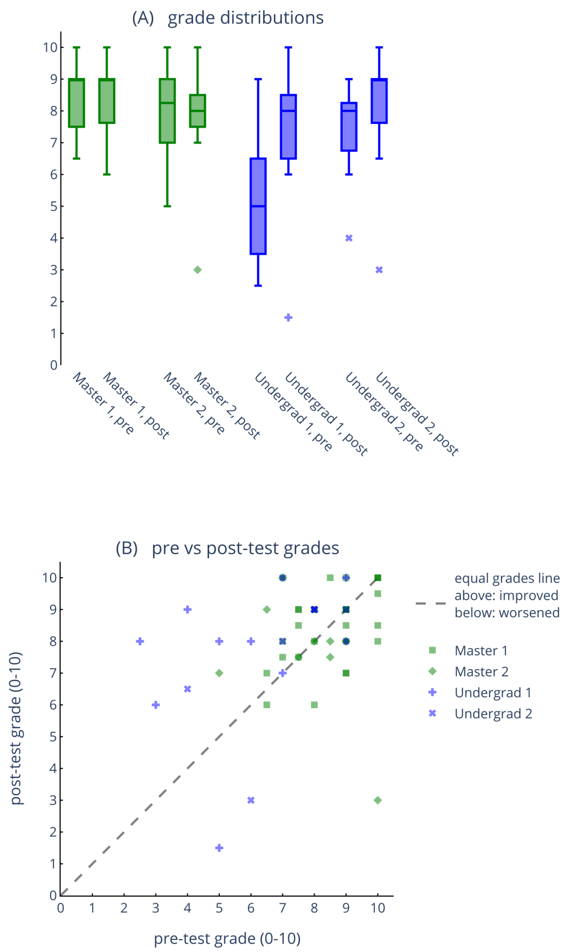

Figure 3: Pre-post score differences in players, on a 0-to-10 scale, 10 =best. (A) shows general score distribution in pre and post-tests for each group; (B) shows improvement as a scatterplot. Master's students generally had higher grades, and improved less than undergraduate students.

Qualitative Results

The satisfaction survey (activity 8) contained free-text answers, which the three authors encoded collaboratively, agreeing on keywords to assign for each response of each of the 56 participants before any quantitative analysis. When asked for aspects that they liked, 23% of players reported finding the game simple and intuitive. A large portion of players (21.4%) also mentioned that they liked the theoretical explanations of the tutorial, as well as the game's educational content.

Likewise, 21.4% of players found the game practical, and 17.9% mentioned that they found it very useful and liked being able to jump back and forth between practical examples and the applicable formulas and theoretical background at any time while playing the game. This same percentage of players (17.9%) indicated that they found the game entertaining and enjoyable. Players commented positively on the dynamics of the game, learning by playing, the incremental progression of difficulty, the reminder/help formulas button, hint availability, the brevity and clarity explanations, and the design of practical exercises. In the same satisfaction survey, another question asked for free-text answers on what they disliked. The most common response (17.9%) was to leave the field blank or comment "nothing". Next, many com-

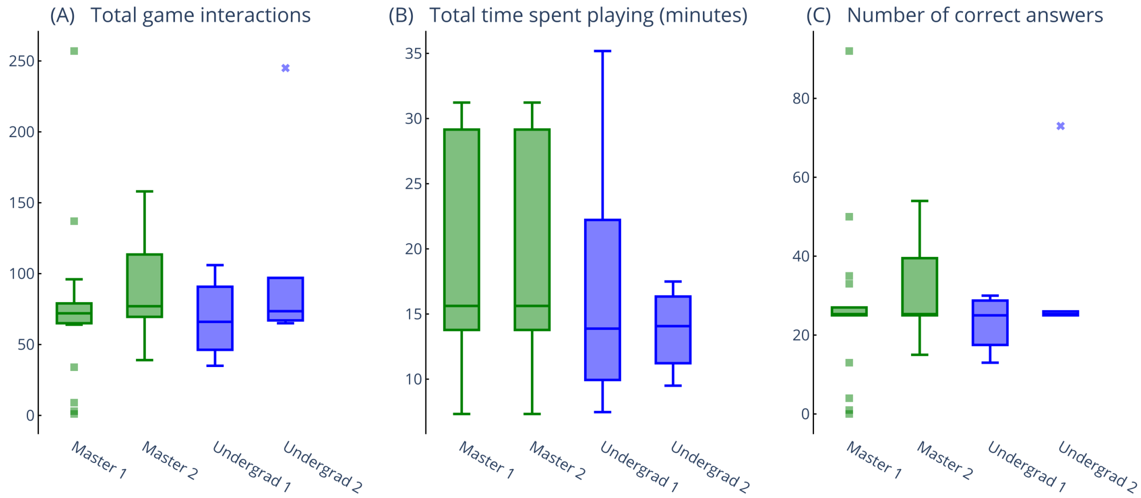

Figure 4: Analyzing in-game interactions by players, segmented by groups. In (A), all groups of players can be seen to have interacted a similar amount of times with the game, while in (B), master's students can be seen to have taken longer to finish playing. Outliers that took long breaks (3 in total, in 2 of the groups) have been removed from (B). In (C), all students can be seen answering a similar number of correct answers.

plained about the calculator (visible at the right of Figure 2), as some values had to be entered by using a keyboard; that the exercises were somewhat repetitive, with the same calculations using different values (7.1%); limited levels (7.1%); the music, which sometimes made it difficult to concentrate on solving the exercises (7.1%); and the fact that, while feedback did detect incorrect results, it did not always pinpoint their cause (3.6%). A similar number complained that the game was too short, or that there should be levels with a higher difficulty.

Some participants also reported technical problems: the game failed to work as expected on Mac operating systems (7.1%), downloading it could be slow (5.7%), with further slowdowns if running in an emulated environment (5.7%); finally, reporting game analytics was not always easy (3.6%). In the words of one of the master's students, "It is a very engaging and innovative way to learn. I am looking forward to becoming a teacher so I can apply it to a session as part of the statistics and probability teaching unit. It allows students to interact with mathematical concepts in a more dynamic and participatory way, which can facilitate learning and understanding. Games can promote learning through problem-solving and decision-making."

5 CONCLUSIONS

Our findings align with previous literature, showing that students highly appreciate the integration of educational content within serious games. The game captured student attention and kept them highly engaged and motivated by practicing statistical concepts, with students reporting that they specially valued the immediate feedback, personalized learning, the experience of hands-on learning, and the self-assessment facilitated by the game, continuing to play even after completing all mandatory exercises (a behavior known as free choice in motivation theory). Thus, the game design is promising, despite poor calculator mechanics and a certain lack of variety in the exercises; and some minor technical problems related to cross-platform deployment, all of which can be solved in future versions.

Our future steps will be to develop and validate a new version of the game based on improvements suggested by participants in our pilot experiment, including:

- Better contextual help to aid students that get stuck on calculations, by modelling frequent errors and preparing canned feedback for each.
- Improved usability of the built-in calculator, with additional tutorial steps to explain changes.
- More exercise types and variety to add both replayability and depth.

The next round of experiments should also feature a larger pool of participants, which will allow us to use separate groups to compare learning and engagement associated with individual mechanics, and adding a control group to better test the effectiveness of the proposal.

Comparing the design proposed for our SG with those presented in the related work of this paper, all four games rely on inquiry-based problems, questions, levels, progressive increase in difficulty, hints, and feedback. Similarly to those games, we also report a (modest) increase in learning, and a much more significant effect in player motivation and engagement. We expect that even better results and player experience could be achieved by incorporating additional elements from other games, such as introducing a narrative thread, reinforcing the current minimal story, adding maps within an encompassing world, a reward system, more customization options, and/or elements of luck or risk that require decision making and provide greater replay values.

Evaluation is critical to ensure that SGs are effective both in terms of knowledge gain and to realize their main potential of boosting student motivation and engagement. We have described how ours has helped us to improve our game design and player experience, with students who, in the future, will not only be able to apply SGs in their classrooms once they become teachers, but may hopefully also design their own.

REFERENCES

Bada, S. O. and Olusegun, S. (2015). Constructivism learning theory: A paradigm for teaching and learning. *Journal of Research & Method in Education*, 5(6):66–70.

Dele-Ajayi, O., Sanderson, J., Strachan, R., and Pickard, A. J. (2016). Learning mathematics through serious games: An engagement framework. *2016 IEEE Frontiers in Education Conference (FIE)*, pages 1–5.

Fraga-Varela, F., Vila-Couñago, E., and Martínez-Piñeiro, E. (2021). The impact of serious games on mathematics fluency: A study in primary education. *Comunicar*, 29(69).

Freire, M., Serrano-Laguna, Á., Manero Iglesias, B., Martínez-Ortiz, I., Moreno-Ger, P., and Fernández-Manjón, B. (2023). Game learning analytics: Learning analytics for serious games. In *Learning, design, and technology: An international compendium of theory, research, practice, and policy*, pages 3475–3502. Springer.

Kislenko, K., Grevholm, B., and Lepik, M. (2007). Mathematics is important but boring : students' beliefs and attitudes towards mathematics. In *Relating practice and research in mathematics education : proceedings of NORMA 05, Fourth Nordic Conference on Mathematics Education, Trondheim, 2nd-6th September 2005*, pages 349–360. Tallin University, Tapir Academic Press.

Macnab, D. S. and Payne, F. (2003). Beliefs, Attitudes and Practices in Mathematics Teaching: Perceptions of Scottish primary school student teachers. *Journal of Education for Teaching*, 29(1):55–68.

Marangunić, N. and Granić, A. (2015). Technology acceptance model: a literature review from 1986 to 2013. *Universal Access in the Information Society*, 14(1):81–95.

M.Connolly, T., A.Boyle, E., MacArthur, E., Hainey, T., and M. Boyle, J. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education*, 59(2):661–686.

Meletiou-Mavrotheris, M., Tsouccas, L., and Paparistodemou, E. (2019). Digital games as tools for enhancing statistics instruction in the early years: A teaching intervention within a grade 2 mathematics classroom. In Gentile, M., Allegra, M., and Sobke, H., editors, *Games and Learning Alliance, GALA 2018*, volume 11385 of *Lecture Notes in Computer Science*, pages 414–417. Serious Games Soc; Natl Res Council Italy, Inst Educ Technol.

Mohamed, H., Ashaari, N. S. ., Judi, H. M., and Wook, T. S. M. T. (2012). Factors Affecting FTSM Students' Achievement in Statistics Course. *Procedia - Social and Behavioral Sciences*, 59:125–129.

Paciarotti, C., Bertozzi, G., and Sillaots, M. (2021). A new approach to gamification in engineering education: the learner-designer approach to serious games. *European Journal of Engineering Education*, 46(6):1092–1116.

Pirker, J. and Gütl, C. (2015). Educational Gamified Science Simulations. In *Gamification in education and business*, pages 253–275. Springer International Publishing.

Saidi, S. S. and Siew, N. M. (2019). Assessing students' understanding of the measures of central tendency and attitude towards statistics in rural secondary schools. *International Electronic Journal of Mathematics Education*, 14(1):73–86.

Serrano-Laguna, Á., Martínez-Ortiz, I., Haag, J., Regan, D., Johnson, A., and Fernández-Manjón, B. (2017). Applying standards to systematize learning analytics in serious games. *Computer Standards & Interfaces*, 50:116–123.

Siepermann, M. (2016). Teaching statistics and risk management with card games. In DeGloria, A. and Veltkamp, R., editors, *Games and Learning Alliance, GALA 2015, Revised Selected Papers*, volume 9599 of *Lecture Notes in Computer Science*, pages 171–180. Serious Games Soc; Univ Genoa.

Sutter, C. C., Givvin, K. B., and Hulleman, C. S. (2024). Concerns and challenges in introductory statistics and correlates with motivation and interest. *The Journal of Experimental Education*, 92(4):662–691.

Westera, W., Slootmaker, A., and Kurvers, H. (2014). The playground game: Inquiry-based learning about research methods and statistics. In Busch, C., editor, *Proceedings of the 8th European Conference on Game-Based Learning (ECGBL 2014), VOLS 1 AND 2*, Proceedings of the European Conference on Games-Based Learning, pages 620–627.

Wronowski, M., Urick, A., Wilson, A. S. P., Thompson, W., Thomas, D., Wilson, S., Elizondo, F. J., and Ralston, R. (2020). Effect of a serious educational game on academic and affective outcomes for statistics instruction. *Journal of Educational Computing Research*, 57(8):2053–2084.

Zakaria, E., Zain, N. M., Ahmad, N. A., and Erlina, A. (2012). Mathematics anxiety and achievement among secondary school students. *American Journal of Applied Sciences*, 9(11):1828.