GENERATIVE ARTIFICIAL INTELLIGENCE IN PRIMARY AND SECONDARY EDUCATION IN PORTUGAL: ACCEPTANCE AND USE BY STUDENTS

Nuno Dorotea and Célia Ribeiras
UIDEF – Unidade de Investigação e Desenvolvimento em Educação e Formação
Institute of Education, University of Lisbon
PORTUGAL

Abstract

Generative Artificial Intelligence (GenAI) has penetrated the world globally, prompting several structural changes and new ways of dealing with knowledge. Major challenges are emerging in education. This research aims to analyse the factors that influence the use and acceptance of GenAI by primary and secondary school students in Portugal through UTAUT2 (Unified Theory of Acceptance and Use of Technology). With 478 participants, the data was collected in 2024 and analysed using the partial least squares method. Results indicate that Habit emerged as the most influential factor on Behavioural Intention, followed by Performance Expectancy, Hedonic Motivation, and Personal Innovation. Habit and Behavioural Intention demonstrated significant impact on Behavioural Intention.

Introduction

Generative Artificial Intelligence (GenAI) is a consequence of technological evolution and of the human desire to surpass their own limits, creating systems that reproduce intelligent behaviour in an artificial way (Oliveira, 2019, p. 2). Like any other technology, its presence in education is inevitable. Since November 2022, with the public release of ChatGPT, this topic has flooded the global educational landscape. On the one hand, Liu et al. (2023, p. 73) points out that GenAI "…can improve the learning process and experience for students"; on the other hand, its constant and ongoing emergence and development demands "more research (…) to determine its effectiveness in different contexts" (Su & Yang, 2023, p. 362) and requires understanding their capabilities and limitations. This is a new, emerging, and overwhelming technology, creative and generative, that, when it appears, marks only the beginning of something much greater, in constant growth, and whose impact on education is not yet fully understood. What is clear is that AI will have significant consequences for education. As Holmes, Bialik, and Fadel (2019) point out, "however, while many assume that artificial intelligence in education means students being taught by robot teachers, the reality is more prosaic yet still has the potential to be transformative. Nonetheless, the application of AI to education raises far-reaching questions" (p. 80). The interaction between

Generative AI and education extends beyond the classroom to include teaching about AI and preparing for Human-AI collaboration. The introduction of AI into education raises questions about pedagogy, access, ethics, equity, and sustainability, and calls for a continuous reassessment of the foundational principles of education. The pedagogical advantages of using generative AI in education are numerous. Liu et al. (2023) argue that GenAI technologies, "together with other forms of AI, can enhance the learning process and experience for students due to their ability to access and generate information" (p. 73).

However, as an emerging technology, there is still much to learn, identify, and explore, especially due to the challenges it presents, such as the errors and inaccuracies it may produce, the biases in its results, not only due to the algorithms used but also the data employed for machine learning, as well as the so-called "hallucinations" (Adiguzel et al. 2023; Su & Yang, 2023; Sullivan et al., 2023; Tlili et al., 2023; Liu et al., 2023), and the ethical, moral, and legal issues associated with it. In order to better understand how students in non-higher education perceive, accept, and use Generative AI, it is crucial to investigate whether studies exist on the acceptance and use of this technology in primary and secondary education, not only in Portugal but also abroad. And ultimately, if no alternatives are available, at other levels of education. Based on these premises, the present study aims to identify and analyse the factors that influence the acceptance and use of Generative AI in academic contexts by students in primary and secondary education in Portugal.

Methodology

To analyse the factors influencing the adoption and use of Generative AI, a literature review was conducted to identify the most appropriate theoretical model for measuring levels of technology acceptance.

Theories of technology acceptance and use have emerged over many years of research, resulting in different models with similar purposes. Some of the most used models include the TAM (Technology Acceptance Model), the TPB (Theory of Planned Behaviour), the MM (Motivation Model), the Model of PC Utilization (MPCU), the IDT (Innovation Diffusion Theory), and the SCT (Social Cognition Theory. All these models contributed to the construction of the UTAUT model (Unified Theory of Acceptance and Use of Technology) and, consequently, to the UTAUT2 model. Venkatesh et al. (2003) selected the constructs and respective theories they considered most effective in identifying the factors that most impact technology acceptance and use "both in organizational and non-organizational contexts" (Venkatesh et al., 2012). For the present study, the UTAUT2 model was adopted, which is an extension of the UTAUT model developed by Venkatesh, Thong, and Xu (2012).

We considered the following constructs from UTAUT2: Performance Expectancy (PE), Effort Expectancy (EE), Social Influence (SI), Facilitating Conditions (FC), Hedonic Motivation (HM), Habit (HB), Personal Innovation (PI), Confidence (CO), and Perceived Risk (PR). The aim was to identify the impact of these factors on the constructs Behavioural Intention (BI) and Frequency of Use (FU). Additionally, two moderating variables were considered: Gender and School Level.

Performance Expectancy (PE) considers the extent to which individuals believe they can improve their performance by using a given technology.

Effort Expectancy (EE) refers to the ease of using a technology, that is, the level of effort required to use the technology.

Social Influence (SI) refers to the influence that other people (whether relevant or not to the individual) have on an individual's use of a given technology. This construct is a key determinant of Behavioural Intention, as it is known that an individual's behaviour is influenced by how they believe others will perceive them as a result of using the technology (Venkatesh et al., 2003).

Facilitating Conditions (FC) reflect the degree to which an individual believes there is support for using a technology. Directly linked to the technological and organizational environment surrounding individuals, this support should contribute to solving problems that may arise and consequently remove barriers.

Hedonic Motivation (HM) is synonymous with pleasurable feelings: "the fun or pleasure derived from using a technology, and it has been shown to play an important role in determining technology acceptance and use" (Venkatesh et al., 2012, p. 161).

Habit (HB) aims to determine the extent to which individuals behave automatically when handling a technology. This automatism is directly linked to the learning individuals have acquired or developed through the use of that technology.

Personal Innovation (PI) "refers to an individual's willingness and ability to adopt and use new technology in their daily life" (Strzelecki, 2023, p. 4). According to the same author, this construct is an essential addition to the UTAUT2 model, generally defined as the level of willingness to embrace new technologies while simultaneously demonstrating comfort and confidence in handling them.

Confidence (CO) in technology refers to "the users' belief that the use of technology is reliable and trustworthy" (Al-Azawei & Alowayr, 2020, p. 5). In other words, it concerns confidence in the outputs of Generative AI, and their credibility, which should not be “blind” but rather measured, informed, and cautious.

Perceived Risk (PR) is directly related to security, knowledge about dangers and issues associated with Generative AI, data protection and privacy, ethical concerns, as well as misinformation and biases/prejudices exhibited by these applications. These are linked to overconfidence, which can result in a lack of critical thinking and creativity. These challenges are so widespread that there exists, both within and beyond education, a sense of distrust, threat, and discomfort regarding Generative AI. In other words, this construct determines an individual’s assessment of the potential risks or uncertainties of a given situation, which in this case is the use of Generative AI. According to Yao et al. (2024), “previous research has shown that risk perception plays a crucial role in shaping individuals’ attitudes and intentions towards adopting new technologies” (p. 6).

Behavioural Intention (BI) refers to the “likelihood or subjective intention of an individual to use a particular technology in the future” (Venkatesh et al., 2012, cited in Strzelecki, 2023, p. 4).

Frequency of Use (FU) refers to how often an individual uses a particular technology.

In all constructs, the concept of technology was adapted to Generative AI (GenAI). In total, 49 items were considered, aiming to identify the factors that contribute to the adoption and frequency of use of Generative AI (GenAI) by students in primary and secondary education. Based on these constructs, 13 hypotheses were developed to demonstrate the relationships among them.

Validation of the Questionnaire

The quality analysis of the questionnaire was conducted for all constructs, except for Frequency of Use, as it contained only a single item. Table 1 presents the results of the reliability analysis of the different constructs in the questionnaire, as well as of the instrument as a whole.

According to George and Mallery (2003), the Cronbach’s Alpha value for each construct should be above 0.7 to ensure that the internal consistency of the data is acceptable, i.e., to ensure data reliability. As shown in Table 1, the Cronbach’s alpha values for all constructs are above 0.8, with most being very close to or above 0.9, indicating that the questionnaire demonstrates near-excellent internal consistency. The overall reliability of the scale in this study is 0.952. Similarly, the composite reliability values are very close to or above 0.9.

Table 1
Measurement Model Indicators

Constructs	Number of items	Average	Standard Deviation	Cronbach's Alpha (α)	Composite reliability	Average Variance Extracted (AVE)
Facilitating Conditions (FC)	5	4.91	1.27	.808	.880	.602
Confidence (CO)	4	4.20	1.52	.892	.924	.753
Performance Expectancy (PE)	6	4.95	1.42	.934	.948	.752
Effort Expectancy (EE)	4	5.27	1.36	.903	.932	.774
Habit (HB)	4	3.69	1.66	.897	.928	.765
Behavioural Intention (BI)	3	4.60	1.61	.912	.944	.850
Personal Innovation (PI)	4	4.16	1.54	.879	.918	.738
Social Influence (SI)	8	3.94	1.41	.928	.941	.671
Hedonic Motivation (HM)	4	4.75	1.44	.917	.942	.804
Perceived Risk (PR)	6	4.80	1.28	.871	.898	.598
Global	**48**	**4.51**	**.94**	**.952**	**-**	**-**

The various constructs were also analysed in terms of convergent validity and discriminant validity.

The Average Variance Extracted (AVE) is a metric used to assess the convergent validity of a construct in Structural Equation Modeling (SEM). It helps measure how much of the variance in a set of indicators is explained by the construct they are intended to measure. Thus, it ensures that the constructs in a model are adequately represented by the indicators and helps to ensure the robustness of the analyses based on that model.

For a construct to demonstrate convergent validity, an AVE value greater than 0.50 is a mandatory requirement (Hair et al., 2014; Henseler et al., 2009). This indicates that more than 50% of the variance in the indicators is explained by the construct. Analysing the AVE values for each construct (Table 1), it is observed that all are above 0.598, which suggests that the constructs are capable of explaining at least 60% of the variance they are intended to represent.

Discriminant validity was also assessed using the Fornell and Larcker (1981) criteria, including the Heterotrait-Monotrait ratio (HTMT), and cross-loadings.

According to Fornell and Larcker (1981) and Hair et al. (2014), discriminant validity is assessed using the Fornell-Larcker criterion, which is confirmed when the square roots of the AVE values are greater than the correlations between constructs (i.e., the bold value for each construct must be higher than all the values in the intersection of that construct with the others). In this study, this requirement is met. On the other hand, it is important to analyse the cross-loadings, whose ideal values require that each item loads more highly on the construct to which it is theoretically linked than on any other construct. The Heterotrait-Monotrait Ratio (HTMT) is another metric, considered more precise for assessing discriminant validity in structural equation models (SEM). Discriminant validity refers to the extent to which a construct is truly distinct from other constructs in the model.

A value below 1.0 generally indicates good discriminant validity, suggesting that the constructs are distinct from one another. HTMT values below 0.85 are acceptable and suggest adequate discriminant validity.

Based on the analysis of the model HTMT values (below 0.83), it can be concluded that the model meets the criteria for convergent and discriminant validity, thereby ensuring the consistency of its structure and the reliability of subsequent statistical inferences.

Since all conditions for all measured constructs were met, it can be concluded that these constructs are suitable for estimating their impacts on Behavioural Intention and Frequency of Use of Generative AI applications.

Participants

A total of 478 students from Primary and Secondary Education participated in the study, with an average age of 15 years. Of these, 50.4% were male and 49.6% were female. The data were collected between January and May 2024 and subsequently analysed statistically using the Smart PLS-SEM software.

Results

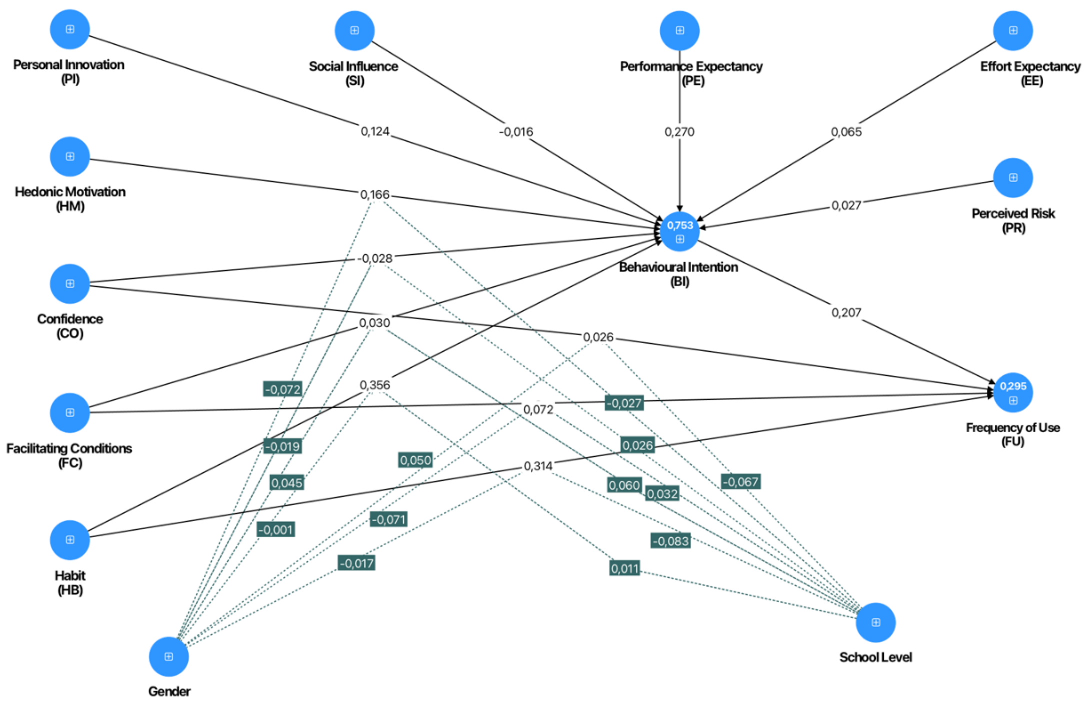
Partial Least Squares Structural Equation Modelling (PLS-SEM) provides diagrammatic representations that visually illustrate hypotheses and relationships between constructs (Hair et al., 2021). In these models, constructs, or latent variables, are depicted as circles or ovals. The relationships between constructs, and between constructs and indicators, are represented by unidirectional arrows, suggesting predictive or, when supported by robust theory, causal relationships. PLS-SEM comprises two main components: the structural model (or inner model), which connects the constructs and displays the relationships among them, and the measurement models (or outer models), which show the relationships between constructs and their indicators.

To estimate the model, we used the PLS-SEM algorithm with the path weighting scheme through the SmartPLS4 software (Version 4.1.0.8), running 5,000 bootstrap samples to determine the statistical significance of the PLS-SEM results, as recommended by Ringle et al., (2022).

Below, we present the results of the relationships between constructs and their influence on Behavioural Intention and Frequency of Use.

Figure 1

Results of the GenAI Acceptance and Utilisation Model: Structural Model of Student Acceptance and Use of AI Defined for the Study



The coefficient of determination R^2 is used to determine the explanatory power of each construct and of the overall model. Ranging between 0 and 1, higher R^2 values indicate greater explanatory power. According to Hair et al. (2021), R^2 values of 0.25 are considered weak, 0.50 moderate, and 0.75 substantial. Figure 1 presents the results of the PLS-SEM analysis, indicating the relationships between the constructs and the R^2 values, which are displayed inside the circles. As we can see, 75.3% of the variance in Behavioural Intention (BI) can be explained by the other constructs, and only 29.5% of Frequency of Use (FU) can be explained by Behavioural Intention (BI), Confidence (CO), Facilitating Conditions (FC) and Habit (HB).

According to Sarstedt, Ringle, and Hair (2022), f^2 (f-squared) determines the effect size of a construct. Values around 0.35 correspond to large effects, 0.15 to medium effects, and 0.02 to small effects. Values of f^2 below 0.02 suggest no effect. Table 2 shows the effect size of the hypotheses confirmed for this study.

Table 2

Path Coefficients and Significance Test Results

Hypo-thesis	Relationships	*Path Coefficients*	*P values*	f^2	Confirmed
H1	(PE) Performance Expectancy -> (BI) Behavioural Intention	.270	**.000**	.114 ++	**Yes**
H2	(EE) Effort Expectancy -> (BI) Behavioural Intention	.065	.097	.008	No
H3	(SI) Social Influence -> (BI) Behavioural Intention	-.016	.570	.001	No
H4	(FC) Facilitating Conditions -> (BI) Behavioural Intention	.030	.404	.002	No
H5	(FC) Facilitating Conditions -> (FU) Frequency of Use	.072	.096	.005	No
H6	(HM) Hedonic Motivation -> (BI) Behavioural Intention	.166	**.001**	.039 +	**Yes**
H7	(HB) Habit -> (BI) Behavioural Intention	.356	**.000**	.180 ++	**Yes**
H8	(HB) Habit -> (FU) Frequency of use	.314	**.000**	.047 +	**Yes**
H9	(BI) Behavioural Intention -> (FU) Frequency of use	.207	**.003**	.019	**Yes**
H10	(PI) Personal Innovation -> (BI) Behavioural Intention	.124	**.013**	.022 +	**Yes**
H11	(CO) Confidence -> (BI) Behavioural Intention	-.028	.395	.002	No
H12	(CO) Confidence -> (FU) Frequency of Use	.026	.816	.000	No
H13	(PR) Perceived Risk -> (BI) Behavioural Intention	.027	.406	.002	No

Note*: (+) $f^2 > .02$ = low effect; (++) $f^2 > 0.15$ medium effect (Sarstedt et al., 2022)*

The values of the relationships between the different constructs are also shown in Table 2, under the Path Coefficients (pc) column. A Path Coefficient closer to +1 indicates a strong positive relationship (as one construct increases, so does the other). A path coefficient closer to -1 indicates a strong negative relationship (as one construct increases, the other decreases). A Path Coefficient of 0 means there is no relationship between constructs.

The analysis of these internal relationships between the model's constructs, which allows identifying its capacity to predict "Behavioural Intention" and "Frequency of Use", suggests that the strongest predictors of "Behavioural Intention" in descending order, are "Habit" (pc = .356; p = .000), "Performance Expectancy" (pc

= .270; p = .000), “Hedonic Motivation” (pc = .166; p = .001), and “Personal Innovation” (pc = .124; p = .013), which together explain 75.7% of the variance in “Behavioural Intention”.

Regarding “Behavioural Intention,” positive effects were also observed for “Effort Expectancy” (pc = .065; p = .074), “Facilitating Conditions” (pc = .030; p = .404), and “Perceived Risk” (pc = .027; p = .406), but the effect size (f^2) of these relationships is not significant (< .02).

Concerning the predictors of “Frequency of Use” the results suggest that the strongest predictor is “Habit” (pc = .314; p = .000), followed by “Behavioural Intention” (pc = .207; p = .003). These constructs account for 29.1% of the variance in “Frequency of Use”. Positive effects were also observed for “Facilitating Conditions” (pc = .072; p = .096) and “Confidence” (pc = .026; p = .011), but again, the effect size (f^2) of these relationships is not significant (< .02).

Thus, of the 13 hypotheses defined for the study, only H1, H6, H7, H8, H9, and H10 are confirmed, as they show statistical significance at the 5% level. The remaining hypotheses are not accepted. Based on the statistical analysis results, the study corroborated six of the thirteen initially proposed hypotheses.

These confirmed hypotheses highlight the influence of Performance Expectancy (PE), Hedonic Motivation (HM), Habit (HB), and Personal Innovation (PI) on Behavioural Intention (BI), as well as the influence of Behavioural Intention (BI) and Habit (HB) on Frequency of Use (FU). Notably, Habit showed a medium influence (pc = 0.356, $p < .001$, $f^2 = 0.180$) and Performance Expectancy (PE) also had a medium effect (pc = 0.270, $p < .001$, $f^2 = 0.114$) on Behavioural Intention ($p < 0.05$).

Conversely, hypotheses H2, H3, H4, H5, H11, H12, and H13 did not receive sufficient statistical support to be accepted in the context of this investigation.
The model presented in the study also considers the moderating effects of “Gender” and “School Level”. The results show that the moderating variable “School Level” had a significant impact on Frequency of Use (pc = .107; p = .015) and influenced the relationship between Facilitating Conditions and Behavioural Intention (pc = .060; p = .031). On the other hand, the moderating variable “Gender” had no significant impact on the tested relationships between the predictors and the dependent variables.

Conclusions

Digital technologies are now inseparable from analogue ones in all aspects of human life, including education. The use of technologies in education has been a growing and widely discussed topic over the past 30 years. Artificial Intelligence, more specifically its generative aspect, represents a further step in this evolution, especially since November 2022, when ChatGPT was made available to the public. Regarding the influence of the various dimensions on Behavioural Intention and Frequency of Use, Habit directly and significantly impacts the former, making it the primary factor influencing the intention to use Generative AI (GenAI). Other important factors include the expectation that GenAI can contribute to better performance, both in terms of time efficiency and success, the enjoyment derived from using these technologies, and the feeling among students that they are learning something new. Simultaneously, both Habit and Behavioural Intention have a positive direct impact on Frequency of Use. In other words, the stronger the habit and the greater the intention to use GenAI, the higher the frequency of its use.

Among the moderating variables, only the effect of the School Level on Frequency of Use and on the relationship between Facilitating Conditions and Behavioural Intention was confirmed. These were the 6 hypotheses supported out of the 13 initially proposed.

The results obtained can thus provide valuable contributions to the understanding of the adoption and use of GenAI in the context of primary and secondary education, as well as support the development of educational strategies that effectively integrate this technology.

References

Adiguzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology*, *15*(3), ep429. https://doi.org/10.30935/cedtech/13152

Al-Azawei, A., & Alowayr, A. (2020). Predicting the intention to use and hedonic motivation for mobile learning: A comparative study in two Middle Eastern countries. *Technology in Society, 62*, 101325. https://doi.org/10.1016/j.techsoc.2020.101325

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50. https://doi.org/10.2307/3151312

George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference* (4th ed.). Pearson Higher Education.

Hair, J. F., Sarstedt, M., Hopkins, L. & Kuppelwieser, V. G. (2014). Partial least squares structural equation modeling (PLS-SEM): An emerging tool in business research. *European Business Review, 26*(2), 106–121.

Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., & Ray, S. (2021). *Partial least squares structural equation modeling (PLS-SEM) using R. Classroom Companion: Business*. Springer International Publishing. https://doi.org/10.1007/978-3-030-80519-7

Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. *New Challenges in International Marketing (Advances in International Marketing, Vol. 20*), 277-319. Emerald Group Publishing Limited.

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial Intelligence in education: Promises and implications for teaching & learning.* Center for Curriculum Redesign.

Liu, B. L., Morales, D., Roser Chinchilla, J. F., Sabzalieva, E, Valentini, A., Vieira do Nascimento, D., & Yerovi, C. (2023). *Harnessing the era of artificial intelligence in higher education: A primer for higher education stakeholders.* UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000386670

Oliveira, A. (2019). *Inteligência Artificial*. Lisboa: Fundação Francisco Manuel dos Santos.

Ringle, C. M., Wende, S., & Becker, J. M. (2022). *SmartPLS 4.* SmartPLS. https://www.smartpls.com/

Sarstedt, M., Ringle, C. M., & Hair, J. F. (2022). Partial least squares structural equation modeling. In C. Homburg, M. Klarmann, & A. Vomberg (Eds.), *Handbook of market research* (pp. 587–632). Springer. https://doi.org/10.1007/978-3-319-57413-4_15

Strzelecki, A. (2023). To use or not to use ChatGPT in higher education? A study of students' acceptance and use of technology. Interactive Learning Environments, 1–14. https://doi.org/10.1080/10494820.2023.2209881

Su, J., & Yang, W. (2023). Unlocking the power of ChatGPT: A framework for applying generative AI in education. *ECNU Review of Education, 6*(3), 355-366. https://doi.org/10.1177/20965311231168423

Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning & Teaching, 6*(1), 31-40. https://doi.org/10.37074/jalt.2023.6.1.17

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments, 10*, 15. https://doi.org/10.1186/s40561-023-00237-x

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*(3), 425–478. https://www.jstor.org/stable/30036540

Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly, 36*(1), 157–178. https://doi.org/10.2307/41410412

Yao, E., Guo, D., Liu, S., & Zhang, J. (2024). The role of technology belief, perceived risk and initial trust in users' acceptance of urban air mobility: An empirical case in China. *Multimodal Transportation, 3*(4), 100169. https://doi.org/10.1016/j.multra.2024.100169

Author Details

Nuno Dorotea
UIDEF, Unidade de Investigação e Desenvolvimento em Educação e Formação.
Institute of Education, University of Lisbon
nmdorotea@ie.ulisboa.pt

Célia Ribeiras
Institute of Education, University of Lisbon
celiaribeiras@edu.ulisboa.pt

This work was supported by National Funds through FCT-Portuguese Foundation for Science and Technology, I.P., under the scope of UIDEF - Unidade de Investigação e Desenvolvimento em Educação e Formação, UIDB/04107/2020, https://doi.org/10.54499/UIDB/04107/2020