Preprints.org

Article

Assessing the Psychological Impact of Generative AI on Data Science Education: An Exploratory Study

Pedro Henrique Ramos Pinto [*] , Vitor Meneghetti Ugulino de Araujo , Cleydson de Souza Ferreira Junior , Lutero Lima Goulart , João Vitor Cardoso Beltrão , Gabriel Silva Aguiar , Samuel José Fernandes Mendes , Filipe de Lima Vaz Monteiro , Erlon Lacerda Avelino

Article

Assessing the Psychological Impact of Generative AI on Data Science Education: An Exploratory Study

Pedro Henrique Ramos Pinto [1,*] 📙, **Vitor Meneghetti Ugulino de Araujo** [2],
Cleydson de Souza Ferreira Junior [3], **Lutero Lima Goulart** [3], **João Vitor Cardoso Beltrão** [3],
Gabriel Silva Aguiar [3], **Samuel José Fernandes Mendes** [3], **Filipe de Lima Vaz Monteiro** [3]
and **Erlon Lacerda Avelino** [3]

[1]   Center for Education, Federal University of Paraíba, Brazil; pedro.clave22@gmail.com (P.P.)
[2]   Department of Informatics, Federal University of Paraíba, Brazil; vitor@ci.ufpb.br (V.A.)
[3]   Center for Informatics, Federal University of Paraíba, Brazil; csfj@academico.ufpb.br (C.J.),
      luterogoulart@gmail.com (L.G.), joaovitorbeltrao@gmail.com (J.B.), filipe.lvm@gmail.com (F.M.),
      gsa3@academico.ufpb.br (G.A.), sjfm@academico.ufpb.br (S.M.) and erlon.avelino@academico.ufpb.br (E.A.)
*   Correspondence: pedro.clave22@gmail.com

Abstract: The integration of AI in education, particularly through Large Language Models (LLMs), is accelerating, reshaping pedagogical approaches and student interactions with new learning tools. This study assesses the impact of generative AI on the educational experiences of data science students at the Center for Informatics, University of Paraíba (CI/UFPB), Brazil. Through the validation of five psychometric scales, we analyze the students' acceptance of LLMs, their associated burnout levels, technology anxiety and the prevalence of metacognitive and dysfunctional learning strategies. Results indicate a significant adoption of AI-driven technologies, with a low incidence of technology anxiety, such as fear of job displacement due to AI. However, a significant correlation between burnout and dysfunctional learning strategies was observed, which could likely be attributed to the rigorous academic environment. Additionally, the employment of metacognitive strategies in conjunction with LLMs reflects an advanced learning approach, yet challenges with functional learning strategies persist. This study contributes to the discourse on AI in education, highlighting the need for educational frameworks that support effective AI adoption while addressing the psychological demands on students.

Keywords: AI for education; ChatGPT; educational technologies; LLMs; data science education; psychological impact of AI; generative AI; validity

---

1. Introduction

In the era of Artificial Intelligence (AI), the academic landscape is undergoing a rapid transformation (Dwivedi et al. 2023; Okonkwo and Ade-Ibijola 2021). AI-powered technologies, such as ChatGPT, now provide nuanced responses to a broad range of topics in mere seconds, aiding undergraduate students in their coursework and learning processes (Cooper 2023; Tu et al. 2023; Zhai 2023). Despite these significant benefits, professors express concerns about the responsible use of such technologies, especially regarding students becoming overly reliant on them, potentially undermining their long-term creative and problem-solving skills (Choi et al. 2023; Hung and Chen 2023; Tu et al. 2023). Additionally, ethical issues such as AI system bias, plagiarism, and lack of transparency need to be considered (Dwivedi et al. 2023; Zhai 2023).

As AI continues to influence the academic landscape, universities worldwide are grappling with emerging challenges and opportunities. For instance, the Chinese University of Hong Kong was among the first Chinese institutions to ban the use of ChatGPT, aiming to uphold academic integrity (Hung and Chen 2023). Students caught using ChatGPT could face penalties ranging from grade reduction to course failure, or even dismissal, on grounds of academic plagiarism and misconduct (Hung and Chen 2023). In contrast, the Hong Kong University of Science and Technology has embraced the use

of ChatGPT and other Large Language Models (LLMs), asserting their responsibility as educators to prepare students for an AI-driven world where tasks can be completed in a timely and cost-effective manner (Hung and Chen 2023).

Popularized by ChatGPT and commonly referred to as generative AI, LLMs are powerful tools with remarkable capabilities for generating human-like text. Tu et al. (2023) emphasizes that LLMs are revolutionizing data science education, necessitating a shift in both curriculum content and pedagogical approaches. The study by these authors revealed that LLMs can also execute all stages of data analysis with just a few command prompts. This fact underscores the potential for students to manipulate conventional exam questions, highlighting the need to adapt assessment practices. However, in alignment with the findings of Cooper (2023); Chiu et. al. (2023); Rahman et al. (2023) and Zhai (2023), the authors also expressed optimism regarding the integration of LLMs into the educational landscape.

Beyond major tech giants such as Google and OpenAI, smaller research groups are increasingly training their own LLMs (Gao and Gao 2023). Stemming from the machine learning branch of AI, these models require vast amounts of training data, which can be easily sourced from the internet. As of 12 PM (GMT -5) on July 18, 2023, there were 15,821 LLMs registered with Hugging Face, a popular machine learning repository (Gao and Gao 2023). The expanding number of LLMs spans various architectures, settings, training methods, and families, reflecting not only the growing presence of these models but also the emergence of diverse types. Each type comes with unique capabilities and applications, marking a varied and rapidly evolving landscape in AI technology (Gao and Gao 2023; Hadi et al. 2023).

Given the advancements in LLMs and the new possibilities they introduce for education, a novel era of teaching and learning is clearly being ushered in. The paradigm shift in educational settings is just the beginning. LLMs are now employed in a range of applications, including virtual assistants, customer service, content creation, and researchers are also benefiting from their support in academic pursuits (Dwivedi et al. 2023; Ramos 2023; Rahman et al. 2023). Several prestigious publishers, including Taylor and Francis, Nature, and Elsevier, have revised their authorship policies to accommodate this new research paradigm. As standard practice, these publishers disallow listing LLMs like ChatGPT as authors, emphasizing a "human-centric" approach by detailing the use of LLMs in the methods section without granting them co-authorship (Dwivedi et al. 2023). Furthermore, researchers remain responsible for the integrity of their academic publications.

1.1. State-of-the-art

As a leading-edge educational tool, AI chatbots offer the unique advantage of scalable, individualized tutoring, providing support that autonomously adapts to the learning pace and style of each student (Lee; Hwang and Chen 2023). Expanding on this concept, Yin et al. (2020) investigated a micro-learning chatbot environment compared to a traditional classroom setting with 91 students over a brief 40-minute intervention. Although the average post-test scores were similar, chatbot users surpassed their peers by half a standard deviation. In a related study, Essel et al. (2022) provided 68 undergraduates with access to an AI assistant, KNUSTbot, and these students also outperformed the control group by half a standard deviation. Corroborating these findings, a systematic review by Okonkwo and Ade-Ibijola (2021) emphasized that AI chatbots, akin to KNUSTbot, have become a mainstay in online education and serve as an effective technological tool to boost students' learning engagement.

However, it's important to distinguish between older chatbots and LLMs. The advent of transformer architecture has profoundly reshaped the Natural Language Processing (NLP) landscape in a relatively short period (Pahune and Chandrasenkharan 2023), and currently, ChatGPT is hailed as the state-of-the-art in conversational technologies (Hadi et al. 2023; Zhai 2023). These AI models have evolved to the degree that they can proficiently tackle standard assessments in law schools (Choi et al. 2023) and even devise intricate solutions for programming paradigms (Tu et al. 2023). Moreover, Cooper (2023) underscores the benefits of ChatGPT for educators and institutions, emphasizing its

prowess in personalizing coursework, tailoring assessments, and meticulously crafting quizzes and science units.

Regrettably, while there is an abundance of reviews and opinion pieces on ChatGPT, there is a dearth of experimental studies evaluating student performance using LLMs as a one-to-one tutoring method. This gap in empirical research may be attributed to the novelty of the subject, suggesting that it might be premature for comprehensive, controlled experiments to have been conducted and published. However, in an encouraging development, Urban et al. (2023) recently released a preprint detailing their research with a control group, focusing on creative problem-solving performance among university students. Their findings suggest that students utilizing ChatGPT demonstrated a capacity to formulate solutions that were more innovative, detailed, and closely aligned with task objectives compared to those not using such advanced tools. This study represents a preliminary but significant step in understanding the potential educational benefits of ChatGPT and similar LLMs in enhancing student learning outcomes.

Delving deeper, two surveys have provided valuable insights into the current state of research on the impact of ChatGPT in educational settings. In their study, Hanum; Hasmayni and Lubis (2023) highlighted the significant positive effect of ChatGPT on students' learning motivation. Leveraging validated psychometric scales for their analysis, they found that approximately 57.3% of the variance in student motivation could be attributed to the use of ChatGPT. Separately, Sallam et al. (2023) introduced a TAM-based survey instrument, the TAME-ChatGPT (Technology Acceptance Model Edited for ChatGPT Adoption), specifically designed to assess the successful integration and application of ChatGPT in healthcare education. Their primary goal was to develop and validate an appropriate psychometric scale for measuring the usage of ChatGPT, facilitating subsequent studies on this construct and related variables such as student anxiety, perceived risk, and behavioral/cognitive factors.

When inferring psychological variables, it is crucial to utilize validated psychometric scales, as such variables cannot be directly observed (Damasio 2012). Therefore, the credibility of a psychological instrument is intrinsically linked to the rigor of its validation process, and a critical aspect of this process is factorial analysis (Goretzko; Pham and Bühner 2021). This statistical method evaluates the common variance among items on a psychometric scale to ensure they measure the same underlying construct (Taherdoost et al. 2022). Given the nascent stage of research on ChatGPT, especially in its application within educational contexts, there is a noticeable gap in the psychometric assessment of its impact. In response to this knowledge deficit, an exploratory study was embarked upon. Applying factorial analysis, the goal was to validate several psychometric scales, ensuring their reliability and consistency for measuring the psychological effects of generative AI on data science education.

2. Materials and Methods

2.1. Objectives

The primary objective of this research is to explore five psychological constructs potentially related to the impact of generative AI on data science education. To address this, psychometric scales specifically designed to measure those core constructs were developed, encompassing AI acceptance, motivation to learn, technology anxiety related to AI, academic burnout, and metacognitive and dysfunctional learning strategies when studying with Large Language Models (LLMs). Consequently, a fundamental part of the research objectives includes validating these tailor-made psychological instruments. This validation process is critical to ensure that the scales accurately capture the intended constructs and are capable of providing meaningful information for understanding and improving the educational process in the context of emerging AI technologies.

2.2. Procedures and Participants

The study was conceptualized as a probabilistic sampling-based survey, leveraging opinion-driven questionnaires (detailed in Appendix A). Data collection was conducted exclusively at the Center for Informatics, University of Paraíba (CI/UFPB), in Brazil, using Google Forms. This platform automatically notified participants of any missing values, ensuring the completeness and accuracy of each submission. To maximize participation and reach, a diverse dissemination strategy was employed. This strategy included using WhatsApp for communication within CI/UFPB academic groups, distributing informative leaflets with QR codes for straightforward access to the online form, and facilitating educational discussions in classroom environments led by faculty members and students. This effort yielded a substantial dataset, with 178 respondents: 143 males, accounting for 80.3% of the total, and 35 females, comprising 19.7%.

Designed as a non-interventional study, the procedures adhered strictly to the highest ethical standards in research involving human subjects. Sociocultural values, participant autonomy, and anonymity were respected throughout the data collection process. All potential risks were carefully measured and mitigated. The deployment of the Informed Consent Statement was integral to this process, ensuring participants' complete awareness of the research's nature and their rights. To maintain data integrity and reduce potential response bias, the psychometric scales were presented to participants in a randomized sequence. This approach is effective in minimizing the influence of question order on responses, thereby providing a more balanced and objective view of the results.

2.3. Instruments

Five psychometric scales were made, each featuring items rated on a 7-point Likert scale to capture the nuances of participants' responses. The scales ranged from "strongly disagree" to "strongly agree" and from "never" to "always". Each scale contributes to a comprehensive understanding of the multifaceted relationship between students and generative AI technologies. The data derived from these instruments are poised to provide a rich foundation for analysis, aiming to elucidate the psychological and educational dynamics at play. Below is a description of each scale employed in the study, outlining their respective domains and the constructs they are intended to measure:

- Academic Burnout Model, 4 items (ABM-4): Initially designed to assess work-related burnout (Kristensen et al. 2005), this scale measures the extent of specific aspects students might feel as a result of prolonged engagement with intense academic activities. These aspects include study-demand exhaustion, the emotional impact of academic pressures, and the depletion experienced from academic endeavors.
- AI Technology Anxiety Scale, 3 items (AITA-3): This scale measures the anxiety students may feel when interacting with generative AI technologies, including fears of job displacement or subject matter obsolescence. The scale was adapted from Wilson et al. (2023, p. 186), which defined technology anxiety as "the tension from the anticipation of a negative outcome related to the use of technology deriving from experiential, behavioral, and physiological elements".
- Intrinsic Motivation Scale, 3 items (IMOV-3): Designed to assess the level of students' intrinsic motivation towards learning, this scale, adapted from (Hanum; Hasmayni and Lubis 2023), evaluates the inherent satisfaction and interest in the learning process itself. This scale is particularly relevant given the literature suggesting a significant relationship between the use of AI technologies and students' motivation to learn.
- Learning Strategies Scale with Large Language Models, 6 items (LS/LLMs-6): This scale encompasses two distinct dimensions involving LLMs, one focusing on Dysfunctional Learning Strategies (DLS/LLMs-3) and the other on Metacognitive Learning Strategies (MLS/LLMs-3), each consisting of three items. The DLS/LLMs-3 sub-scale investigates potential counterproductive learning strategies that students might adopt when using LLMs, which could impede effective learning, and includes both unique and conventional items of dysfunctional

strategies. In contrast, the MLS/LLMs-3 sub-scale assesses the self-regulatory practices that students employ while learning with LLMs, aimed at enhancing learning outcomes. This sub-scale is specifically tailored to include items directly related to LLMs. Based on (Oliveira and Caliatto 2018; Pereira; Santos and Ferraz 2020), these scales provide a comprehensive assessment of learning strategies in the context of LLMs, covering both the metacognitive techniques that enhance learning and the dysfunctional methods that may potentially hinder it.

- LLMs Acceptance Model Scale, 5 items (TAME/LLMs-5): This scale, based on the Technology Acceptance Model (TAM), evaluates data science students' readiness to integrate Large Language Models (LLMs) into their learning process. It has been adapted from Sallam et al. (2023) to provide a more streamlined approach for assessing AI technology acceptance and intention to use.

2.3.1. Adaptation Review

All scales used in the study were specifically adapted to align with the unique characteristics of the participants and the research context, which focused on the impact of generative AI on data science education at CI/UFPB. This institution is renowned for its technologically proficient student body, many of whom are familiar with or actively engaged in technology and programming. Adapting these tools was crucial to accurately capture the nuances of an environment where students are both learning about and working with AI technology, ensuring relevance and sensitivity to the depth of their AI engagement.

Moreover, to enhance participant engagement and ensure response accuracy, the study employed a concise questionnaire design. This brevity was vital to prevent participant fatigue and maintain their interest, enabling the collection of accurate and meaningful data reflective of their experiences and perceptions in a rapidly evolving educational context. The study focused on using established tools with a smaller set of items, ideal for foundational research. By distilling these scales to their key elements, we ensured their psychometric soundness while making them more user-friendly and less burdensome for the respondents.

2.4. Data Analysis

The collected data underwent a meticulous statistical examination, with Exploratory Factorial Analysis (EFA) serving as a pivotal technique for dimensionality reduction. This approach was crucial in simplifying the complex data structure and unveiling the fundamental dimensions within the observed variables, thereby validating the psychological properties of the scales. For descriptive statistics, both box plots and violin plots were utilized to provide visual representations of data distribution and variance. Additionally, Spearman correlation analysis, particularly suited for the data's non-parametric nature, was employed to comprehend the relationships between variables. These combined methods offered a comprehensive understanding of the data's characteristics and interrelations. Python and R were used for the analysis, with the study's code and dataset available in the GitHub Repository. See Appendix B for a summary of the key libraries used and their applications.

2.4.1. EFA Strategic Considerations

Factorial analysis is not a singular technique but rather a group of associated methods that should be considered and applied in concert (Damasio 2012). The objectives of EFA are multifaceted and include the reduction of variables to a smaller number of factors, assessment of multicollinearity, development of theoretical constructs, and testing of proposed theories (Taherdoost et al. 2022). The sequential and linear approach to EFA demands careful consideration of various methodological steps to ensure the validity and reliability of the results. The decisions made throughout this process are detailed below:

1. Sample Adequacy: Prior to factor extraction, it's imperative to evaluate whether the data set is suitable for factor analysis. To this end, Bartlett's test of Sphericity and the Kaiser-Meyer-Olkin

(KMO) measure were employed (Damasio 2012; Taherdoost et al. 2022). Bartlett's test is used to test the hypothesis that the correlation matrix is not an identity matrix, essentially assessing whether the variables are interrelated and suitable for structure detection. A significant result from Bartlett's test allows for the rejection of the null hypothesis, indicating the factorability of our data (Taherdoost et al. 2022). On the other hand, the KMO measure evaluates the proportion of variance among variables that could be attributed to common variance. With the KMO index ranging from 0 to 1, values above 0.5 are considered suitable for factor analysis (Damasio 2012; Taherdoost et al. 2022).

2. Factor Retention: Determining the optimal number of factors to retain is a crucial aspect of EFA, as it defines the dimensionality of the constructs. In this study, three distinct methods were adopted to identify the optimal number of factors: the Kaiser-Guttman criterion (Taherdoost et al. 2022), parallel analysis (Crawford et al. 2010), and the factor forest approach, which involves a pre-trained machine learning model (Goretzko and Bühner 2020). Factor extraction plays a significant role in simplifying data complexity and revealing the dataset's underlying structure, thereby ensuring the dimensions of the constructs measured are captured accurately, and the analysis truly reflects the data's nature.

3. Internal Reliability Assessment: A critical component of EFA is the evaluation of the internal reliability of the scales. Reliability assessment refers to the process of examining how consistently a scale measures a construct. Ensuring consistent measurement is pivotal, as it confirms that any observed variations in data accurately reflect differences in the underlying construct, rather than resulting from measurement error or inconsistencies (Dunn; Baguley and Brunsden 2013). For this purpose, we employed two key metrics: McDonald's omega (ω) and Cronbach's alpha (α). The omega metric provides an estimate of the scales' internal consistency, presenting a robust alternative to the traditionally utilized Cronbach's alpha (Dunn; Baguley and Brunsden 2013).

4. A fundamental next step involves selecting an appropriate method for factor extraction, which dictates how factors are derived from the data. In this research, Principal Axis Factoring (PAF) was employed. Unlike methods that require multivariate normality, PAF is adept at handling data that may not fully meet these criteria, making it a suitable choice in exploratory contexts, especially with smaller sample sizes (Goretzko; Pham and Bühner 2021; Taherdoost et al. 2022). PAF's capability to uncover latent constructs within the data without imposing stringent distributional assumptions aligns well with the exploratory nature of the survey (Goretzko; Pham and Bühner 2021).

5. Rotation Method: The final step in factorial analysis often involves choosing a rotation method to achieve a theoretically coherent and interpretable factor solution. In this study, promax rotation, a widely used method for oblique rotation, was selected Sass and Schmitt (2010). The rationale for using an oblique rotation like promax lies in its suitability for scenarios where factors are presumed to be correlated. Unlike orthogonal rotations, which assume factors are independent, oblique rotations acknowledge and accommodate the possibility of inter-factor correlations (Damasio 2012; Sass and Schmitt 2010).

2.5. Use of AI Tools

The ChatGPT-4 model played a substantial role in this project, being utilized not only for correcting grammar but also for refining paragraphs, and assisting with coding in data analysis. Additionally, a specialized model, SciChat, was developed, specifically designed to assist in enhancing the writing process for this research. This model was custom-designed to respond to queries related to scientific writing, providing more focused and effective support. The diverse application of these AI tools ensured the maintenance of high scientific accuracy and integrity throughout the project, with every output rigorously checked for precision and reliability.

3. Results

3.1. LLMs Usage Preferences Among Students

In the assessment of LLMs' preference at the CI/UFPB, the data revealed a predominant usage of ChatGPT 3.5, with a staggering 92.7% of the respondents utilizing this free version. ChatGPT 4, despite being a paid version, is used by 5.6% of the participants, showcasing a willingness to invest in more advanced AI tools. Bing Chat, another LLM, is used by 23% of the students, indicating a diversity in the AI platforms engaged by the students for their educational pursuits. A smaller, yet significant, fraction of students have adopted Bard, comprising 18% of the users, and only a minority, 4.5%, reported not using any LLMs at all, which underscores the widespread penetration of these technologies in the academic environment.

3.2. Validity Assessment Through Factorial Analysis

The results of the EFA revealed diverse psychometric properties across the five validated scales. Primarily, both Bartlett's test and the KMO measure affirmed the suitability of the data for factor analysis. As indicated in Table 1, the KMO values for all scales exceeded the benchmark of 0.5, suggesting the sample adequacy for each scale (Taherdoost et al. 2022). Moreover, the significance of Bartlett's tests across the board (p-value < 0.05) rejected the null hypothesis that our correlation matrix is an identity matrix, underscoring the data's aptness for structure detection (Damasio 2012).

Table 1. Sample adequacy.

Scales	KMO Test	Bartlett's Test of Sphericity
ABM-4	0.742	0.000
AITA-3	0.562	0.000
IMOV-3	0.581	0.000
LS/LLMs-6	0.682	0.000
TAME/LLMs-5	0.689	0.000

As detailed in Table 2, the factorial retention procedure further elucidates the dimensionality of the scales, reinforcing the diverse psychometric properties observed during EFA. Remarkably, the ABM-4, AITA-3, and IMOV-3 scales each indicated a singular factorial structure, as evidenced by uniform retention values across the chosen methods: Kaiser criterion, parallel analysis, and factor forest algorithm. Conversely, the LS/LLMs-6 and TAME/LLMs-5 scales exhibited a more complex structure. The LS/LLMs-6 scale steadily retained two factors across all three methods, indicating a clear factorial structure, and the TAME/LLMs-5 scale presented a variation between one and two factors. Hence, the decision to retain only one factor for this scale was a strategic response to the inherent limitations of parallel analysis in processing categorical and rank-ordered data. Recognizing these technical limitations, the study shifted its focus to the factor forest method (Goretzko and Bühner 2020). This approach, weighting multiple state-of-the-art machine learning algorithms, proved more appropriate for the TAME/LLMs-5 scale, offering nuanced insights otherwise concealed in traditional analyses.

Table 2. Factorial retention.

Scales	Kaiser criterion	Parallel analysis	Factor Forest
ABM-4	1	1	1
AITA-3	1	1	1
IMOV-3	1	1	1
LS/LLMs-6	2	2	2
TAME/LLMs-5	1	2	1

After evaluating sample adequacy and determining the number of factors to retain, the study progressed to assess the internal reliability of the scales. This evaluation, detailed in Table 3, involved an analysis of Cronbach's alpha (α), McDonald's omega (ω), and the cumulative variance explained for each scale (Damasio 2012; Dunn; Baguley and Brunsden 2013). These analyses were crucial for understanding the scales' consistency and reliability. Notably, the ABM-4 and LS/LLMs-6 scales displayed strong reliability, with both ω values surpassing the 0.7 benchmark, suggesting high consistency. The TAME/LLMs-5 scale showed a well-balanced reliability profile, evidenced by closely matched α and ω values. In contrast, the AITA-3 and IMOV-3 scales, while still reliable, recorded slightly lower scores, indicating moderate consistency.

Importantly, all the scales exhibited acceptable levels of cumulative variance explained. This measure indicates the proportion of total variance in the observed variables that is accounted for by the factors. Higher values suggest that the factors extracted during the EFA process are effectively capturing the underlying structure of the dataset (Goretzko; Pham and Bühner 2021). In the context of this study, the cumulative variance explained by each scale, although varying, was within an acceptable range. This suggests that the scales are adequately capturing the constructs they are intended to measure, thereby supporting their validity. For instance, even though the TAME/LLMs-5 scale had the lowest cumulative variance explained (0.426), it still provided a significant portion of the variance, contributing to a meaningful understanding of the underlying construct.

Table 3. Internal Reliability Assessment.

Scales	Cronbachs' alpha (α)	McDonalds' omega (ω)	Cumulative variance explained
ABM-4	0.702	0.702	0.528
AITA-3	0.619	0.664	0.579
IMOV-3	0.578	0.625	0.545
LS/LLMs-6	0.640	0.739	0.618
TAME/LLMs-5	0.656	0.663	0.426

At last, the cornerstone of EFA is the factorial extraction procedure. The factorial loadings in EFA are critical as they represent the strength and direction of the relationship between observed variables (items) and underlying latent factors (Damasio 2012; Taherdoost et al. 2022). Essentially, these loadings measure how much variance in an item is explained by the factor, providing insights into how well each item aligns with a particular factor. Higher loadings indicate a stronger association, thus contributing significantly to the interpretation and validation of the factor structure. Analyzing the EFA loadings presented in Table 4, it's evident that most items demonstrate strong correlations with their respective factors, indicating a coherent pattern associations across the scales, reinforcing the construct's validity and the scale's reliability in measuring the intended constructs.

Precisely, in the ABM-4 scale, each item exhibits substantial loadings, ranging from 0.70 to 0.74 on Factor 1, indicating a strong and consistent relationship with this factor. The AITA-3 scale's items also show a solid association with Factor 1, though one item scores lower, at 0.53. The IMOV-3 scale has loadings varying from 0.67 to 0.82, suggesting a moderately strong relationship with the factor. The TAME/LLMs-5 scale displays varied loadings, with Item 5 having a lower loading of 0.47, indicating a weaker association with the underlying factor compared to the other items in the scale. The LS/LLMs-6 scale presents a more complex scenario. Items 1, 2, and 3 have high loadings, all above 0.83, on Factor 1, while Items 4, 5, and 6 show significant loadings, ranging from 0.68 to 0.76, on Factor 2, suggesting that this scale is adept at capturing two distinct yet related constructs. This intricate pattern of factor loadings across various scales underscores the multifaceted nature of the constructs being measured, and the effectiveness of the factorial analysis in simplifying the data's complexity into a small number of psychological constructs.

Table 4. Factor loadings.

Scales	Items	Factor 1	Factor 2
ABM-4	Item 1	0.70	-
	Item 2	0.74	-
	Item 3	0.72	-
	Item 4	0.74	-
AITA-3	Item 1	0.53	-
	Item 2	0.85	-
	Item 3	0.86	-
LS/LLMs-6	Item 1	**0.83**	0.00
	Item 2	**0.85**	0.05
	Item 3	**0.86**	0.04
	Item 4	0.11	**0.68**
	Item 5	0.02	**0.71**
	Item 6	0.09	**0.76**
IMOV-3	Item 1	0.72	-
	Item 2	0.82	-
	Item 3	0.67	-
TAME/LLMs-5	Item 1	0.69	-
	Item 2	0.69	-
	Item 3	0.63	-
	Item 4	0.74	-
	Item 5	0.47	-

3.3. Psychometric Scales' Descriptive Statistics

The exploration into the scales' characteristics will be visualized through the strategic use of box and violin plots, which will illustrate the core tendencies and variations within the students' data. Box plots are designed to highlight central measures — the mean (indicated by a white dot) and median (depicted as a red line) — along with the spread of responses. This spread is represented by quartiles, which divide the data into four equal parts. The interquartile range, marking the distance between the first and third quartiles, encompasses the central 50% of the data, shown as a blue box. In parallel, violin plots will provide insights into the data's density and distribution, offering a more nuanced interpretation of variability and frequency across different values. This dual approach allows for a comprehensive understanding of the dataset, from central tendencies to the diversity of responses.

Beginning with an assessment of students' mental health, the ABM-4 scale uncovers significant patterns in students' exhaustion, stress and depletion (Figure 1). The mean and median, positioned above the midpoint scale of 16, suggest a high level of academic burnout among the students. The box plot displays a wide interquartile range, indicating a diverse dispersion in the severity of students' experiences. Additionally, the violin plot's density is prominently stretched, reinforcing the spectrum of responses, from the median to more extreme reports. In contrast, the AITA-3 scale, which focuses on technology anxiety related to AI, presents a different data distribution (Figure 2). The box plot demonstrates that the interquartile range is below the mid-scale value of 12, indicating a trend toward lower anxiety levels within the group. This observation is further supported by the violin plot, where the data concentration is heaviest at the lower end of the scale. It reveals that 75% of students' anxiety levels are comfortably below the midpoint, and none have reached the maximum level of technology anxiety.
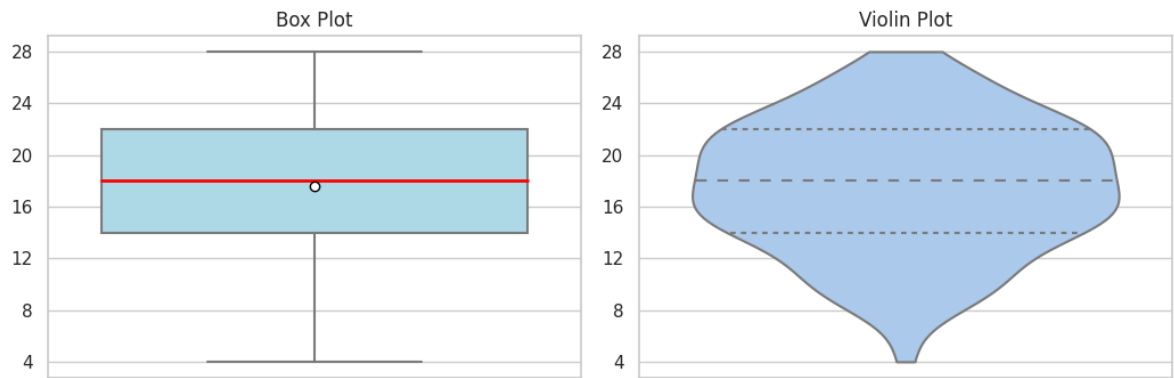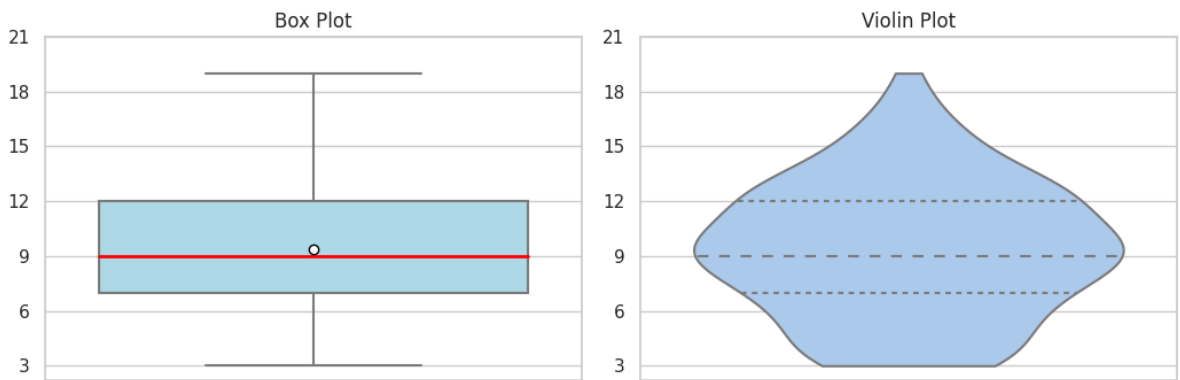
Figure 1. ABM-4.

Figure 2. AITA-3.

Regarding learning strategies, the analysis of the MLS/LLMs-3 sub-scale through its descriptive statistics unveils a median that is marginally above the midpoint, indicating a propensity for higher engagement with metacognitive strategies (Figure 3). Nevertheless, the wide interquartile range indicates a variety in students' responses. The corresponding violin plot supports this conclusion, displaying a response density that increases towards the ends. Furthermore, the DLS/LLMs-3 visualizations shed light on another dimension of student learning strategies (Figure 4). When comparing these sub-scales, it appears that students are generally more consistent in their use of metacognitive strategies than in avoiding dysfunctional ones, which seem to be more scattered. This is reinforced by the violin plot's bimodal peaks, which imply two main clusters of responses among the participants.
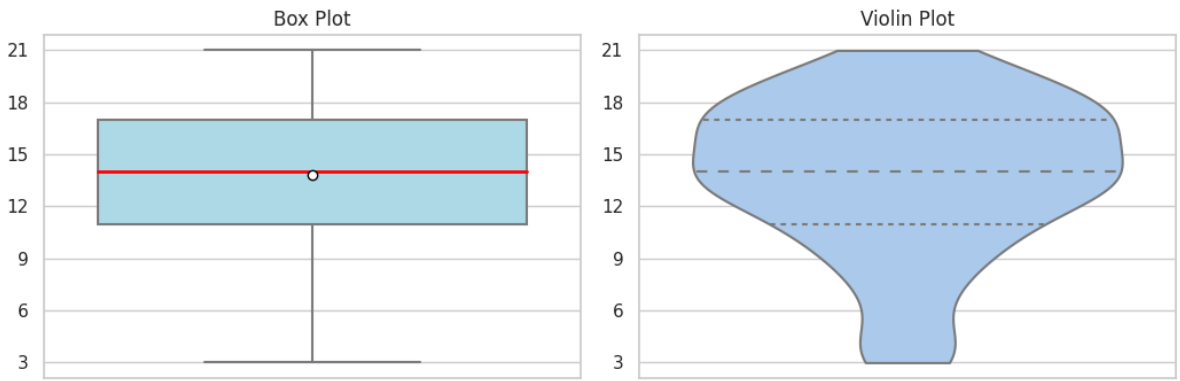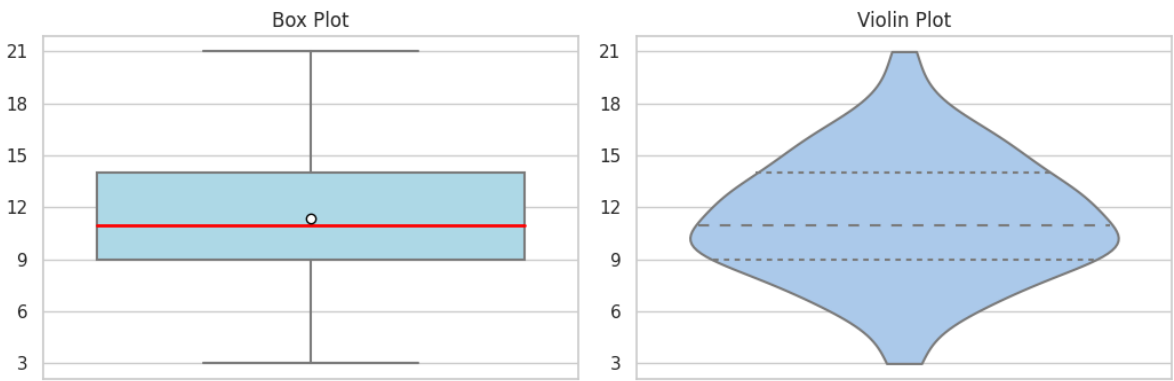

Figure 3. MLS/LLMs-3.

Figure 4. DLS/LLMs-3.

Delving deeper, IMOV-3 provides an intriguing overview of the distribution of students' inherent enthusiasm for learning (Figure 5). The median value of 15, with the entire interquartile range situated above the midpoint of 12, reflects a collective tendency toward a more motivated approach to learning. An outlier, represented as a solitary dot, hints at an exceptional case where a student's motivation significantly diverges from the norm. This data leads us to a plausible conclusion that students demonstrate strong motivation for learning, evidenced by 75% of them scoring beyond the midpoint threshold. Lastly, TAME/LLMs-6 provides insights into students' perceptions and acceptance of LLMs (Figure 6). The box plot reveals a small interquartile range positioned above the scale's midpoint, indicating a cohesive attitude toward LLMs among the respondents. This compact range suggests a consensus in acceptance levels, with only a few outliers indicating some reservations. Mirroring this, the violin plot's expanded middle section reflects a majority consensus, which narrows at both ends to represent fewer students with extreme viewpoints, be they highly skeptical or exceptionally receptive to LLMs.
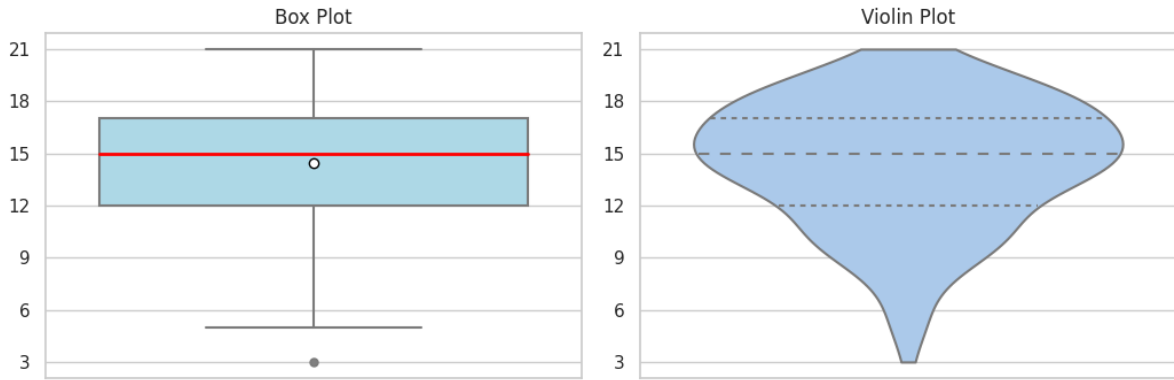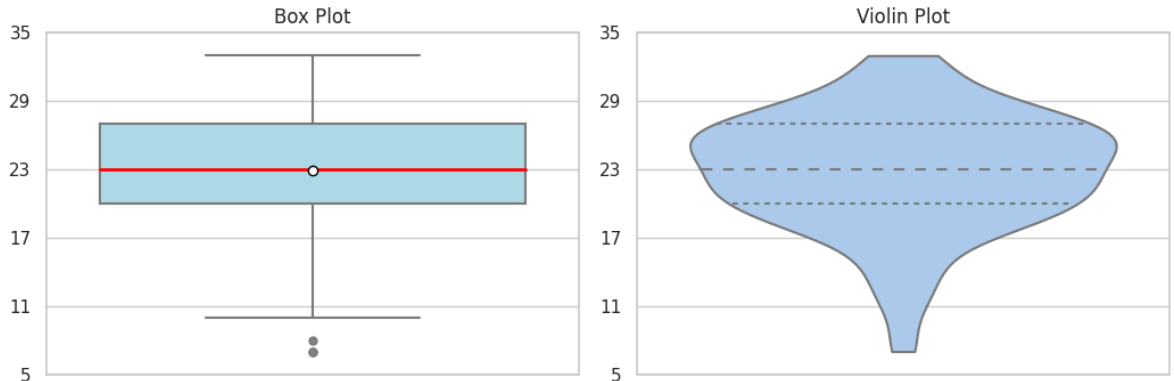

Figure 5. IMOV-3.

Figure 6. TAME/LLMs-6.

3.4. Spearman's Correlation for Categorical Variables

The Spearman's correlation matrix, a sophisticated analytical tool for mapping out rank-order correlations, unveils a complex network of relationships among the scales described before (Tables 5 and 6). A pattern emerges where academic burnout subtly correlates with technology anxiety and more notably, with dysfunctional learning strategies. This triad forms a nexus, indicating a dynamic interplay where each construct potentially exacerbates the influence of the others. Yet, the most pronounced relationship in the matrix is observed between the acceptance of LLMs and the metacognitive strategies employed by students, strongly suggesting that embracing these advanced tools could be aligned with more sophisticated and reflective learning approaches. However, the observed positive correlation between the acceptance of LLMs, academic burnout, and technology anxiety introduces an unexpected complexity into the equation.

Table 5. Spearman's correlation matrix (ρ).

Scales	ABM-4	AITA-3	IMOV-3	MLS/LLMs-6	DLS/LLMs-6	TAME/LLMs-6
ABM-4	1	**0.27**	-0.14	0.05	**0.41**	**0.16**
AITA-3	**0.27**	1	0.00	0.09	**0.34**	**0.16**
IMOV-3	-0.14	0.00	1	-0.03	**-0.31**	-0.04
MLS/LLMs-6	0.05	0.09	-0.03	1	0.11	**0.60**
DLS/LLMs-6	**0.41**	**0.34**	**-0.31**	0.11	1	0.13
TAME/LLMs-5	**0.16**	**0.16**	-0.04	**0.60**	0.13	1

Table 6. Significance matrix (p-values).

Scales	ABM-4	AITA-3	IMOV-3	MLS/LLMs-6	DLS/LLMs-6	TAME/LLMs-6
ABM-4	-	**0.00**	0.06	0.46	**0.00**	**0.04**
AITA-3	**0.00**	-	0.99	0.23	**0.00**	**0.03**
IMOV-3	0.06	0.99	-	0.64	**0.00**	0.60
MLS/LLMs-6	0.46	0.23	0.64	-	0.13	**0.00**
DLS/LLMs-6	**0.00**	**0.00**	**0.00**	0.13	-	0.07
TAME/LLMs-5	**0.04**	**0.03**	0.60	**0.00**	0.07	-

4. Discussion

4.1. An Overwhelming Acceptance of LLMs

The findings unveil nuanced psychological and behavioral patterns among data science students, particularly their overwhelming acceptance of LLMs like ChatGPT and Bard. This trend is indicative of a progressive mindset toward the integration of AI technologies into their academic toolkit. With only eight students reporting no use of LLMs, and considering the descriptive results of the TAME/LLMs-6 scale, the data underscores a pervasive, technology-oriented ethos at CI/UFPB. As documented in recent studies (Cooper 2023; Choi et al. 2023; Dwivedi et al. 2023; Tu et al. 2023; Zhai 2023), the influence of LLMs is reshaping educational practices across institutions worldwide. Notably, these AI technologies are not merely transient tools but are becoming integral to the future of teaching and learning, with their impact evolving more rapidly in fields like data science that are inherently intertwined with technological advancements.

4.2. The Role of AI in Data Science Education

This technology-friendly environment likely contributes to the notably low levels of AI-related anxiety, as observed in the descriptive results of the AITA-3 scale. The students' frequent interactions with advanced technological tools seem to buffer them against the typical apprehensions concerning new technological integrations. Rather than viewing AI as a threat to their skills or future job prospects, they appear to recognize its potential to enhance their capabilities.

However, a minor positive correlation of 0.16 (p-value = 0.03) was observed between the acceptance of LLMs and technology anxiety. This finding somewhat diverges from Wilson et al. (2023), which suggested that anxiety regarding technology typically has a negative correlation with the acceptance, usage, and integration of such tools. Nevertheless, it is essential to distinguish that the ATAS scale from Wilson et al. (2023) is concerned with general technology anxiety and not crafted for the evolving AI context, whereas the AITA-3 is dedicated to exploring the societal issues triggered by those technologies. This distinction implies that for students regularly using LLMs, the perceived effectiveness of AI might paradoxically induce more anxiety about its societal integration, highlighting a nuanced relationship between familiarity with AI and perceptions of its broader implications.

The concern about the displacement of programming jobs by AI mirrors a broader expectation of profound changes across various sectors. This anticipated shift underscores the critical need for strategic upskilling in education, encouraging programmers to expand their expertise beyond the conventional pipeline. Now, "students need to learn to view themselves as product managers rather than software engineers" (Tu et al. 2023, p. 3), which not only prepares them for the evolving demands of the job market but also positions them to navigate the future of work with agility and foresight.

It is known that the emergence of digital technologies, such as calculators, smartphones, and GPS systems, has profoundly impacted human cognition (Dwivedi et al. 2023). Similarly, AI is poised to bring about significant psychological changes, but the specifics of these changes remain largely unknown. These evolving cognitive landscapes, influenced by AI's unique interactions and capabilities, underscore the need for new forms of literacy and adaptability in the 21st century, which goes beyond traditional digital navigation skills. A crucial aspect of effective AI interaction is the skill to craft precise prompts, a capability that varies among individuals; some find it easier to formulate than others (Dwivedi et al. 2023; Shanahan 2022). As AI becomes increasingly integral in various aspects of life, the skill of prompt formulation should be recognized and developed with the same emphasis as the overall digital literacy.

4.3. Metacognition in Modern Learning Environments

The integration of LLMs with metacognitive strategies, which refer to the conscious control over cognitive processes involved in learning such as organizing, prioritizing, and actively monitoring one's comprehension and progress, indicates a sophisticated approach to learning (Oliveira and Caliatto 2018; Pereira; Santos and Ferraz 2020). According to the MLS/LLMs-6 descriptive results, students are not merely relying on LLMs; rather, they are thoughtfully incorporating them into their study habits, utilizing their capabilities to enhance understanding and refine problem-solving skills. This strategic application likely contributes to the favorable perception of LLMs, as demonstrated by the strong positive correlation between the TAME/LLMs-6 and the MLS/LLMs-3 ($\rho = 0.60$, $p < 0.01$). This indicates that increased acceptance of AI corresponds with heightened metacognitive engagement, suggesting a deliberate and resourceful management of their learning processes by the students.

In contrast, the analysis of the DLS/LLMs-3 sub-scale reveals that despite general confidence in using LLMs, students recognize certain challenges. Difficulties in identifying inaccuracies in outputs from LLMs highlight the complexities of relying on AI for learning. This complexity is further elucidated by the moderate negative correlation between the IMOV-3 and the DLS/LLMs-6 ($\rho = -0.31$, $p < 0.01$), indicating that students employing fewer dysfunctional learning strategies tend to be more intrinsically motivated. The DLS/LLMs-3 also reflects conventional dysfunctional strategies, which include lack of self-regulation and subject understanding. This reinforces the necessity of teachers' support in optimizing students' use of AI for educational purposes, corroborating Chiu et. al. (2023)'s findings which indicate that both student expertise and teacher assistance are crucial for effectively fostering intrinsic motivation and learning competence with AI-based chatbots.

In a broader context, the meta-analysis by Theobald (2021) highlights the intricate relationship between various factors in academic settings, emphasizing the positive impacts of cooperative learning on cognitive and metacognitive strategies. The analysis suggests that programs centered around

feedback more effectively enhance metacognitive skills, resource management, and motivation. Notably, programs grounded in a metacognitive theoretical framework achieve greater success in academic achievement compared to those that focus solely on cognitive aspects. This insight is particularly relevant in the context of AI-driven technologies, such as chatbots, which excel in offering personalized, immediate feedback. Studies by Chiu et. al. (2023); Okonkwo and Ade-Ibijola (2021); Yin et al. (2020) have demonstrated the effectiveness of AI technologies in these areas. They argue that such technologies, by providing personalized feedback, can significantly aid in the development of self-regulated learning skills and autonomy.

Metacognitive strategies encompass the processes of planning, monitoring, and evaluating one's understanding and learning. These higher-order cognitive processes involve self-regulation and control over learning activities (Oliveira and Caliatto 2018; Pereira; Santos and Ferraz 2020). The dichotomy between LLMs' usage in metacognitive versus dysfunctional strategies highlights the necessity for a balanced integration of LLMs in educational settings. It underscores the importance of equipping students with the skills to critically evaluate and effectively employ LLMs outputs, enabling them to discern and rectify errors, thus optimizing their learning journey and outcomes. As suggested by recent studies (Cooper 2023; Tu et al. 2023; Urban et al. 2023; Zhai 2023), strategic LLMs use, supported by a thorough understanding of their limitations, can empower students to navigate potential pitfalls and maximize the benefits of these advanced tools.

4.4. LLMs and Mental Health Among Students

The high scores on academic burnout highlight a critical aspect of student life, emphasizing how the pressures for academic achievement and the competitive nature of academia significantly contribute to elevated stress levels and impact students' overall educational experiences. An important aspect to consider is the interaction between this widespread stress and students' use of Large Language Models (LLMs). Research findings reveal a positive and moderate correlation between ABM-4 and DLS/LLMs-3 ($\rho = 0.41$, p-value = 0.00). This correlation suggests that higher levels of academic burnout are associated with an increased reliance on ineffective learning strategies involving LLMs. Additionally, a small but significant correlation exists between ABM-4 and AITA-3 ($\rho = 0.27$, p-value = 0.00), indicating a relationship between academic stressors and students' apprehensions regarding AI's role in society.

In the context of the scholarly consensus highlighted by Mofatteh (2020), which elucidates the impact of psychological and academic variables on stress and anxiety levels in university students, the correlations identified in the referenced study acquire considerable significance. These correlations, when placed within a broader context, contribute to a nuanced understanding of student experiences. Recognized factors such as low self-esteem, personality traits like high neuroticism and low extraversion, and feelings of loneliness are known to increase susceptibility to stress, anxiety, and depression during university years. The findings from the study in question add an additional layer to this complex dynamic, which is particularly relevant in the current educational landscape where technology and digital tools have become increasingly integral to the learning process.

5. Conclusion

In summary, the study sheds light on the complex dynamics between students' engagement with LLMs and their psychological and academic well-being. The findings reveal a dichotomy: on one side, there is a discernible trend of acceptance and strategic utilization of LLMs among data science students, indicating a positive shift towards the integration of AI in educational paradigms. On the other side, the study also unveils a nuanced interplay between the use of these advanced tools and academic stressors. The identified correlations between academic burnout, dysfunctional learning strategies, and AI-related anxiety highlight the necessity for educational institutions to cultivate not just digital literacy but also a supportive environment for student development.

5.1. Limitations and Prospects for Future Research

This study's insights must be contextualized within the scope of its methodological constraints. The relatively small sample size may limit the generalizability of our findings to broader populations. Additionally, the brevity of the psychometric scales, while necessary to cover a range of constructs without overburdening respondents, may yield less granularity compared to more extensive conventional scales. Such constraints could potentially affect the accuracy and depth of our insights into the multifaceted impacts of generative AI on education. Moreover the scales were not reviewed by professionals specializing in the constructs being measured, which may have provided further validation of the instruments used.

Future research endeavors could aim to mitigate these limitations by employing larger and more diverse samples to enhance the representativeness of the results. Further refinement and validation of the scales by subject matter experts would bolster the reliability of the measurements. An exploration of correlations between scale results and sociodemographic variables could yield rich insights; for instance, comparing the metacognitive and disfunctional learning strategies of newer students with those more advanced in their studies could reveal how adaptability and coping mechanisms evolve through the university experience. Such investigations could offer valuable information on the developmental trajectory of learning strategies in relation to the integration of AI in academic settings.

Author Contributions: Conceptualization, Araújo, V. M. U., Beltrão, J. V. C., Aguiar, G. S., Ferreira Junior, C. S., Avelino, E. L. and Mendes, S. J. F.; methodology, Araújo, V. M. U., Ramos, P. H. R., Ferreira Junior, C. S. and Mendes, S. J. F.; validation, Ramos, P. H. R.; formal analysis, Ramos, P. H. R.; investigation, Mendes, S. J. F., Monteiro, F. L. V.; Aguiar, G. S., Ramos, P. H. R. and Beltrão, J. V. C.; resources, Ramos, P. H. R., Monteiro, F. L. V. and Goulart, L. L.; data curation, Ramos, P. H. R., Goulart, L. L.; writing—original draft preparation, Ramos, P. H. R.; writing—review and editing, Araújo, V. M. U. and Ferreira Junior, C. S.; visualization, Ramos, P. H. R.; supervision, Araújo, V. M. U., Ramos, P. H. R. and Ferreira Junior, C. S.; project administration, Araújo, V. M. U.; All authors have read and agreed to the published version of the manuscript.

Abbreviations

The following abbreviations are used in this manuscript:

ABM-4	Academic Burnout Model Scale - 4 items.
AI	Artificial Intelligence.
AITA-3	Artificial Intelligence Technology Anxiety Scale - 3 items.
CI/UFPB	Center for Informatics, Federal University of Paraíba.
DLS-3	Dysfunctional Learning Strategies Sub-scale - 3 items.
EFA	Exploratory Factorial Analysis.
IMOV-3	Intrinsic Motivation Scale - 3 items.
LLMs	Large Language Models.
LS/LLMs-6	Learning Strategies Scale with LLMs - 6 items.
MLS-3	Metacognitive Learning Strategies Sub-scale - 3 items.
TAME/LLMs-5	Technology Acceptance Model for LLMs - 5 items.

Appendix A. Questionnaire item's transcription

Appendix A.1. ABM-4

1. Original item: *Nunca me sinto capaz de alcançar meus objetivos acadêmicos.*

Suggested translation: I never feel able to achieve my academic goals.

2. Original item: *Tenho dificuldade para relaxar depois das aulas.*

    Suggested translation: I have trouble relaxing after school.

3. Original item: *Fico esgotado quando tenho que ir à universidade.*

    Suggested translation: I get exhausted when I have to go to college.

4. Original item: *As demandas do meu curso me deixam emocionalmente cansado(a).*

    Suggested translation: The demands of my course make me emotionally tired.

Appendix A.2. AITA-3

1. Original item: *Sinto como se não pudesse acompanhar as mudanças causadas pelos modelos de inteligência artificial.*

    Suggested translation: I feel like I can't keep up with the changes caused by artificial intelligence models.

2. Original item: *Me preocupo que programadores sejam substituídos pelos modelos de inteligência artificial.*

    Suggested translation: I worry that programmers will be replaced by artificial intelligence models.

3. Original item: *Tenho medo que modelos de inteligência artificial tornem conteúdos que aprendi na faculdade obsoletos.*

    Suggested translation: I'm afraid artificial intelligence models will render content I learned in college obsolete.

Appendix A.3. IMOV-3

1. Original item: *Gosto de ir em todos as aulas do meu curso.*

    Suggested translation: I like to go to every class in my class.

2. Original item: *Muitas vezes, fico tão empolgado que perco a noção do tempo quando estou envolvido em um projeto ou atividade acadêmica.*

    Suggested translation: I'm often so excited that I lose track of time when I'm involved in a project or academic activity.

3. Original item: *Para mim, aprender sobre programação é um interesse pessoal.*

    Suggested translation: For me, learning about programming is a personal interest.

Appendix A.4. LS/LLMs-6

1. Original item: *Utilizo LLMs (ChatGPT, Bard etc.) para tirar dúvidas e preencher lacunas no meu conhecimento sobre programação.*

    Suggested translation: I use LLMs (ChatGPT, Bard, etc.) to clarify doubts and fill gaps in my knowledge about programming.

2. Original item: *Utilizo LLMs (ChatGPT, Bard etc.) para formular e resolver atividades de programação.*

    Suggested translation: I use LLMs (ChatGPT, Bard, etc.) to formulate and solve programming activities.

3. Original item: *Corrijo meus códigos utilizando LLMs (ChatGPT, Bard etc.).*

    Suggested translation: I correct my codes using LLMs (ChatGPT, Bard etc.).

4. Original item: *Deixo para estudar para as provas de última hora.*

    Suggested translation: I leave it to study for last-minute exams.

5. Original item: *Tenho dificuldade para encontrar erros em respostas e códigos gerados por LLMs (ChatGPT, Bard etc.).*

    Suggested translation: I have difficulty finding errors in responses and codes generated by LLMs (ChatGPT, Bard etc.).

6. Original item: *Sinto que estou apenas memorizando informações em vez de realmente entender os conteúdos.*

Suggested translation: I feel like I'm just memorizing information instead of really understanding the contents.

Appendix A.5. TAME/LLMs-5

1. Original item: *LLMs (ChatGPT, Bard etc.) tornam a programação mais democrática e acessível para as pessoas.*

Suggested translation: LLMs (ChatGPT, Bard etc.) make programming more democratic and accessible for people.

2. Original item: *Acredito que LLMs (ChatGPT, Bard etc.) podem ser melhor exploradas pelos professores nas aulas, atividades e/ou provas.*

Suggested translation: I believe that LLMs (ChatGPT, Bard etc.) can be better explored by teachers in classes, activities and/or tests.

3. Original item: *Me sinto confiante com os textos e/ou códigos gerados por LLMs (ChatGPT, Bard etc.).*

Suggested translation: I feel confident with the texts and/or codes generated by LLMs (ChatGPT, Bard etc.).

4. Original item: *Penso que LLMs (ChatGPT, Bard etc.) são muito eficientes em programação.*

Suggested translation: I think LLMs (ChatGPT, Bard etc.) are very efficient in programming.

5. Original item: *Prefiro programar sem ajuda de LLMs (ChatGPT, Bard etc.).*

Suggested translation: I prefer programming without the help of LLMs.

Appendix B. Libraries used for data analysis

- Pandas: As the backbone of our data manipulation efforts, Pandas provided an efficient platform for organizing, cleaning, and transforming the collected survey data into a structured format suitable for analysis.
- Pingouin: This library was instrumental in performing advanced statistical calculations. It allowed us to conduct Spearman's correlation analysis, offering insights into the relationships between various psychometric scales.
- NumPy: Essential for handling complex linear algebra operations, NumPy supported our data processing needs, especially in managing multi-dimensional arrays and matrices.
- Matplotlib and Seaborn: These visualization tools were pivotal in creating intuitive and informative graphs. Matplotlib, with its flexible plotting capabilities, and Seaborn, which extends Matplotlib's functionality with advanced visualizations, together provided a comprehensive view of our data's distribution and trends.
- RPy2: Serving as a bridge between Python and R, RPy2 enabled us to leverage R's robust statistical and graphical capabilities within our Python-centric analysis pipeline.
- FactorAnalyzer: This library facilitated the Exploratory Factor Analysis (EFA), allowing us to uncover the underlying structure of our psychometric scales and validate their dimensionalities.
- Scipy: We utilized its statistical functions, particularly Spearman's rank correlation, to explore the associations between different variables, enhancing the robustness of our findings.
- R Libraries (mlr, psych, ineq, BBmisc, ddpcr, MVN, GPArotation): These R packages, integrated into our analysis via RPy2, offered specialized functions ranging from machine learning (mlr) and psychological statistics (psych) to multivariate normality testing (MVN) and factor rotation methods (GPArotation).

References

Cooper, G. Examining Science Education in ChatGPT: AN Exploratory Study of Generative Artificial Intelligence. **2023**. *Journal of Science Education and Technology*; *32*, pp. 444–452. Available online: DOI: 10.1007/s10956-023-10039-y

Chiu, T. K. F.; Moorhouse, B. L.; Chai, C. S.; and Ismailov, M. Teacher support and student motivation to learn with Artificial Intelligence (AI) based chatbot. **2023**. *Interactive Learning Environments*. DOI: 10.1080/10494820.2023.2172044

Choi, J. H.; Hickman, K. E.; Monahan, A., and Schwarcs, D. B. ChatGPT Goes to Law School. **2023**. *71 Journal of Legal Education 387 (2022), submitted*; pp. 1–16. DOI: 10.2139/ssrn.4335905

Crawford, A. V.; Green, S. B., Levy, R.; Lo, W.; Scott, L.; Svetina, D.; and Thompson, M. S. Evaluation of parallel analysis methods for determinning the number of factors. *Educational and Psychological Measurement*, **2010**; *79(6)*, pp. 885–01. Available online: https://doi.org/10.1177/0013164410379332

Damasio, B. Uses of Exploratory Factorial Analysis in Psychology. **2012**. *Avaliação Psicológica*; *11(2)*, pp. 213–228. DOI: S1677-04712012000200007

Dunn, T. J.; Baguley, T.; and Brunsden, V. From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. **2013**. *British Journal of Psychology*; *105(3)*, pp. 399–412. DOI: 10.1111/bjop.12046

Dwivedi, Y. K.; Kshetri, N.; Huges, L.; Slade, E. L., Jeyaraj, A.; Kar, A. K.; Baabdullah, A. M.,; Koohang, A.; Raghavan, V.; Ahuja, M.; Albanna, H.; Albashrawi, M. A.; Al-Busaidi, A.; Balakrishnan, J.; Barlette, Y.; Basu, S.; Bose, I.; (...) and Wright, R. **2023**. Opinion Paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*; *71*. DOI: 10.1016/j.ijinfomgt.2023.102642

Essel, H. B.; Vlachopoulos, D.; Tachie-Menson, A.; Johnson, E. E.; and Baah, P. K. **2022**. The impact of a virtual teaching assistant (chatbot) on students' learning in Ghanaian higher education. *International Journal of Educational Technology in Higher Education*; *19(1)*, pp. 1–19. DOI: 10.1186/s41239-022-00362-6

Gao, S., and Gao, A. K. **2023**. On the Origin of LLMs: An Evolutionary Tree and Graph for 15,821 Large Language Models. *ArXiv preprints, submitted*; pp. 1–14. DOI: 10.48550/arXiv.2307.09793

Gilbert, W.; Bureau, J. S.; Poellhuber, B.; Guay, F. **2022**. Educational contexts that nurture students' psychological needs predict low distress and healthy lifestyle through facilitated self-control. *Current Psychology*; *42*, pp. 29661–29618. DOI: 10.1007/s12144-022-04019-5

Goretzko, D.; Pham, T.T.H. and Bühner, M. **2021**. Exploratory factor analysis: Current use, methodological developments and recommendations for good practice. *Current Psychology*; *40*, pp. 3510–3521. DOI: 10.1007/s12144-019-00300-2

Goretzko D. and Bühner M. **2020**. One model to rule them all? Using machine learning algorithms to determine the number of factors in exploratory factor analysis. *Psychological Methods*; *25(6)*, pp. 776-786. DOI: 10.1037/met0000262

Hadi, M. U.; Al-Tashi, G.; Qureshi, R.; Shah, A.; Muneer, A.; Shaikh, M. B.; Al-Garadi, M. A.; Wu, J.; and Mirjalili, S. **2023**. Large Language Models: A Comprehensive Survey of Applications, Challenges, Limitations, and Future Prospects. *TechRxiv preprints, submitted*; pp. 1–44. DOI: 10.36227/techrxiv.23589741.v4

Hanum S. F.; Hasmayni, B.; and Lubis, A. H. **2023**. The Analysis of Chat GPT Usage Impact on Learning Motivation among Scout Students. *International Journal of Research and Review*, *10(7)*, pp. 632-–638. DOI: 10.52403/ijrr.20230774

Hung, J.; Chen, J. The Benefits, Risks and Regulation of Using ChatGPT in Chinese Academia: A Content Analysis. **2023**. *Social Sciences*; *12(9)*, pp. 380. DOI: 10.3390/socsci12070380

Kristensen, T. S.; Borritz, M.; Villadsen, E. and Christensen B. C. **2005**. The Copenhagen Burnout Inventory: A new tool for the assessment of burnout. *Work and Stress: An International Journal of Work, Health and Organisations*; *19(3)*, pp. 192-207. DOI: 10.1080/02678370500297720

Lee, Y. F.; Hwang, G. J.; and Chen, P. Y. **2023**. Impacts of an AI-based chabot on college students' after-class review, academic performance, self-efficacy, learning attitude, and motivation. *Educational technology research and development*; *70*, pp. 1843–1865. DOI: 10.1007/s11423-022-10142-8

Mofatteh M. **2020**. Risk factors associated with stress, anxiety, and depression among university undergraduate students.*AIMS Public Health.*; *8(1)*, pp. 36–65. DOI: 10.3934/publichealth.2021004

Neji, W.; Boughattas, N.; Ziadi, F. **2023**. Exploring New AI-Based Technologies to Enchance Students' Motivation. *Informing Science Institute*; *20*. DOI: 10.28945/5149

Oliveira, A. F.; Caliatto, S. G. **2018**. Exploratory Factor Analysis of a Scale of Learning Strategies. *Educação: Teoria e Prática*; *18(59)*, pp. 548-565. DOI: 10.18675/1981-8106.vol28.n59.p548-565

Okonkwo, C. W.; and Ade-Ibijola; A. **2021**. Chatbots applications in education: A systematic review. *SComputers and Education: Artificial Intelligence*; *2(2021)*, pp. 380. DOI: 10.1016/j.caeai.2021.100033

Pahune, S. and Chandrasekharan, M. **2023**. Several Categories of Large Language Models (LLMs): A Short Survey. *International Journal of Research in Applied Science and Engineering Technology*; *11(7)*. DOI: 10.22214/ijraset.2023.54677

Patil, V. H.; Singh, S. N.; Mishra, S.; and Donavan, D. T. **2008**. Efficient theory development and factor retention criteria: Abandon the 'eigenvalue greater than one' criterion. *Journal of Business Research*; *61(2)*, pp 162-170. DOI: 10.1016/j.jbusres.2007.05.008

Pereira, C. P. da S.; Santos, A. A. A. dos; Ferraz, A. S. **2020**. Learning Strategies Assessment Scale for Vocational Education: Adaptation and psychometric studies.*Revista Portuguesa de Educação*; *33(1)*, pp. 75–93. DOI: 10.21814/rpe.19127

Ramos, A. S. M. **2023**. Generative Artificial Intelligence based on large language models - tools for use in academic research. *Scielo preprints*, *submitted*. DOI: 10.1590/SciELOPreprints.6105

Rahman, M.; Terano, H. J. R.; Rahman, N. Salamzadeh, A. and Rahaman, S. **2023**. ChatGPT and Academic Research: A Review and Recommendations Based on Practical Examples. *Journal of Education, Management and Development Studies*; *3(1)*, pp. 1–12. DOI: 10.52631/jemds.v3i1.175

Sallam M.; Salim N. A.; Barakat M.; Al-Mahzoum K.; Al-Tammemi A. B.; Malaeb D.; Hallit R. and Hallit S. **2023**. Assessing Health Students' Attitudes and Usage of ChatGPT in Jordan: Validation Study. *Journal of Medical Internet Research*; *9*. DOI: 10.2196/48254

Sass, D. A.; Schmitt, T. A. **2010**. A comparative investigation of rotation criteria within exploratory factor analysis. *Multivariate Behavioral Research*; *45(1)*, pp. 73–103. DOI: 10.1080/00273170903504810

Shanahan, M. **2022**. Talking About Large Language Models. *ArXiv preprint*, *submitted*; *5*, pp. 1–13. DOI: 10.48550/arXiv.2212.03551

Taherdoost, H.; Sahibuddin, S.; and Jalaliyoon, N. **2022**. Exploratory Factor Analysis: concepts and theory. In *Advances in Applied and Pure Mathematics*. Jerzy Balicki. *27*, WSEAS, pp.375- 382. Available from: https://hal.science/hal-02557344/document

Theobald, M. **2021**. Self-regulated learning training programs enhance university students' academic performance, self-regulated learning strategies, and motivation: A meta-analysis.*Contemporary Educational Psychology*; *66*. DOI: 10.1016/j.cedpsych.2021.101976

Tu, X., Zou, J. Su, W. J., and Zhang, L. **2023**. What Should Data Science Education Do with Large Language Models? *arXiv preprints*, *submitted*; pp. 1–18. DOI: 10.48550/arXiv.2307.02792

Urban, M.; Děchtěrenko, F.; Lukavsky, J.; Hrabalová, V.; Švácha, F.; Brom, C.; Urban, K. **2023** ChatGPT Improves Creative Problem-Solving Performance in University Students: An Experimental Study. *PsyArXiv preprints*. DOI: 10.31234/osf.io/9z2tc

Yin, J.; Goh, T.; Yang, B. Xiaobin, Y. **2023**. Conversation Technology With Micro-Learning: The Impact of Chatbot-Based Learning on Students' Learning Motivation and Performance. *Journal of Educational Computing Research*; *59(1)*. DOI: 10.1177/0735633120952067

Zhai, X. ChatGPT User Experience: Implications for Education. **2023**. *SSRN Eletronic Journal*, *submitted*. DOI: 10.2139/ssrn.4312418

Wilson, M. L.; Huggins-Manley, A. C.; Ritzhaupt, A. D.; Ruggles, K. **2023**. Development of the Abbreviated Technology Anxiety Scale (ATAS). *Behavior Research Methods*; *55*, pp. 185-199. DOI: 10.3758/s13428-022-01820-9