

Lehren und Lernen mit Intelligenten Systemen

Bachelorarbeit

**Gender Differences in the Impact of Gamification Elements on
Performance and Anxiety**

Robin Gebert

Studiengang: Informatik - M. Ed. (Erw)

Prüfer/in: Jun.-Prof. Dr. rer. nat. Maria
Wirzberger

Betreuer/in: Nadine Nicole Koch, M.Sc.
Niklas Meißner, M.Sc.

Beginn am: 5. Juni 2023

Beendet am: 5. Dezember 2023

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Keywords: Gamification, Gender

Gender Differences in the Impact of Gamification Elements on Performance and Anxiety

Theoretical Background

Tutorial systems

Tutorial systems have become increasingly popular in educational settings, offering a wide range of features and flexibility to support students learning processes. These systems can range from simple instructive texts to simulations and virtual realities, serving as models that simplify aspects of the real world to reduce complexity mostly from interconnection and context of knowledge for both the machine and the user (Psołka et al., 1988). Those Environments are in use at higher education institutions and since the COVID-19 pandemic, they have become more important (El Hadbi et al., 2024) and widely researched, especially in the field of computer science (Zawacki-Richter et al., 2019).

Tutorial systems are often enhanced with some sort of intelligence (ITS). A dynamic adaptation to learner's needs, incorporating usage factors such as performance, but also external factors such as age, culture or gender (González et al., 2014; Nkambou et al., 2010). The role of Artificial Intelligence in this process could also become more important in the future, as it could have a significant, yet unknown impact on ITS (Zawacki-Richter et al., 2019).

Tutoring Systems are often used in combination with gamified elements, significantly improving the learning experience (Dermeval et al., 2019). Jackson and McNamara (2013) found that the use of gamified elements in tutoring systems significantly improved the motivation and performance of students compared to traditional tutoring systems which lead to boredom and disengagement after long periods of use.

Incorporating gamified elements not only enhances the engagement and motivation within the ITS but also necessitates mechanisms for tracking progress, such as content unlocking (González et al., 2014). The evolving landscape of ITS research also includes

emotional and relational dynamics, linking student emotions and teacher-student relationships to learning efficacy and motivation (Woolf et al., 2010). These insights have led to the development of digital companions, often named pedagogical agents, within ITS that significantly boost the learning potential and self-concept of students, particularly those who are low-achieving. Intriguingly, a study noted that ITS programs with a male companion were muted twice as often as those with a female companion, highlighting potential gender differences that could be explored to enhance the predictive capabilities of the student model (Woolf et al., 2010).

Gender and Stereotype threat

Gender, as a concept within social sciences, refers to more than the binary categorization of male and female. It encompasses a range of identities and experiences that are shaped by a complex interplay of biological, psychological, and social factors. Gender is not solely determined by biological characteristics; instead, it is increasingly recognized as a spectrum, acknowledging the presence of diverse gender identities beyond the traditional binary understanding (Lindqvist et al., 2021). Socialization plays a critical role in shaping gender identity. It influences how individuals perceive themselves and interact with their surroundings based on the gender norms prevalent within their society. These norms dictate behaviours, roles, and expectations, which are often internalized from an early age through various socialization agents like family, media, educational institutions, and peer groups (Kampshoff & Wiepcke, 2012). While acknowledging the spectrum of gender identities, this thesis will focus primarily on the binary categorization of gender—male and female. This approach does not negate the validity of non-binary or genderqueer identities but rather limits the scope of investigation to traditional gender roles within the binary framework.

Stereotype threat occurs when "one can be judged by, treated in terms of, or self-fulfill negative stereotypes about one's group". Although this study does not aim to

eliminate stereotype threat it is an important factor as it can explain at least some differences different genders experience while studying computer science (Cheryan et al., 2011), especially regarding math (Spencer et al., 1999). Stereotype threat even leads to lower identification with academics and specific subjects (Christy & Fox, 2014).

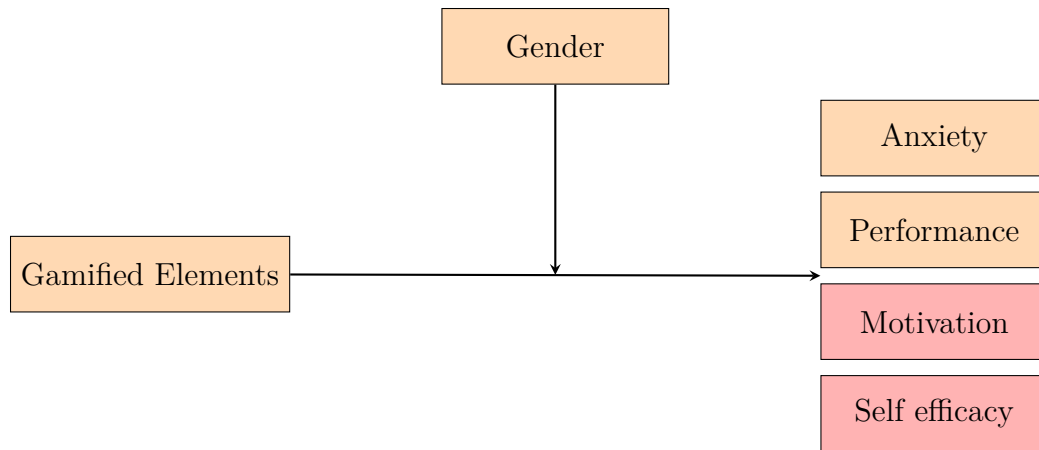
Gamification

Gamification can be defined as "the idea of using game design elements in non-game contexts" (Deterding et al., 2011) to further increase motivation and user activity within interaction design (Deterding et al., 2011). These game-design elements, subsequently called gamified elements, are elements often found in classical video games. However, the concept of gamification is different from designing a game, the focus lies on using the addictive component (González et al., 2014). Often used elements are points, badges, leaderboards and avatars, other mechanisms include content unlocking, storytelling and memes (Zainuddin et al., 2020). Often those elements are used specific constellations like the PBL triad described by Werbach and Hunter (2012), which contains points, badges and leaderboards. A system that is not only known from games, but also everyday enterprise features like loyalty programs and employee competitions (Werbach & Hunter, 2012). Points because they add an absolute scale, badges because they represent a status symbol and work like a temporary goal to strive toward and leaderboards to compare yourself to peers (Werbach & Hunter, 2012). One of the positive effects of gamification is brought by the feedback in different forms (task, process, self-regulation, self) either immediate or delayed. Feedback is one of the most important factors in the relation between education and learning Sailer and Homner (2020). The use of gamified elements showed positive outcomes in multiple studies, in general (Hamari et al., 2014) as well as in education specific contexts (Sailer & Homner, 2020). But gamification, especially some elements like leaderboards, can also lead to negative outcomes. Leaderboards, while motivating through comparison, have been reported to demotivate participants (Almeida et al., 2021).

"Pavlovication" as Klabbers (2018) calls it, Gamification, as it is often a short question-answer-reward-cycle, conditions the user to learn conditional and narrows the possible ways to solve a problem down (Klabbers, 2018). Some studies also suggested that gamified learning platforms also lack individualism regarding choice and display of gamification elements, resulting in discomfort and negative emotions (Santos et al., 2023). To combat this missing individualism, Dehghanzadeh et al. (2024) and Oliveira et al. (2023) suggest using more independent variables to tailor the use of gamification elements.

Hypotheses

As noted in the first chapter, there are open questions regarding the efficiency of various gamified elements and how different genders relate to these gamified elements. The question of the efficiency of certain elements and combinations of elements remains unresolved (Dehghanzadeh et al., 2024). To explore the connection between gender and gamification elements, we have created the following model:



This model additionally incorporates concepts of motivation and self-efficacy, which, although not featured in my thesis, are included in the doctoral thesis of **Nadine Koch**. Since males perform better than females in solving progressive matrices from age 15 onward, Hypothesis **H1a** one-sidedly formulated (**ravenStandardProgressiveMatrices2003**). The hypotheses we want to investigate in

this work are:

H1 Males and females differ in their cognitive and affective states.

- a) Male performance is better compared to female.
- b) Male and female students differ regarding their anxiety levels.
- c) Male and female students differ regarding their motivation.
- d) Males have a higher self-efficacy compared to females.

H2 Different gamified elements have a varying impact on the cognitive and affective states.

- a) Gamified elements impact performance differently.
- b) Different gamified elements impact anxiety levels differently.
- c) Different gamified elements impact motivation differently.
- d) Different gamified elements impact self-efficacy differently.

H3 Different gamified elements differently impact the cognitive and affective states of males and females.

- a) The influence of different gamified elements on performance differs between males and females.
- b) The influence of different gamified elements on anxiety levels differs between males and females.
- c) The influence of different gamified elements on motivation differs between males and females.
- d) The influence of different gamification elements on self-efficacy differs between males and females.

All hypothesized effects result from interacting with the gamified digital learning environment.

Methods

Participants

119 participants were recruited from the university campus, with 69.74% identifying as 'male', 28.57% as 'female' and 1.68% as 'other'. Participants were aged between 18 and 27, the average age was 22 years, with a standard deviation of 2.4 years. The study programs most frequently represented were computer science, software engineering and aerospace engineering. Participants each received €15 as monetary compensation for their involvement in the study.

Design

This study explored the impact of various gamified elements and participant gender on performance and anxiety. The independent variables were gamified elements, with participants randomly assigned to one of eight conditions: Avatars (A), Badges (B), Points (P), Leaderboards (L), Narrated Content (N), combinations of Points, Badges, Leaderboards, and Avatars (PBLA), Points, Badges, Leaderboards, Avatars, and Narrated Content (PBLAN), and a control group with no gamified elements. Each participant experienced three distinct conditions, which were sent by the server out of a randomized pregenerated batch, ensuring that all conditions were evenly distributed across participants. Participants underwent a series of tests in a fixed order during each round, beginning with a gamified performance test in a digital learning environment followed by not gamified assessments for anxiety, self-efficacy, and motivation. At the end participants were given a monetary compensation of €15. The performance tests utilized standard progressive matrices, adapted with gamification techniques to engage and challenge participants uniquely in each round. The dependent variables included:

- **Performance**, assessed through accuracy and response times in the gamified progressive matrices.

- **Anxiety**, evaluated using a standardized questionnaire immediately after the performance test. Anxiety was measured using a shortened form of the State-Trait Anxiety Inventory (STAI) with 6 items (Marteau & Bekker, 1992).

Although self-efficacy and motivation were also assessed (Chen et al., 2001; Guay et al., 2000) through subsequent questionnaires, these variables were not analyzed within the scope of this bachelor thesis. The collected data for self-efficacy and motivation are intended for use in the doctoral dissertation of **Nadine Koch**. This research employed a repeated-measures design, where each participant was exposed to three different gamification conditions chosen randomly. This within-subjects approach facilitated the analysis of individual responses to each condition across the different rounds, providing insights into how variations in gamification can affect psychological states and performance. The sequence and consistency of the testing procedure, including the series of questions asked in the gamified digital learning environment were always maintained to ensure the reliability of measurements and comparability of results across the various stages of the experiment.

Materials

Physical environment

The study was conducted in two separate rooms in the cellar of a university building, one equipped with five and one with seven iMacs. As Christy and Fox (2014) suggested that the physical environment can influence the results, so both rooms are equipped with the same furniture and lighting and are furnished very dry, like a typical software laboratory.

Virtual environment

The software used in this study was build by the author using SvelteKit in frontend and KTor in backend. Its UI is designed after the study by Albuquerque et al. (2017). On

the iMac's the study was displayed full-screen mode using the Safari web browser to ensure no further distractions. The study consisted of 4 screens. A consent screen to give an overview and explain the data collection to the user. A personal detail screen to collect said data; gender, age and study program. Participants also had to enter a deletion code in order to request their data's deletion after the collection.

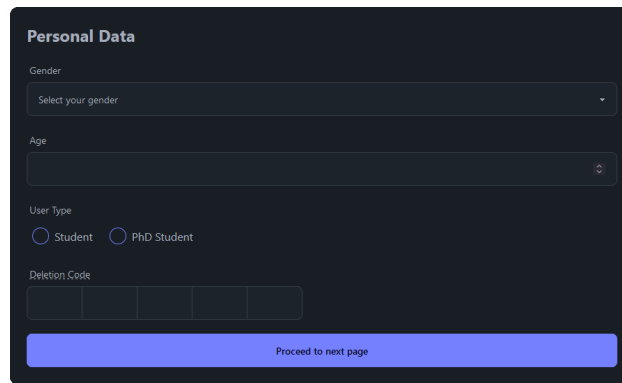


Figure 1

The personal details collection form

The next screen was the gamified learning environment, where the participants had to solve 20 questions in a row while being exposed to the gamified elements. The matrices were taken from Albuquerque et al. (2017), to generate 60 questions out of the 20, the 40 questions for iteration two and three were slightly altered versions of the original 20 made by this author.

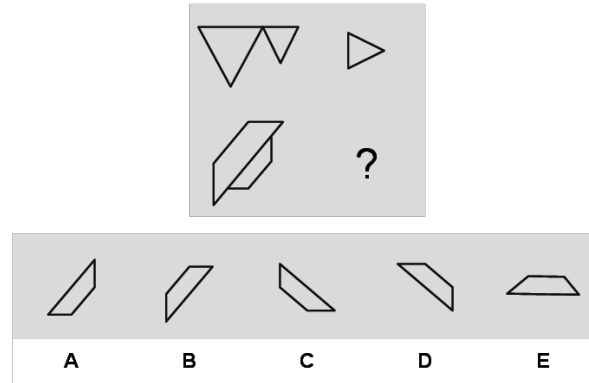


Figure 2

A standard progressive matrix, one of the tasks given to the participants

The gamified learning environment consists of different UI elements representing the gamified elements.

Leaderboards : A list of participants and scores, including the current participant. The other players shown are not real.

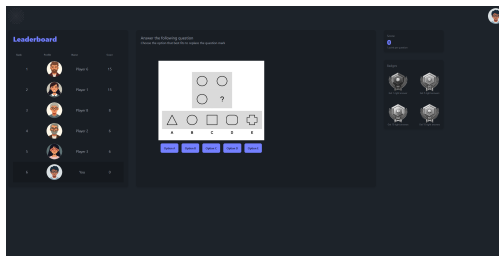
Badges : An array of four badges that are awarded for 1, 5, 10 and 18 correctly answered questions.

Avatars : A small avatar that is shown in the top right corner of the screen and on the leaderboard. To increase identification with the avatar further, the participants were asked to choose one of 15 different avatars before the iteration.

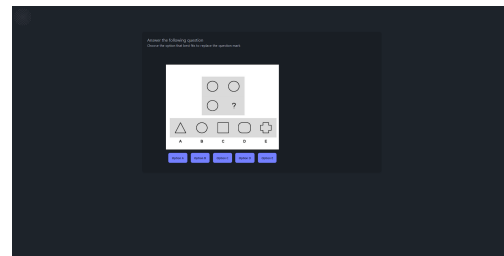
Narrated content : Narrated content is shown in the bottom right corner of the screen. It is presented as a speech bubble with an avatar next to it, in case avatars are enabled. It shows a random praise or encouragement sentence every three questions.

Points : A counter next to the question frame shows the current points. One point is awarded for each correctly answered question. The narrated content is shown every three questions.

After answering one question the next question has a one-second delay which increases to four seconds if narrated content is shown.



(a) *The Digital Learning Environment with Points, Badges, Leaderboards and Avatars enabled*



(b) *The Digital Learning Environment with no Gamified Elements enabled*

Figure 3

Comparison of the Digital Learning Environment with and without gamified elements enabled

After the gamified learning environment, the participants were shown a questionnaire for anxiety, motivation and self-efficacy. The three questionnaires were a six-question shortened form of the State-Trait Anxiety Inventory (STAI) (Marteau & Bekker, 1992), the eight-question General Self-Efficacy Scale (GSE) (Guay et al., 2000) and the 16-question Situational Intrinsic Motivation Scale (SIMS) (Chen et al., 2001).

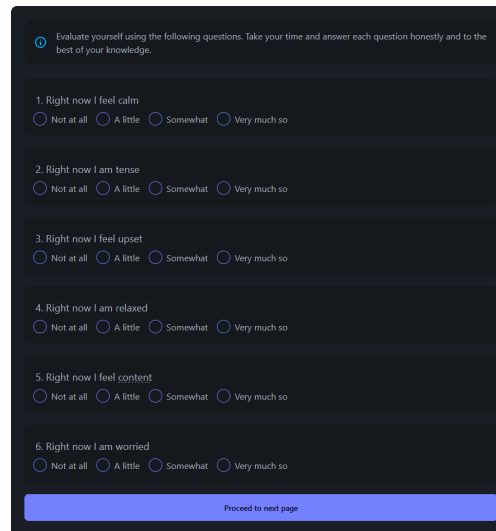
The image shows a digital questionnaire interface on a dark background. At the top, there is a blue header bar with a white circle containing the number '1' and the text: 'Evaluate yourself using the following questions. Take your time and answer each question honestly and to the best of your knowledge.' Below this, there are six numbered questions, each followed by four radio button options. The questions are: 1. 'Right now I feel calm', 2. 'Right now I am tense', 3. 'Right now I feel upset', 4. 'Right now I am relaxed', 5. 'Right now I feel content', and 6. 'Right now I am worried'. The options for each question are: 'Not at all', 'A little', 'Somewhat', and 'Very much so'. At the bottom of the form, there is a blue button with the text 'Proceed to next page'.

Figure 4

The anxiety questionnaire

To submit data to the backend there is a data submission screen that guides the participants to the next iteration or in case of the third iteration to the end of the study.

Procedure

Participants were enlisted at the university campus and invited to engage in a study concerning gamification, with an incentive of 15€ offered upfront for their involvement. Following a brief overview of the study's framework, they were directed to select both a room and a computer. The initial screens presented were those seeking consent and outlining the study details. Subsequently, participants inputted their data, leading into a series of three questioning phases. Each phase initiated with the gamified learning environment, followed by three questionnaires, and concluded with a data submission interface. Participants were advised to proceed at their own pace and refrain from communicating with fellow participants throughout the duration of the study. Upon completion of the third iteration, they were acknowledged for their contribution and compensated with the 15€.

Scoring

Scoring was done in R manually. The data was cleaned up before the analysis. The scores for the different conditions were calculated as follows:

Performance was calculated as the sum of the correctly answered questions divided by all questions. If this value was below 0.25 the particular dataset was excluded from the analysis.

STAI was calculated using the formula provided by Marteau and Bekker (1992). As participants had answers from "Not at all" to "Very much so" the answers are represented by numbers from zero to five. As the original test has 20 questions, weights according to Marteau and Bekker (1992) were applied to the answers. Negative weights were applied for negative questions like "Right now I am worried".

New GSE was calculated using the formula provided by Guay et al. (2000). The mean of the number representation of the answers was calculated.

References

- Albuquerque, J., Bittencourt, I. I., Coelho, J. A., & Silva, A. P. (2017). Does gender stereotype threat in gamified educational environments cause anxiety? An experimental study. *Computers & Education*, 115, 161–170.
<https://doi.org/10.1016/j.compedu.2017.08.005>
- Almeida, C., Kalinowski, M., & Feijó, B. (2021). A Systematic Mapping of Negative Effects of Gamification in Education/Learning Systems. *2021 47th Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*, 17–24.
<https://doi.org/10.1109/SEAA53835.2021.00011>
- Chen, G., Gully, S. M., & Eden, D. (2001). Validation of a New General Self-Efficacy Scale. *Organizational Research Methods*, 4(1), 62–83.
<https://doi.org/10.1177/109442810141004>
- Cheryan, S., Meltzoff, A. N., & Kim, S. (2011). Classrooms matter: The design of virtual classrooms influences gender disparities in computer science classes. *Computers & Education*, 57(2), 1825–1835. <https://doi.org/10.1016/j.compedu.2011.02.004>
- Christy, K. R., & Fox, J. (2014). Leaderboards in a virtual classroom: A test of stereotype threat and social comparison explanations for women’s math performance. *Computers & Education*, 78, 66–77. <https://doi.org/10.1016/j.compedu.2014.05.005>
- Dehghanzadeh, H., Farrokhnia, M., Dehghanzadeh, H., Taghipour, K., & Noroozi, O. (2024). Using gamification to support learning in K-12 education: A systematic literature review. *British Journal of Educational Technology*, 55(1), 34–70.
<https://doi.org/10.1111/bjet.13335>
- Dermeval, D., Albuquerque, J., Bittencourt, I. I., Isotani, S., Silva, A. P., & Vassileva, J. (2019). GaTO: An Ontological Model to Apply Gamification in Intelligent Tutoring Systems. *Frontiers in Artificial Intelligence*, 2.
<https://doi.org/10.3389/frai.2019.00013>

- Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining "gamification".
- El Hadbi, A., Hachem, E. K., Bourray, H., Rziqi, M. H., & Oubalahcen, H. (2024). Review Study on the Adaptive and Personalized Intelligent Tutoring Systems in E-learning. In M. Ezziyyani, J. Kacprzyk, & V. E. Balas (Eds.), *International Conference on Advanced Intelligent Systems for Sustainable Development (AI2SD'2023)* (pp. 45–68). Springer Nature Switzerland.
https://doi.org/10.1007/978-3-031-54288-6_5
- González, C., Mora, A., & Toledo, P. (2014). Gamification in intelligent tutoring systems. *Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality*, 221–225. <https://doi.org/10.1145/2669711.2669903>
- Guay, F., Vallerand, R. J., & Blanchard, C. (2000). On the Assessment of Situational Intrinsic and Extrinsic Motivation: The Situational Motivation Scale (SIMS). *Motivation and Emotion*, 24(3), 175–213. <https://doi.org/10.1023/A:1005614228250>
- Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does Gamification Work? — A Literature Review of Empirical Studies on Gamification.
<https://doi.org/10.1109/HICSS.2014.377>
- Jackson, G. T., & McNamara, D. S. (2013). Motivation and performance in a game-based intelligent tutoring system. *Journal of Educational Psychology*, 105(4), 1036–1049.
<https://doi.org/10.1037/a0032580>
- Kampshoff, M., & Wiepcke, C. (Eds.). (2012). *Handbuch Geschlechterforschung und Fachdidaktik*. VS Verlag für Sozialwissenschaften.
<https://doi.org/10.1007/978-3-531-18984-0>
- Klabbers, J. H. (2018). On the Architecture of Game Science. *Simulation & Gaming*, 49(3), 207–245. <https://doi.org/10.1177/1046878118762534>
- Lindqvist, A., Sendén, M. G., & Renström, E. A. (2021). What is gender, anyway: A review of the options for operationalising gender. *Psychology & Sexuality*. Retrieved

May 18, 2024, from

<https://www.tandfonline.com/doi/abs/10.1080/19419899.2020.1729844>

Marteau, T. M., & Bekker, H. (1992). The development of a six-item short-form of the state scale of the Spielberger State—Trait Anxiety Inventory (STAI). *British Journal of Clinical Psychology*, 31(3), 301–306.

<https://doi.org/10.1111/j.2044-8260.1992.tb00997.x>

Nkambou, R., Bourdeau, J., & Mizoguchi, R. (Eds.). (2010). *Advances in Intelligent Tutoring Systems* (Vol. 308). Springer. <https://doi.org/10.1007/978-3-642-14363-2>

Oliveira, W., Hamari, J., Shi, L., Toda, A. M., Rodrigues, L., Palomino, P. T., & Isotani, S. (2023). Tailored gamification in education: A literature review and future agenda. *Education and Information Technologies*, 28(1), 373–406.

<https://doi.org/10.1007/s10639-022-11122-4>

Psotka, J., Massey, L. D., & Mutter, S. A. (1988). *Intelligent tutoring systems: Lessons learned*. Lawrence Erlbaum Associates, Inc.

Sailer, M., & Homner, L. (2020). The Gamification of Learning: A Meta-analysis. *Educational Psychology Review*, 32(1), 77–112.

<https://doi.org/10.1007/s10648-019-09498-w>

Santos, J., Andrade, E., Benevides, K., Silva, K., Nascimento, J., Bittencourt, I., Pereira, M., Fernandes, S., & Isotani, S. (2023). Does gender stereotype threat affects the levels of aggressiveness, learning and flow in gamified learning environments?: An experimental study. *Education and Information Technologies*, 28(2), 1637–1662. <https://doi.org/10.1007/s10639-022-11220-3>

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype Threat and Women's Math Performance. *Journal of Experimental Social Psychology*, 35(1), 4–28.

<https://doi.org/10.1006/jesp.1998.1373>

Werbach, K., & Hunter, D. (2012, January 1). *For the Win: How Game Thinking can Revolutionize your Business*.

- Woolf, B. P., Arroyo, I., Cooper, D., Burleson, W., & Muldner, K. (2010). Affective Tutors: Automatic Detection of and Response to Student Emotion. In R. Nkambou, J. Bourdeau, & R. Mizoguchi (Eds.), *Advances in Intelligent Tutoring Systems* (pp. 207–227). Springer. https://doi.org/10.1007/978-3-642-14363-2_10
- Zainuddin, Z., Chu, S. K. W., Shujahat, M., & Perera, C. J. (2020). The impact of gamification on learning and instruction: A systematic review of empirical evidence. *Educational Research Review*, 30, 100326. <https://doi.org/10.1016/j.edurev.2020.100326>
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1), 39. <https://doi.org/10.1186/s41239-019-0171-0>