Available online at https:
https://journal.uitm.edu.my/ojs/index.php/JIKM

e-ISSN: 2289-5337

Journal of Information and Knowledge Management (2024) Vol. 15, No. 2

Journal of Information and Knowledge Management

How Ready Are Students to Use ChatGPT? A Study on Instrument Validity and Reliability

Muhammad Naim Mohd Noh[1], A'dillah Mustafa[2*], Kasmarini Baharuddin[2]

[1]Pengurusan Air Selangor Sdn. Bhd. Jalan Templer PJS 5/1, 46050 Petaling Jaya Selangor
[2]Faculty of Information Science Selangor, Puncak Perdana Campus,,UiTM Selangor Branch, 40150 Shah Alam, Selangor, Malaysia

Corresponding Authors' Email Address: adillah973@uitm.edu.my

ARTICLE INFO

Article history:
Received: 3 August 2025
Revised: 1 September 2025
Accepted: 7 September 2025
Online first
Published: 1 October 2025

Keywords:
ChatGPT adoption
Behavioral intention
UTAUT model
Instrument validity
Instrument reliability

https://doi.10.24191/jikm.v15i2.9051

ABSTRACT

This study reports the construct validity and reliability of an instrument designed to measure students' behavioural intention to use ChatGPT. The instrument is developed based on the Unified Theory of Acceptance and Use of Technology (UTAUT) model. Data were collected from 65 undergraduate students of the Faculty of Information Management at UiTM Puncak Perdana using proportionate stratified sampling. Analyses conducted with SPSS and Smart PLS confirmed that the instrument demonstrated acceptable construct validity and moderate reliability. Internal consistency values included Cronbach's alpha (0.810–0.929), composite reliability (0.70–0.90), and average variance extracted above 0.50. Model accuracy and relevance (R^2 and Q^2) were within acceptable and moderate ranges, with no collinearity issues detected. Content validity was further established using CVI, S-CVI, and S-CVI/Ave, which exceeded 0.80, indicating strong agreement. Overall, the instrument achieved robust validity and reliability, offering a useful inventory for future research on ChatGPT adoption in education.

INTRODUCTION

In recent years, Artificial Intelligence (AI) has made significant progress and it is now used across a wide range of various industries and applications (Nasim et al., 2022). Moreover, AI is altering the way people live and work that expecting AI to evolve and play a more significant part in society as research and technology advancements continue. AI also has been utilizing in education sector, due to its ability to transform education through improving teaching and learning experiences, customizing education, streamlining administrative duties, and opening up new avenues for research (Zhai et al., 2021). AI technology can be used in a variety of ways in the education, including intelligent tutoring systems, adaptive learning platforms, automated grading systems, virtual reality simulations, and data analytics (Holmes & Tuomi, 2022). In term of intelligent tutoring systems, an AI program like Chat GPT is a tool that can assist students in producing writing assignments, providing feedback and revision guidelines, and providing writing assistance. Consequently, it can be used to help them develop their skills. (Abdullayeva & Musayeva, 2023) and enhance the learning process (Ausat et al., 2023). However, the usage of Chat GPT among students become contradiction, some parties agrees that this technology is able to improve students learning process. Conversely, some of them put on as the opposite that resistance of the use of Chat GPT (Crawford et al., 2023). The pro parties claimed that Chabot such Chat GPT can improve productivity, communication, learning, and teaching support. A new educational platform can address the most difficult educational problems by utilizing this technology as an engagement tool (Sandu & Gide, 2019). In contrast, the other parties stated that students might have a potential for overreliance on Chat GPT. Instead of honing their critical thinking and problem-solving skills, they become dependent on the model for answers and solutions. Their inability to think freely and creatively, which are essential for both academic success and personal growth, may be hampered by this over-reliance (Fuchs, 2023). In line with this, a study from Ali et al. (2023) pointed out that the use of Chat GPT for learning is able to enhance the students' learning motivation. This study intended to evaluate the construct validity and reliability of an instrument designed to measure students' behavioral intention to use ChatGPT, based on the Unified Theory of Acceptance and Use of Technology (UTAUT) model (Venkatesh, Morris, Davis & Davis, 2003). This study also aim to investigate the suitability of the Behavioral Intention Model in ChatGPT adoption study.

Problem Statement

Limited Availability of ChatGPT-Specific Instruments in the Malaysian Context

There is currently a scarcity of validated instruments tailored for studying ChatGPT usage among Malaysian students. This gap is due to two primary factors:
Technical and Domain Constraints of ChatGPT: Large language models like ChatGPT exhibit notable limitations in understanding highly complex or specialized content. For instance, the model may struggle with nuanced legal or medical queries, often generating incorrect, generic, or even misleading responses (Gopalakrishnan, 2023; Lai & Adebayo, 2024). Specifically, it lacks deep contextual awareness and domain-specific precision, which makes it unreliable for critical or expert-level tasks. Accessibility Challenges in Malaysia: In the Malaysian context, accessing ChatGPT—particularly for research—can be difficult. Although ChatGPT Plus, offering improved access and elevated performance, is available, it comes at a subscription cost of around US$20 (approximately RM85) per month (The Star, 2023; TechTRP, 2023). The free version, while available, often suffers from performance issues like slow responses or unavailability during peak usage times (TechTRP, 2023). These constraints pose practical barriers for researchers attempting to conduct consistent, reliable studies using ChatGPT in Malaysia.

Lack of ChatGPT Studies Applying Behavioral Intention Theory to Student Adoption

There is a notable gap in the literature regarding the application of behavioral intention theories to specifically examine students' adoption of ChatGPT. While several studies have applied theoretical frameworks such as the Unified Theory of Acceptance and Use of Technology (UTAUT), their focus has been outside the student population. For instance, Emon et al. (2023)

investigated factors influencing ChatGPT adoption behavior among professionals in Bangladesh, applying the UTAUT model; however, the study was limited to knowledge workers, excluding students. Similarly, Menon et al. (2023) analyzed user intentions toward ChatGPT adoption using UTAUT, but their sample comprised self-employed and employed workers, not learners in academic contexts.

Given that ChatGPT is a relatively new technology, adoption studies remain at an early stage, with most focusing on general or professional users rather than students. Moreover, the application of established models such as the Technology Acceptance Model (TAM) or the Theory of Planned Behavior (TPB) to ChatGPT-specific student adoption is still lacking (Zhou et al., 2023; Dwivedi et al., 2023). This highlights the need for empirical research targeting the educational domain, where students represent a critical user group for AI-driven technologies like ChatGPT.

Research Objective

The objective of this study is to assess and report the construct validity and reliability of student adoption of ChatGPT in academic environment instruments. This research covers the adoption of ChatGPT among university students. The instrument aims to investigate the influence of performance expectancy, effort expectancy, social influence, facilitating condition, attitude and anxiety towards the behavioural intention. It was developed by adapting constructs and items from recent research and the development of new items. The assessment analysed an overall 26 items from seven constructs developed using the UTAUT Model.

LITERATURE REVIEW

Artificial Intelligence Overview

Artificial intelligence is a branch of computer science that concerns the ability of computers to perform intelligent tasks, such as those requiring recognition, reasoning, and learning. Artificial intelligence, as a discipline, initiated by scientist John McCarthy (McCarthy et al., 2006) and formally publicized at Dartmouth Conference in 1956 (Moor, 2006), now occupies a center stage for many organizations (Balugani et al., 2018; Sreedharan et al., 2018; Balugani et al., 2019; Kumar, 2019). Artificial intelligence forms a part of computer or computing science that generates expert systems, algorithms and programs (Aldasoro et al., 2019; Chen et al., 2019; Colicchia et al., 2019). The real purpose of artificial intelligence is to imitate human brain and perform decision-making like human beings during various situations and circumstances (De Sousa Jabbour et al., 2018; Deshpande et al., 2018). Practically, it is the ability of a machine or equipment to think, learn and act as humans. Although the concept of artificial intelligence was initiated in 1956, but it has gained momentum recently (Dolgui et al., 2018; Felfel et al., 2018; Garza-Reyes, 2018; Dolgui et al., 2019).

Artificial intelligence is categorized as narrow artificial intelligence and general artificial intelligence. The former deals with simple issues like organizing business events/calendars and resolving customer-service enquiries. However, the latter handles complex issues like driving cars, robotics and reducing language barriers (Breja et al., 2011; Choiet al., 2018; Dupontet al., 2018). Business creation by artificial intelligence is expected to touch $3.9 tn in 2022, while it was $1.2 tn in 2018, reported an increase of 70 % from 2017 (Brown, 2013; Fatorachian and Kazemi, 2018; Richards et al., 2019). Further, expert systems, agent-based systems, big data analytics and genetic algorithms are few other concepts that fall in close proximity of artificial intelligence. Expert systems constitute one of the most essential aspects of AI. It mimics the decisiveness of an expert human to solve complex issues and problems by applying reasoning and analytical abilities like computers (Keeble-Ramsay and Armitage, 2010; Hamdi et al., 2018; Hazen et al., 2018; Horng et al., 2018). Agent- based systems, also known as agent-based models, are computer-based models, which perform as autonomous agents. Such agents can work either in groups or individually towards assigned tasks. While, Big Data denotes huge volume of data, and big data analytics is the management of such data. It does systematic organization of unorganized data that might be large and complex. The technique is extremely useful to avoid information asymmetry (Kazancoglu et al., 2018; Kumar et al., 2018; Kumar et al., 2019).

ChatGPT

"GPT" in ChatGPT stands for "generative pretrained transformer", the latest AI development in generative machine learning (ML), which is more powerful than its previous generated forms. Not only can it make predictions; it can also create and develop answers to almost any question (McKinsey, 2023). For testing purposes, OpenAI released it to the public on November 30, 2022. It uses a large language model (LLM) trained on billions of words from online sources such as books, articles and conversations, and it produces answers by learning from this. Its abilities include, but are not limited to, accurate text translation, generating marketing copy, summarizing reports and news and coding. It can understand the meaning of large texts and generate accurate responses. Benuyenah (2023) is convinced that the impact of AI in higher education is undeniable and impossible to ignore. If ethical considerations accompany its integration, it can greatly contribute to the health of higher education.

ChatGPT was an outcome of the huge investments in AI which have more than doubled over the last five years and are projected to continue growing (McKinsey, 2022). Yet despite this trend, the effects of ChatGPT and similar AI tools on the education sector remain uncertain. It has the potential to drastically alter the landscape of education, causing us to re-evaluate traditional methods of learning, assessment and evaluation. As AI continues to shape the job market, concerns are growing about the necessary skills for the new work environment, tools like ChatGPT demonstrate the potential to automate routine tasks, saving time and cost, and even to perform parts of more complex and creative work (Zhai, 2022). However, it should be recognized that ChatGPT has certain limitations, for instance, those related to reasoning, factual errors, math, coding and bias (Winner, 2009 and Borji, 2023). As ChatGPT has brought. Its implications for education should concern us and despite our great excitement need to be carefully evaluated, so that we are alert to its potential influence on education. ChatGPT's technology has propelled its rapid integration into search engines, such as Microsoft's Bing, through OpenAI's ChatGPT integration, and the introduction of Google's rival tool, Bard (Financial the New York Times, 2023). Notably, other companies are also striving to develop more advanced alternatives, such as Baidu's Ernie Bot in China. According to the MIT Technology Review, Ernie Bot resembles ChatGPT in performance but only the former can generate new images. Furthermore, GPT-4, which was introduced in

March 2023, is reported to be a hundred times more powerful than GPT-3 (OpenAI, 2023) and thus ensures significant advances based on the data gathered from its predecessor.

Artificial Intelligence Application in Academic

Artificial Intelligence (AI) has emerged as a transformative force reshaping various sectors of society, and academia is no exception. In recent years, AI technologies have revolutionized teaching, learning, research, and administrative tasks in academic institutions worldwide (Vieriu and Petrea (2025). A study by Augustine and Ali (2021) focused on the application of AI in academic libraries in Nigeria, showing that AI could be applied in academic library services in Nigeria like Expert Systems in Reference Services, Technical, Indexing, Acquisition, and its application in Natural Language Processing, Pattern Recognition, and Robotics in library activities. Additionally, Pathak (2023) and Ahmad (2023) emphasized the AI-powered educational platforms utilize machine learning algorithms to provide personalized learning experiences tailored to each student's individual needs, preferences, and learning styles. These platforms analyze student performance data to offer adaptive content, recommendations, and feedback, enhancing student engagement and learning outcomes. AI-based tutoring systems provide interactive, individualized support to students, offering explanations, feedback, and guidance aligned with their learning pace and needs. These systems improve student understanding and mastery of concepts across various subjects, supplementing traditional classroom instruction. AI model predict student outcomes, such as academic performance and retention, based on historical data and student characteristics. These predictive analytics tools enable academic institutions to identify at-risk students early and provide targeted interventions to support their success, fostering student engagement and retention. AI technologies automate repetitive tasks in the research process, such as literature reviews, data collection, and experiment design, freeing up researchers' time for higher-level tasks. This automation enhances research productivity, accelerates knowledge generation, and facilitates interdisciplinary collaboration.

In conclusion, the widespread adoption of AI technologies in academia is transforming teaching, learning, research, and administrative tasks, revolutionizing the academic landscape. From personalized learning platforms to predictive analytics tools, AI offers a myriad of opportunities to enhance educational experiences, advance research frontiers, and improve administrative efficiency in academic institutions. Embracing AI in academia promises to unlock new possibilities for innovation, collaboration, and knowledge creation, shaping the future of education and scholarship in the digital age. Table 1 below highlights the selected studies on ChatGPT in education applying UTAUT and a systematic review to provide a holistic views on the perspectives.

Table 1: Related studies on ChatGPT in education

Author	Aim of Study	Research Method	Main Findings
Shoufan, A (2023)	To analyze its impact on teaching and learning it is crucial to understand how students perceive ChatGPT and assess its potential and challenges.	Questionnaire	Perception's student about ChatGPT

Pathak,M (2023)	To examine the awareness regarding ChatGPT among students.	-Online survey - Descriptive statistics and independent sample t-test has been employed using IBM SPSS 27 software to examine the awareness level of ChatGPT.	The study attempts to find whether is there any difference in awareness of ChatGPT based on gender or the field of study among the students.
Tlili, A (2023)	To examine chatbots in education and for this purpose, the study approaches ChatGPT as a representative case of an advanced chatbotamong early adopters	-Qualitative method	Investigate social network analysis of tweets, content analysis of interviews, and user experiences using ChatGPT
Albanna, S (2023)	To explore the advantages of integrating a new generative artificial intelligence (AI) technology in education.	-Literature search -Screening and Selection -Analysis -Synthesis of Findings	ChatGPT can be effectively integrated into education to automate routine tasks and enhance the learning experience for students, ultimately increasing productivity and efficiency and fostering adaptive learning
Yusoff, N. (2022)	To explore the perceive benefits of AI in ELS and language.	-ESL -Language Learning	Focuses on AI in ESL classrooms; identifies benefits like adaptive feedback, and challenges such as teacher preparedness and cost.
Shoaib, M. et al.(2021)	To highlight the opportunities of AI applications in secondary and tertiary education.	Systematic Review	Provides a comprehensive review of AI applications in education; discusses opportunities (personalization, assessment) and challenges (privacy, ethics).
Lu, X & Yang, J (2021)	Review the contributions of AI application in personalization education.	Systematic Review Sustainability & Education	Reviews how AI contributes to sustainable education practices; highlights benefits such as personalized learning and challenges such as bias.
Hussain, I. (2020)	Reports generally positive attitudes toward AI's instructional role; highlights teacher concerns about training and control	University students & teachers (survey)	Directly addresses *attitude* toward AI: showing positive attitude is present, which is a key factor in many acceptance models. Does not appear to deeply explore anxiety or UTAUT but relevant for the attitude part.
Shon, K & Kwon, O (2020)	Useful crossover: while not education-only, shows what constructs (usefulness, enjoyment) tend to drive intention; can inform which constructs you might include for AI in education.	Users of AI-based intelligent products (consumer / product context)	Compared different acceptance theories (TAM, UTAUT, VAM etc.) for AI-based intelligent products; found that enjoyment, perceived usefulness, etc. are important; VAM did well in modeling acceptance.

Theoretical Framework

In understanding the students' behavioral intention in using ChatGPT, Figure 1 illustrates the framework of this research. To analytically accomplish the objective of this study, development of theoretical framework is necessary so that further conclusions can be made.

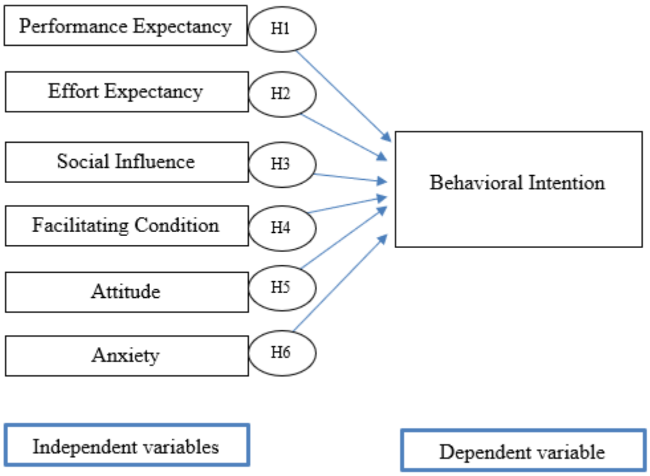

Figure 1: Theoretical Framework for assessing students' behavioral intention in using ChatGPT.

Description of Variables

Performance Expectancy

Performance Expectancy is a concept within the Unified Theory of Acceptance and Use of Technology (UTAUT) Venkatesh et, al. (2003) model, which seeks to explain individuals' acceptance and use of new technologies. Specifically, Performance Expectancy refers to the degree to which an individual believes that using a particular technology will help them to achieve gains in job performance or task accomplishment. In simpler terms, Performance Expectancy assesses the perceived usefulness of the technology in enhancing an individual's performance or productivity. It reflects the user's belief that using the technology will lead to improved outcomes, such as increased efficiency, effectiveness, or convenience in performing tasks.

Effort Expectancy

Effort expectancy focuses on the perceived ease of use or the degree of effort required to use the technology. It assesses the user's perception of how easy or difficult it is to use a particular technology. It reflects the user's belief that using the technology will require minimal effort and be straightforward, intuitive, and user-friendly. Effort Expectancy is a critical factor in determining the acceptance and adoption of technology or systems, as users tend to prefer technologies that are easy to use (Kwak et al., 2022).

Social Influence

Social influence aims to explain individuals' acceptance and use of new technologies. Social Influence refers to the perceived impact of social factors, such as the expectations, opinions, and behaviors of others, on an individual's intention to use a technology. Social Influence reflects the extent to which an individual's decision to adopt and use a technology is influenced by the attitudes, norms, and behaviors of their social network, including peers, colleagues, supervisors, and other influential individuals. Social influence is defined as the extent of social pressure exerted on an individual to adopt new technology (Chaouali et al., 2016; Kesharwani and Singh Bisht, 2012; Martins et al., 2014)

Facilitating Condition

Facilitating Conditions aims to explain individuals' acceptance and use of new technologies. Facilitating Conditions refer to the perceived resources, support, and infrastructure available to individuals that facilitate their use of the technology. Facilitating Conditions assess the extent to which individuals perceive that they have the necessary resources, support, and organizational infrastructure to effectively use the technology. According to Hong et al. (2008), "if users would not have necessary operational skills, they would have lower intention to adopt information technology."

Attitude and Anxiety

In many UTAUT extensions, attitude is added as an additional predictor of behavioral intention, defined as an individual's overall positive or negative evaluation of using the technology (Or, 2023). Likewise, anxiety (or technology anxiety) has been incorporated in extended UTAUT models as an emotional factor capturing worry, stress, or apprehension about using the system; empirical studies find that anxiety negatively influences intention (Gunasinghe, 2020) and more recent work adds AI anxiety in UTAUT extensions (Wang et al.,2024).

INSTRUMENT DEVELOPMENT

In designing the instrument, this study adapted and adopted questions from the previous studies. A detail literature review that has been done to identify related studies. All the items adapted have been tested the validity and reliability. The instrument is the major contributors for this study. The instrument consists of items that use a five-point Likert Scales for measuring all the variables. It will be five-point from 1-5: (1) Strongly Disagree, (2) Disagree, (3) Neutral, (4) Agree and (5) Strongly Agree. The instrument is divided into 7 parts, namely A: Performance Expectancy, B: Effort Expectancy, C: Social Influence, D: Facilitating

Condition, E: Attitude, F: Anxiety and G: Behavioral Intention. Table 2 below explains the instrument items and references.

Table 2: Instrument Items and References

No	Variable/Theory	Item	Reference
1	Performance Expectancy	1.I would find ChatGPT useful in my study	D, J., Srinivasan, M., Dhanunjay, G. S., & Shamala, R. (2023), Pangaribuan, C. H., & Wulandari, Y. (2019)
		2. Using ChatGPT enables me to accomplish task more quickly	
		3. Using ChatGPT increases my productivity.	
		4. Using ChatGPT is a good idea.	
2	Effort Expectancy	1.My interaction with ChatGPT would be understandable	D, J., Srinivasan, M. Dhanunjay, G. S., & Shamala, R. (2023), Pangaribuan, C. H., &Wulandari, Y. (2019)
		2.It would be easy for me to become skillful at using ChatGPT	
		3.I would find ChatGPT easy to use	
		4.Learning to use ChatGPT easy to use	
3	Attitude	1.Using ChatGPT is a good idea	Kanwal, A., Hassan, S., & Iqbal, I. (2023)
		2. ChatGPT makes study more interesting.	
		3.Using ChatGPT is fun	
		4.I like interacting with ChatGPT	
4	Social Influence	1. People who influence my learning behavior think that I should use ChatGPT.	Ventakesh et al. (2003), Pangaribuan, C. H., & Wulandari, Y. (2019)
		2.People who are important to me think that use of ChatGPT	
		3. In general, the university has supported the use of ChatGPT.	
5	Facilitating Conditions	1. I have the resources necessary to use ChatGPT.	Pangaribuan, C. H., & Wulandari, Y. (2019)
		2.I have necessary knowledge to use ChatGPT	
		3.ChatGPT is not compatible with other application I use	
		4. I can get assistant whenever I come across with ChatGPT with ChatGPT difficulties.	
6	Anxiety	1. I feel uneasy about using ChatGPT.	Shen et al. (2022)
		2. It scares to me think that I could lose a lot of information using ChatGPT by giving the wrong command.	
		3. ChatGPT is somewhat intimidating.	
		4.I doubt to use ChatGPT for fear of making mistakes	
7	Behavioral Intention	1.I intend to use ChatGPT in the coming semester	Davis et al. (1989), Shen at al (2022)
		2. I predict I will be using ChatGPT in coming semester	
		3. I will encourage others to use ChatGPT in coming semester.	

METHODOLOGY

Research Design

This study employs a quantitative method using survey research techniques. The survey instrument was assessed using smartPLS-SEM 3.0 to determine its validity and reliability. The smartPLS-SEM 3.0 analysis of validity and reliability was divided into two analyses: i) analysis of the measurement model (outer model) and ii) analysis of the structural model (inner model).

Population and Sampling

The average population of undergraduate students at UiTM Puncak Perdana's Faculty of Information Management is around 2000. The sample size of the respondents has been run from Raosoft and the results from the calculation is 323 respondents. The sample size for the pilot test is around 10 to 20% of full-scale survey sample size or at least 30 to 50 respondents according to Hertzog (2008). So, from 20% of 323 will be 65 respondents. Hence, the questionnaires will be distributed to 65 respondents. The sample targeted are undergraduate students from the Faculty of Information Management

Data Collection

This study distributed the questionnaires to the respondents of the undergraduate students from the Faculty of Information Management, UiTM Puncak Perdana. After the process of creating questionnaires using the Google Form platform, the link was distributed to the respondents through a social media platform such as WhatsApp and email and had to ensure they respond within the specified time frame.

DATA ANALYSIS

The following were the main statistical techniques used to analyze the data and obtain the results of the study:

i. Descriptive statistics to analyze the respondent's profile.
ii. Cronbach's Alpha, a reliability measure that indicates how well a set of items relates to each other (that is if the items have internal consistency). A higher value of Cronbach's Alpha (0.6 or higher) indicates higher reliability of the item's measures.
iii. Structural Equation Modelling (SEM) using SPSS and SmartPLS to estimate the causal-effect relationship among multiple and interrelated variables.

This study undertakes two validity procedures, which are face validity and content validity testing. Rahman et al. (2016) denote that validity is the extent which specific items on a tool accurately assess the concept being measured in the research. The instrument face and content validity were concurrently assessed by a panel of expert. The validation procedure took approximately 1 month to complete. The content validity test for the I-CVI and S-CVI are calculated based on the following formula:

I-CVI are calculated as by dividing the number of experts that rated 3 (Relevant) and 4 (Very Relevant) on the relevancy scale by the total number of the number of experts. S-CVI are calculated as I-CVI/Ave, which calculate the average of all I-CVI. S-CVI= I-CVI/Ave

This study also undertakes analyses of the instrument construct validity and reliability were performed using the Smart PLS-SEM 3.0. According to Ibrahim and Tain (2016), the analysis involved two-level models. First model is analyze the measurement model to assess the instrument construct validity using internal consistency of the Cronbach's alpha (CR) value, the convergent validity using factor loading and average variance extracted (AVE) values, and discriminant validity using Fornell-Larcker's criterion and cross-loading value (Hair, Hult, Ringle, & Sarstedt, 2017; Hair, Sarstedt, Hopkins, & Kuppelwieser, 2014). Second is the analysis of the structural model assessing the construct reliability using the model predictive accuracy analyses. The model predictive accuracy involved coefficient determination (R^2), predictive relevance (Q^2), effect size (f^2), and VIF (Collinearity Statistics).

RESULT

Instrument Assessment

Validity

a) Face Validity

The instrument is sent to an identified panel of experts in the field of librarianship and academia for the validity assessment. The expert eligibility is set according to their fields of expertise, experience in the field of library and information science, total number of working experience, professional certification and academic certification. Below table 3 describes the expert panels.

Table 3: Expert Panel

Panel	Expert in Field
Academician (2)	Subject Expert & Methodology
Librarian (1)	Subject Expert

An invitation letter is sent to the panels to seek their agreement to participate in the validation process. The instrument was sent upon receiving their agreement and consent. They are required to provide comments and suggestions on the following face validity criteria as suggested by Oluwatayo (2012) namely: i) Appropriateness of grammar, ii) The clarity and unambiguity of items, iii) The correct spelling of words, iv) The correct structuring of sentences, v) Appropriateness of font size, vi) Structure and format, vii) Appropriateness of difficulty level for respondents, and viii) Adequacy of instruction on the instrument. The face validity assessment analysis is represented in Table 4 describes the 8 elements of the assessment.

Table 4: Analysis of Face Validity

Criteria	P1	P2	P3	Interpretation
Appropriateness of grammar	Yes	Yes	Yes	Appropriate
The clarity and unambiguity of items	Yes	Yes	Yes	Appropriate
The correct spelling of words	Yes	Yes	Yes	Appropriate
The correct structuring of sentences	Yes	Yes	Yes	Appropriate
Appropriateness of font size	Yes	Yes	Yes	Appropriate
Structure and format	No	Yes	Yes	Need for revision
Appropriateness of difficulty level for respondents	Yes	Yes	Yes	Appropriate
Adequacy of instruction on the instrument	No	Yes	Yes	Need for revision

Based on the comments from the panel, there are two (2) elements of the instrument need to be reviewed and improved. The structure and format and adequacy of instruction on the instrument was received two disagreements (with "No" remark). The instrument then undergone a revision exercise, accordingly to the recommendations.

b) Content Validity

In the content validity assessment, panels are required to rate the item according to the scoring guide based on the relevancy and clarity if the items. The rating of the score is from 4 (Very relevant) to 1 (Not relevant) and 4 (Very clear) to 1 (Not clear). Below table 5 represents the items and constructs for content validation.

Table 5: Total items

Name of Construct	Total Items	Number of items
Performance Expectancy	4	1-4
Effort expectancy	4	5-8
Attitude	4	9-12
Social influence	4	13-16
Facilitating conditions	4	17-20
Anxiety	4	21-24
Behavioral Intention	3	25-27
Total	27	

I-CVI, and S-CVI

The content validity assessment is analyzed using two methods: I-CVI, and S-CVI calculation. I-CVI and S-CVI is divided into two subtopic which is relevancy and clarity. Table 6 and Table 7 describe the relevancy and clarity assessments.

Table 6: Calculation I-CVI Based on Relevancy

Items	Relevant	Non-relevant	I-CVI	Interpretation
1	3	0	1	Appropriate
2	3	0	1	Appropriate
3	3	0	1	Appropriate
4	3	0	1	Appropriate
5	2	1	0.67	Need for refinement
6	3	0	1	Appropriate
7	3	0	1	Appropriate

8	3	0	1	Appropriate
9	3	0	1	Appropriate
10	3	0	1	Appropriate
11	3	0	1	Appropriate
12	3	0	1	Appropriate
13	3	0	1	Appropriate
14	3	0	1	Appropriate
15	3	0	1	Appropriate
16	3	0	1	Appropriate
17	3	0	1	Appropriate
18	3	0	1	Appropriate
19	3	0	1	Appropriate
20	3	0	1	Appropriate
21	3	0	1	Appropriate
22	3	0	1	Appropriate
23	3	0	1	Appropriate
24	3	0	1	Appropriate
24	3	0	1	Appropriate
26	3	0	1	Appropriate
27	3	0	1	Appropriate

Table 7: Calculation I-CVI based on Clarity

Items	Clarity	Unclear	I-CVI	Interpretation
1	3	0	1	Appropriate
2	3	0	1	Appropriate
3	3	0	1	Appropriate
4	3	0	1	Appropriate
5	2	1	0.67	Need for refinement
6	3	0	1	Appropriate
7	3	0	1	Appropriate
8	2	1	0.67	Need for refinement
9	3	0	1	Appropriate
10	3	0	1	Appropriate
11	3	0	1	Appropriate
12	3	0	1	Appropriate
13	3	0	1	Appropriate
14	3	0	1	Appropriate
15	3	0	1	Appropriate
16	3	0	1	Appropriate
17	3	0	1	Appropriate
18	3	0	1	Appropriate
19	3	0	1	Appropriate
20	3	0	1	Appropriate
21	3	0	1	Appropriate
22	3	0	1	Appropriate
23	3	0	1	Appropriate
24	3	0	1	Appropriate
25	2	1	0.67	Need for refinement
26	3	0	1	Appropriate
27	3	0	1	Appropriate

The calculation of I-CVI (Content Validity Index) according to Larsson et al. (2015) and Polit, Beck & Owen (2007) the I-CVI was calculated for both "relevant" and "clarity". Schilling, Dixon and Knafi, et al (2007) define I-CVI as the proportion of expert that provided a rating of 3 (very relevant) or 4 (extremely relevant) on the relevance scale and rating 3 (fairly clear) or 4 (very clear) on the clarify scale. The I-CVI cutoff as suggested by Lynn (1986) and Polit, Beck & Owen (2007) is a below than 0.78. As for this assessment, a conservative value of ≥.80 for both relevant and clarify are used as suggested by Paul, et al (2016). Items rated below 0.80 are eliminated. Based on the I-CVI, from Relevance table 4.5 shows that, 1 was remarked as need some revision and 26 items rated very relevant. After that, from Clarity table 4.6 shows that, 3 are remarked as need some revision also and 24 items rated very clear. So, the analysis of the overall items originates 1 item are rated below 0.80 for relevancy and 3 items are rated below 0.80 for clarity. A total of 4 items (0.67) are eliminated based on the relevancy and clarity rates, 3 items are reworded or revised for clarity. Table 8 & 9 described the analysis of the S-CVI for inter-ratter agreement and calculation of S-CVI /Ave.

Table 8: Scale-level content validity index (S-CVI) (Relevance)

Item	Expert in Agreement	Item I-CVI
1	3	1
2	3	1
3	3	1
4	3	1
5	3	1
6	3	1
7	3	1
8	2	0.67
9	3	1
10	3	1
11	3	1
12	3	1
13	3	1
14	3	1
15	3	1
16	3	1
17	3	1
18	3	1
19	3	1
20	3	1
21	3	1
22	3	1
23	3	1
24	3	1
25	3	1
26	3	1
27	3	1

S-CVI (Relevance) = 0.98

Table 9: Scale-level content validity index (S-CVI) (Clarity)

Items	Expert in Agreement	Item I-CVI
1	3	1
2	3	1
3	3	1
4	3	1
5	2	0.67
6	3	1
7	3	1
8	2	0.67
9	3	1
10	3	1
11	3	1
12	3	1
13	3	1
14	3	1
15	3	1
16	3	1
17	3	1
18	3	1
19	3	1
20	3	1
21	3	1
22	3	1
23	3	1
24	3	1
25	2	0.67
26	3	1
27	3	1

S-CVI (Clarity) = 0.96

The S-CVI (Scale-level Content Index) of this study are calculated based on Polit, Beck & Owen (2007) and Davis (1992). Where they suggested the value of S-CVI should be greater than 0.8 or 80% or better agreement among reviewers. The S-CVI based on relevance and clarity for this study instrument are 0.98 and 0.96 and considered achieve "high-level agreement" which is acceptable and consistence

Reliability

Analyses of the instrument construct validity and reliability were performed using the Smart PLS-SEM 3.0. According to Ibrahim and Tain (2016), the analysis involved two-level models which is analysis of the measurement model to assess the instrument construct validity using internal consistency of the Cronbach's alpha (CR) value, the convergent validity using factor loading and average variance extracted (AVE) values, and discriminant validity using Fornell-Larcker's criterion and cross-loading value (Hair, Hult, Ringle, & Sarstedt, 2017; Hair, Sarstedt, Hopkins, & Kuppelwieser, 2014) and analysis of the structural model to assess the construct reliability using the model predictive accuracy analyses. The model predictive accuracy involved coefficient determination (R^2), predictive relevance (Q^2), effect size (f^2), and VIF (Collinearity Statistics). This analyses also is constructed in two sections which is analysis of the measurement model (outer model) and analysis of the structural model (inner model). Both the outer and inner models follow the Smart PLS-SEM 3.0 requirements of structural model analysis.

Analysis of Measurement Model

Analysis of the construct validity and reliability of the measurement model was based on Internal Consistency. Internal consistency analysis is using Cronbach's alpha and composite reliability resulted in all dimensions being accepted at a score from 0.718 to 0.955, which exceeded the cut-off point of >.07 Cronbach's alpha. The composite reliability (CR) value was acceptable (Gefen, Straub, & Boudreau, 2000) with all dimension values of n > 0.70. Next, analysis of the construct validity and reliability of the measurement model also was based on Convergent Validity. Convergent validity was assessed using indicator loading and the average variance extracted (AVE) value. It was considered satisfactory since all dimension values exceeded the cut-off point of >.50 (Byrne, 2010; Fornell & Larcker, 1981; Hair, Sarstedt, et al., 2014). For item loading, the cut off value for this report was accepted at 0.6, where 0.4 – 0.6 was considered acceptable (Hair, Black, Babin, & Anderson, 2010). Thus, the model achieved internal consistency and convergent validity. Table 10 describes the construct's Cronbach's alpha, CR, AVE value, and item loadings.

Table 10: Measurement Model Analysis

Construct	Items	Items Loading	CA	CR1	AVE
PERFORMANCE EXPECTANCY			**0.929**	**0.932**	**0.824**
	B1: I would find ChatGPT useful in my study	0.923			
	B2: Using ChatGPT enables me to accomplish. tasks more quickly.	0.904			
	B3: Using ChatGPT increases my productivity.	0.888			
	B4: Using ChatGPT is a good idea.	0.916			
EFFORT EXPECTANCY			**0.929**	**0.947**	**0.826**
	C1: My interaction with ChatGPT would be understandable.	0.941			
	C2: It would be easy for me to become skillful at using ChatGPT.	0.828			
	C3: I would find ChatGPT easy to use.	0.941			
	C4: Learning to use ChatGPT is easy for me.	0.920			
ATTITUDE			**0.909**	**0.912**	**0.786**
	D1: Using ChatGPT is a good idea.	0.877			
	D2: ChatGPT makes study more interesting.	0.873			
	D3 : Using with ChatGPT is fun.	0.954			

	D4 : I like interacting with ChatGPT.	0.88			
SOSIAL INFLUENCE			**0.810**	**0.821**	**0.643**
	E1: People who influence my learning behavior think that I should use ChatGPT. *	**0.658**			
	E2: People who are important to me think that I should use ChatGPT.	0.856			
	E3: Seniors in this program has been helpful in the use of ChatGPT.	0.855			
	E4: In general, the university has supported the use of ChatGPT.	0.821			
FACILITATING CONDITIONS			**0.834**	**0.888**	**0.681**
	F1: I have the resources necessary to use ChatGPT.	0.910			
	F2: I have the necessary knowledge to use ChatGPT.	0.936			
	F3: ChatGPT is not compatible with other applications I use. *	**0.553**			
	F4: I can get assistant whenever I come across with ChatGPT difficulties.	0.844			
ANXIETY			**0.887**	**0.817**	**0.609**
	G1: I feel uneasy about using ChatGPT. *	**0.603**			
	G2: It scares me to think that I could lose a lot of information using ChatGPT by giving the wrong command.	**0.140**			
	G3: ChatGPT is somewhat intimidating to me.	**0.009**			
	G4: I doubt to use ChatGPT for fear of making mistakes.	**0.232**			
BEHAVIORAL INTENTION			**0.893**	**0.896**	**0.823**
	H1: I intend to use ChatGPT in the coming semester.	0.904			
	H2: I predict I will be using ChatGPT in the coming semester.	0.918			
	H3: I will encourage others to use ChatGPT in the coming semester.	0.900			

*Note. * Item loading below 0.7 are suggested to be retained due to AVE > 0.5*

Discriminant Validity

Discriminat validity is one of the analyses use to construct validity and reliability of the measurement model. Discriminant validity was assessed using the Fornell-Larcker criterion and item cross loading. The Fornell-Larcker test revealed that the individual construct loading was greater than in another construct loading and obtained a value of less than 0.1 as in the model suggested by Chin (1998); Anxiety=0.331, Attitude=0.887, Behavioral Intention=0.907, Effort Expectancy=0.909, Facilitating conditions= 0.825, Performance Expectancy=0.908 and Social influence=0.802. Thus, the model met the requirements for discriminant validity. Table 4.10 presents the Fornell and Larcker criteria. The items were assessed for discriminant validity using their cross-loading values. Item loading must be higher in the construct than in another construct. The loading indicator requirement item loaded at 0.7 or higher was recommended, but lower loadings (0.4) are adequate (Hair, et al., 2014). A loading of 0.6 was acceptable for this study, and discriminant validity was achieved. Table 11 explains the item cross loadings.

i) Fornell-Larcker Criterion

Table 11: Fornell-Larcker Criterion

	Anxiety	**Attitude**	**Behavioral Attention**	**Effort Expectancy**	**Facilitating Conditions**	**Performance Expectancy**	**Social Influence**
Attitude	0.326	0.887					
Behavioral Attention	0.288	0.789	0.907				
Effort Expectancy	-0.259	0.809	0.679	0.909			
Facilitating Conditions	-0.175	0.745	0.789	0.653	0.825		
Performance Expectancy	-0.236	0.884	0.756	0.775	0.762	0.908	
Social Influence	0.069	0.674	0.579	0.502	0.593	0.615	0.802

Items Cross Loading

Table 12: Item Cross Loading

Items	Anxiety	Attitude	Behavioral Intention	Effort Expectancy	Facilitating Conditions	Performance Expectancy	Social Influnce
B1	-0.237	**0.815**	**0.690**	**0.783**	**0.701**	**0.923**	0.523
B2	-0.200	**0.770**	**0.724**	**0.662**	**0.668**	**0.904**	0.557
B3	-0.176	**0.780**	**0.624**	**0.678**	**0.695**	**0.888**	0.530
B4	-0.240	**0.845**	**0.700**	**0.693**	**0.705**	**0.916**	**0.621**
C1	-0.246	**0.765**	**0.686**	**0.941**	**0.609**	**0.697**	0.496
C2	-0.100	**0.642**	0.483	**0.828**	**0.562**	**0.619**	0.461
C3	-0.297	**0.780**	**0.672**	**0.941**	**0.629**	**0.786**	0.474
C4	-0.268	**0.740**	0.596	**0.920**	0.574	**0.705**	0.396
D1	-0.241	**0.877**	**0.684**	**0.675**	0.573	**0.781**	**0.671**
D2	-0.304	**0.874**	**0.636**	**0.685**	**0.643**	**0.765**	0.595
D3	-0.309	**0.954**	**0.737**	**0.785**	**0.743**	**0.818**	**0.607**
D4	-0.172	**0.838**	**0.731**	**0.715**	**0.675**	**0.766**	0.523
E1	0.091	0.421	0.399	0.312	0.507	0.399	**0.659**
E2	0.040	0.595	0.476	0.418	0.472	0.555	**0.856**
E3	0.032	0.532	0.473	0.404	0.397	0.463	**0.855**
E4	0.065	0.597	0.500	0.461	0.531	0.541	**0.821**
F1	-0.067	**0.682**	**0.677**	0.566	**0.910**	**0.690**	**0.633**
F2	-0.165	**0.774**	**0.766**	**0.691**	**0.936**	**0.806**	0.589
F3	0.054	0.333	0.380	0.353	0.553	0.365	0.285
F4	-0.314	0.585	0.708	0.497	**0.844**	0.577	0.397
G1	0.589	-0.198	-0.118	-0.254	-0.054	-0.175	-0.039
G2	0.132	-0.084	-0.029	-0.089	-0.111	-0.087	-0.136
G3	-0.002	-0.115	0.021	-0.092	0.043	-0.082	-0.126
G4	-0.248	0.049	0.102	-0.050	0.067	0.018	-0.125
H1	-0.233	**0.675**	**0.904**	0.596	**0.686**	**0.616**	0.530
H2	-0.348	**0.775**	**0.920**	**0.655**	**0.786**	**0.738**	0.473
H3	-0.107	**0.691**	**0.897**	0.591	**0.668**	**0.698**	0.581

Analysis of The Structural Model

Analysis of the structural model to assess the construct reliability is based on the model predictive accuracy analyses. The model predictive accuracy involved coefficient determination (R^2), predictive relevance (Q^2), effect size (f2) and VIF (Collinearity Statistics).

i) Coefficient Determination (R^2)

This analysis followed the rule of thumb suggested by Hair et al. (2014), which indicated 0.75, 0.50, and 0.25 as substantial, moderate, and weak levels of predictive accuracy, respectively. As this model's R2 score was at behavioral intention (0.717), its predictive accuracy was considered substantial, as shown in Table 13.

Table 13: R^2 Score

	R Square
Behavioral Intention	0.717

ii) Predictive Relevance (Q^2)

The predictive relevance of the model was measured using the Q^2 value. A Q2 value between 0.02 and 0.15 indicates a weak effect, one between 0.15 and 0.35 indicates a moderate effect and one higher than 0.35 indicates a strong effect, as stated by Hair et al. (2012). Q^2 was 0.823 for behavioral intention. This indicates that the model has very high degree of predictive relevance, as the cutoff point was $Q^2 > 0$ (Cohen,1988).

iii) Effect Size

Cohen's f^2 analysis (Cohen, 1988) was used to evaluate the effect size of the predictor construct. This analysis followed the rule of thumb suggested by Cohen (1988), which considers large, medium, and small effect sizes of 0.35, 0.15, and 0.02, respectively. The result indicated that anxiety (0.012), effort expectancy (0.002), performance expectancy (0.002) and social influence (0.004) had a small effect on behavioral intention, while attitude (0.055) had a medium effect on behavioral intention. Facilitating conditions (0.249) had a large effect on behavioral intention. Table 14 displays the effect sizes of these variables.

Table 14: Effect size

	BEHAVIORAL INTENTION
ANXIETY	0.012
ATTITUDE	0.055
EFFORT EXPECTANCY	0.002
FACILITATING CONDITIONS	0.249
PERFORMANCE EXPECTANCY	0.002
SOCIAL INFLUENCE	0.004

iv) (Collinearity Statistics) VIF

Collinearity is critical for the assessment of the structural model. According to (Kock & Lynn, 2012), despite discriminant validity (vertical collinearity), the lateral collinearity issue (predictor criterion collinearity) may create misleading research findings. It was important to assess the VIF value of the predictor constructs to ensure that there was no multi-collinearity among the constructs. The VIF value cut off used in this study followed the recommendation of (Hair et al., 2010), which must be below 5. Any value higher than 5 indicated potential collinearity problems. Table 15 displays the constructs' VIF values. All constructs obtained values below six; thus, this structural model was free from collinearity problems.

Table 15: VIF value

	BEHAVIORAL INTENTION
ANXIETY	1.265
ATTITUDE	4.095
EFFORT EXPECTANCY	3.099
FACILITATING CONDITIONS	2.623
PERFORMANCE EXPECTANCY	3.478
SOCIAL INFLUENCE	2.220

FINDINGS AND DISCUSSION

To report the validity of the ChatGPT adoption study instrument adapting the Behavioral Intention Model

The face and content validity assessment, including face and content analysis, indicated that the instrument was in good order. The face analysis was giving feedback and recommendations for do the improvement for that questionnaire. For content analysis have two methods which are Scale Validity Index (CVI) and Scale level Content Validity Index (S-CVI). The calculation of I-CVI (Content Validity Index) according to Larsson et al. (2015) and Polit, Beck & Owen (2007) the I-CVI was calculated for both "relevant" and "clarity". Schilling, Dixon and Knafi, et al (2007) define I-CVI as the proportion of expert that provided a rating of 3 (very relevant) or 4 (extremely relevant) on the relevance scale and rating 3 (fairly clear) or 4 (very clear) on the clarify scale. The I-CVI cutoff as suggested by Lynn (1986) and Polit, Beck & Owen (2007) is a below than 0.78. As for this assessment, a conservative value of ≥.80 for both relevant and clarify are used as suggested by Paul, et al (2016). Items rated below 0.80 are eliminated. The S-CVI (Scale-level Content Index) of this study are calculated based on Polit, Beck & Owen (2007) and Davis (1992). Where they suggested the value of S-CVI should be greater than 0.8 or 80% or better agreement among reviewers. The S-CVI based on relevance and clarity for this study instrument are 0.98 and 0.96 and considered achieve "high-level agreement" which is acceptable and consistence. Overall, the instrument met all the analysis requirements for face and content validity assessment.

To report the reliability of the ChatGPT adoption study instrument, adapting the Behavioral Intention Model

The validity and reliability assessment, including measurement and structural model analysis, indicated that the instrument was in good order. The internal consistency value of Cronbach's alpha showed that all constructs' scores were between 0.718 and 0.965, indicating that the scale was highly reliable and that the items were highly related to the construct (Cronbach, 1988). Internal consistency measured by the composite reliability (CR) value was within the acceptable range, where the acceptable CR value should be between 0.70 and 0.90 (Gefen et al., 2000). The convergent validity value of AVE should account for at least 50 % of the indicator variance (AVE > 0.50) (Hair et al., 2014) to prove the construct sufficiently explains the indicator's variance. The AVE value of the constructs in this instrument showed a maximum 0.80 and a minimum of 0.10. Constructs with a low AVE value should be complemented with an acceptable CR value to be accepted (Ramayah & Chuah 2017). Discriminant validity, which measures the loadings of each indicator, requires that the loading of the individual indicator be higher in the designated construct compared to other constructs on a diagonal. The loadings are indicated by the Fornell–Larcker criterion and item cross-loading. Analyses of the instrument provided evidence of fulfilling the discriminant validity requirements. All constructs explained the variance of its indicator (high square root of the AVE in its indicator) and the high item cross-loading value on the assigned indicator variable compared to other variables. The measurement model had a satisfactory and acceptable value to meet the requirements of construct validity and reliability. Structural model analysis begins with a model-predictive relevancy assessment. The coefficient of determination (R^2) was used to measure the model predictive accuracy, which assessed the effect of the exogenous construct (independent variables) on the endogenous construct (dependent variables), and the acceptable effect range was between 0.75, 0.50 and 0.25 (Hair et al., 2014). Model R^2 values were considered substantial at behavioral intention (0.717). A predictive relevance analysis of Q^2 was used to compare the original value with the predictive value to calculate the predictive error. A model with a low predictive error has high predictive accuracy. The Q^2 value should be greater than zero in the endogenous construct. Based on the above requirement, endogenous constructs in the instrument were (0.823 = behavioral intention) indicated sufficient prediction relevancy values. The third analysis was effect size (f^2), which measured the relative impact or strength of the explanation of exogenous variables on the endogenous variables (Cohen, 1988).

The effect size of model prediction accuracy was rated for medium and small effect sizes. The collinearity statistical analysis of the variance inflation factor (VIF) used to identify each construct was assessed separately. VIF indicates the overlapping of variables when measuring the same construct. The value according to Hair et al. (2014) was VIF<5.0, indicating potential collinearity. The results of the analysis revealed no collinearity issues in the instrument because all the values obtained were less than 5. Overall, the instrument met all the analysis requirements for construct validity and reliability assessment.

CONCLUSION

The construct validity and reliability assessment of the instrument provided evidence of acceptable validity and reliability, and the structural model revealed a collinearity problem. Thus, the instrument is valid and reliable for measuring the adoption of ChatGPT among student in an academic environment. The instrument was developed based on well-established models UTAUT and been used in various research previously. This has methodologically contributed to a valid instrument for artificial intelligence and student learning behaviour research, which may benefit from the instrument in terms of developing practice guidelines for academic assessments and AI training for academic libraries.

ACKNOWLEDGEMENT

The authors sincerely thank the Perpustakaan Tun Abdul Razak (PTAR) UiTM for their kind support and cooperation in providing assistance in the data collection process and Pengurusan Air Selangor for the invaluable support in making this study possible.

REFERENCES

Abdullayeva, A., & Musayeva, S. (2023). The use of ChatGPT in higher education: Opportunities and challenges. *Journal of Educational Technology and Innovation, 6*(2), 45–56.

Ahmad, M. (2023). AI-powered personalized learning: Opportunities and challenges in education. *International Journal of Emerging Technologies in Learning, 18*(4), 25–37.

Albanna, S. (2023). Integrating generative AI in education: Opportunities and risks. *Education and Information Technologies, 28*(5), 6031–6048.

Aldasoro, J. C., Martín, J. A., & Sanz, A. (2019). Artificial intelligence in logistics and supply chain management: A review. *International Journal of Production Research, 57*(7), 2164–2182.

Ali, A., Khan, S., & Hussain, M. (2023). Exploring the role of ChatGPT in enhancing students' motivation toward learning. *International Journal of Educational Research, 122,* 102185. https://doi.org/10.1016/j.ijer.2023.102185

Augustine, O., & Ali, I. (2021). Artificial intelligence application in academic libraries in Nigeria. *Library Philosophy and Practice, 1–16.*

Ausat, A., Putra, D., & Nugroho, A. (2023). Enhancing learning engagement with ChatGPT: A case study in Indonesian universities. *Journal of Applied Learning and Technology, 5*(1), 77–88.

Balugani, E., Mandelli, A., & Vescovi, T. (2018). Artificial intelligence and big data in decision-making: Toward a new research agenda. *Journal of Business Research, 93,* 109–118.

Balugani, E., Mandelli, A., & Vescovi, T. (2019). AI-based solutions for business innovation: An integrative framework. *European Journal of Innovation Management, 22*(3), 423–440.

Benuyenah, V. (2023). The promise and peril of ChatGPT in higher education. *Journal of Higher Education Policy and Practice, 7*(1), 12–19.

Borji, A. (2023). A categorical archive of ChatGPT failures. *arXiv preprint arXiv:2302.03494.*

Breja, R., Sharma, D., & Singh, A. (2011). Narrow vs general AI: Applications and future. *International Journal of Computer Applications, 30*(8), 1–6.

Brown, J. (2013). Artificial intelligence and business creation. *Harvard Business Review, 91*(7), 56–62.

Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2nd ed.). Routledge.

Chaouali, W., Ben Yahia, I., & Souiden, N. (2016). The interplay of counter-conformity motivation, social influence, and trust in customers' intention to adopt internet banking services. *Journal of Retailing and Consumer Services, 28,* 209–218.

Chen, L., Xu, Q., & Zhang, C. (2019). Artificial intelligence applications in supply chain: A literature review. *Computers & Industrial Engineering, 137,* 106024.

Chin, W. W. (1998). The partial least squares approach to structural equation modeling. In G. A. Marcoulides (Ed.), *Modern methods for business research* (pp. 295–336). Lawrence Erlbaum Associates.

Choi, J., Lee, J., & Park, H. (2018). General AI and robotics: Bridging the gap. *Robotics and Autonomous Systems, 104,* 45–55.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum.

Colicchia, C., Creazza, A., & Dallari, F. (2019). Linking supply chain resilience and artificial intelligence: A conceptual model. *Supply Chain Management, 24*(3), 407–420.

Crawford, K., Paglen, T., & Whittaker, M. (2023). Generative AI in education: Risks of dependency and overreliance. *AI & Society, 38,* 923–936.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science, 35*(8), 982–1003.

De Sousa Jabbour, A. B. L., Jabbour, C. J. C., Foropon, C., & Filho, M. G. (2018). When titans meet—Can industry 4.0 revolutionize the environmentally-sustainable manufacturing wave? The role of critical success factors. *Technological Forecasting and Social Change, 132,* 18–25.

Deshpande, A., Kurhade, S., & Kale, V. (2018). Applications of AI in decision-making. *International Journal of Computer Science and Applications, 15*(2), 14–23.

Dolgui, A., Ivanov, D., & Sokolov, B. (2018). Ripple effect in the supply chain: An analysis and recent literature. *International Journal of Production Research, 56*(1–2), 414–430.

Dolgui, A., Ivanov, D., & Sokolov, B. (2019). Supply chain disruption and recovery: Considerations from supply chain viability perspective. *Annals of Operations Research, 283,* 711–726.

Dupont, J., Kim, H., & Lee, S. (2018). Artificial intelligence in autonomous driving: Recent developments. *IEEE Transactions on Intelligent Vehicles, 3*(2), 123–135.

Dwivedi, Y. K., et al. (2023). Adoption of ChatGPT: A research agenda. *International Journal of Information Management, 71,* 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Emon, M. M. H., Rahman, M., & Sultana, S. (2023). Factors influencing ChatGPT adoption behavior among professionals: A UTAUT model approach. *International Journal of Emerging Technologies in Learning, 18*(5), 32–46.

Fatorachian, H., & Kazemi, H. (2018). A critical investigation of industry 4.0 in manufacturing: Theoretical operationalisation framework. *Production Planning & Control, 29*(8), 633–644.

Felfel, H., Ben Yahia, W., Ayadi, O., & Masmoudi, F. (2018). Stochastic multi-site supply chain planning in textile and apparel industry under demand and price uncertainties with risk aversion. *Annals of Operations Research, 271*(2), 551–574. https://doi.org/10.1007/s10479-018-2980-2

Financial Times/New York Times. (2023). Google and Microsoft race to integrate AI chatbots into search engines. *The New York Times/Financial Times.*

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

Fuchs, C. (2023). ChatGPT and the new AI capitalism. *Monthly Review, 75*(2), 32–47.

Garza-Reyes, J. A. (2018). Green lean and the sustainable manufacturing agenda. *International Journal of Lean Six Sigma, 9*(2), 151–173.

Gefen, D., Straub, D. W., & Boudreau, M. C. (2000). Structural equation modeling and regression: Guidelines for research practice. *Communications of the Association for Information Systems, 4*(1), 1–70.

Gopalakrishnan, S. (2023). Challenges of large language models in specialized domains. *Journal of AI Research and Practice, 12*(3), 211–228.

Gunasinghe, A., Hamid, J. A., Khatibi, A., & Azam, S. M. F. (2020). Role of technology anxiety within UTAUT in understanding non-user adoption intentions to virtual learning environments: The state university lecturers' perspective. *Education and Information Technologies, 25*(6), 5803–5828. https://doi.org/10.1007/s10639-020-10209-z

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Pearson.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling (PLS-SEM)* (2nd ed.). SAGE Publications.

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.

Hamdi, F., Lachhab, M., & Bouslama, F. (2018). Expert systems in artificial intelligence: A literature review. *Expert Systems with Applications, 95,* 233–245.

Hazen, B. T., Boone, C. A., Ezell, J. D., & Jones-Farmer, L. A. (2018). Data quality for data science, predictive analytics, and big data in supply chain management. *International Journal of Production Economics, 193,* 91–99.

Hertzog, M. A. (2008). Considerations in determining sample size for pilot studies. *Research in Nursing & Health, 31*(2), 180–191.

Holmes, W., & Tuomi, I. (2022). State of AI in education: Trends and implications. *European Journal of Education, 57*(4), 567–583.

Hong, S. J., Thong, J. Y. L., & Tam, K. Y. (2008). Understanding continued information technology usage behavior: A comparison of three models in the context of mobile internet. *Decision Support Systems, 42*(3), 1819–1834.

Horng, J.-S., Liu, C.-H. S., Chou, S.-F., Tsai, C.-Y., & Hu, D.-C. (2018). Developing a sustainable service innovation framework for the hospitality industry. *International Journal of Contemporary Hospitality Management, 30*(1), 455–474. https://doi.org/10.1108/IJCHM-12-2015-0727

Hussain, I. (2020). Attitude of university students and teachers towards instructional role of artificial intelligence. *International Journal of Distance Education and E-Learning, 5*(2), 158-177. https://doi.org/10.36261/ijdeel.v5i2.1057

Ibrahim, N., & Tain, J. (2016). Structural equation modeling for instrument validation. *Journal of Social Science Research, 12*(4), 85–94.

Jackson, L. M. (2019). *The psychology of prejudice: From attitudes to social action* (2nd ed.). American Psychological Association. https://doi.org/10.1037/0000168-000

Kazancoglu, Y., Sezer, M. D., & Ozkan-Ozen, Y. D. (2018). Analyzing barriers to industry 4.0 in sustainable food supply chains. *Sustainability, 10*(11), 4080.

Keeble-Ramsay, D. R., & Armitage, A. (2010). Theorising the reflective learner: The role of reflection in work-based learning. *Journal of Workplace Learning, 22*(8), 505–516.

Kesharwani, A., & Singh Bisht, S. (2012). The impact of trust and perceived risk on internet banking adoption in India: An extension of technology acceptance model. *International Journal of Bank Marketing, 30*(4), 303–322.

Kock, N., & Lynn, G. S. (2012). Lateral collinearity and misleading results in variance-based SEM: An illustration and recommendations. *Journal of the Association for Information Systems, 13*(7), 546–580.

Kumar, S. (2019). Artificial intelligence in business strategy. *Journal of Business Strategy, 40*(6), 11–19.

Kumar, S., et al. (2018). Big data analytics in supply chain and logistics management. *Annals of Operations Research, 270,* 367–380.

Kumar, S., et al. (2019). Artificial intelligence and supply chain disruption management. *Computers & Operations Research, 111,* 65–77.

Kwak, D., Lee, H., & Lee, J. (2022). Examining user acceptance of AI-based learning systems. *Educational Technology Research and Development, 70*(2), 543–559.

Lai, M., & Adebayo, T. (2024). Ethical concerns and challenges of large language models in specialized fields. *AI Ethics, 4,* 215–230.

Larsson, H., Hedberg, J., & Nilsen, J. (2015). Measuring inter-rater reliability for content validity. *Nursing Research, 64*(3), 202–208.

Lu, X., & Yang, J. (2021). *Artificial intelligence and its role in education. Sustainability, 13*(22), 12902. https://doi.org/10.3390/su132212902

Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research, 35*(6), 382–386.

Martins, C., Oliveira, T., & Popovič, A. (2014). Understanding the internet banking adoption: A unified theory of acceptance and use of technology and perceived risk application. *International Journal of Information Management, 34*(1), 1–13.

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955. *AI Magazine, 27*(4), 12–14.

McKinsey & Company. (2022). *The state of AI in 2022.* https://www.mckinsey.com/featured-insights/mckinsey-on-books/state-of-ai

McKinsey & Company. (2023). *The state of AI in 2023.* https://www.mckinsey.com/featured-insights/mckinsey-on-books/state-of-ai

Menon, A., Singh, R., & Verma, S. (2023). Users' intention to adopt ChatGPT using UTAUT model. *International Journal of Emerging Technologies in Learning, 18*(6), 120–135.

Moor, J. (2006). The Dartmouth conference: The next 50 years. *AI Magazine, 27*(4), 87–91.

Nasim, S., Rehman, A., & Khan, I. (2022). AI adoption across industries: A systematic review. *Technological Forecasting and Social Change, 175,* 121345.

OpenAI. (2023). GPT-4 technical report. *arXiv preprint arXiv:2303.08774.*

Or, C. (2023). The role of attitude in the Unified Theory of Acceptance and Use of Technology: A meta-analytic structural equation modelling study. *Computers in Human Behavior, 145,* 107729. https://doi.org/10.1016/j.chb.2023.107729

Pathak, M. (2023). Awareness of ChatGPT among students: A gender-based analysis. *International Journal of Advanced Computer Science and Applications, 14*(5), 99–107.

Paul, J., Mittal, A., & Srivastav, G. (2016). Impact of validity measures in instrument development. *International Journal of Management Research, 14*(3), 112–124.

Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health, 30*(4), 459–467.

Rahman, H., Ismail, S., & Ali, N. (2016). Instrument validity in educational research. *Journal of Social Science and Humanities, 11*(2), 105–112.

Ramayah, T., & Chuah, F. (2017). *Partial least squares structural equation modeling (PLS-SEM) using SmartPLS 3.0.* Pearson Malaysia.

Richards, G., Yeoh, W., Chong, A. Y. L., & Popovič, A. (2019). Business intelligence effectiveness and corporate performance management: An empirical analysis. *Journal of Computer Information Systems, 59*(2), 188–196.

Sandu, R., & Gide, E. (2019). The role of Chatbots in education. *International Journal of Learning Technology, 14*(3), 185–202.

Schilling, L., Dixon, J. K., & Knafi, K. (2007). Determining content validity in nursing research. *Nursing Research, 56*(6), 441–445.

Shen, X., Chen, Q., & Liu, H. (2022). Student anxiety in adopting AI-based education platforms. *Computers & Education, 184,* 104495. https://doi.org/10.1016/j.compedu.2022.104495

Shoaib, M., Hussain, S., & Malik, A. (2021). *Artificial intelligence in education: A review of applications and challenges. Education Research International, 2021,* 8812542. https://doi.org/10.1155/2021/8812542

Shoufan, A. (2023). Exploring students' perceptions of ChatGPT: Thematic analysis and follow-up survey. *IEEE Access, 11,* 38805–38818. https://doi.org/10.1109/ACCESS.2023.3268224

Sohn, K., & Kwon, O. (2020). Technology acceptance theories and factors influencing artificial Intelligence-based intelligent products. *Telematics and Informatics, 47,* Article 101324. https://doi.org/10.1016/j.tele.2019.101324

Sreedharan, R., Kumar, R., & Menon, V. (2018). AI in decision science. *Journal of Decision Analytics, 15*(2), 112–123.

TechTRP. (2023). OpenAI ChatGPT Plus subscription in Malaysia: What you need to know. *TechTRP.* https://techtrp.com

The Star. (2023, February 15). OpenAI launches ChatGPT Plus in Malaysia for RM85/month. *The Star.* https://www.thestar.com.my

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,* Article 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.

Vieriu, D., & Petrea, R. (2025). Artificial intelligence and higher education: Impacts on learning and research. *Journal of Educational Futures, 10*(1), 1–15.

Wang, X., Li, Y., Zhang, Z., & Chen, L. (2024). Pre-Service Teachers' GenAI Anxiety, Technology Self-Efficacy, and TPACK: Their Structural Relations with Behavioral Intention to Design GenAI-Assisted Teaching. Behavioral Sciences, 14(5), 373.

Winner, L. (2009). Do artifacts have politics? *Daedalus, 109*(1), 121–136.

Yusoff, N. M. R. N., Jamaludin, K. A., & Alias, N. (2022). *Benefits and challenges in implementing artificial intelligence in education (AIED) in ESL classrooms: A systematic review (2019–2022). International Journal of Academic Research in Business and Social Sciences, 12*(12), 2314–2331. https://doi.org/10.6007/IJARBSS/v12-i12/20422

Zhai, C. (2022). Implications of ChatGPT in higher education: A literature review. *Educational Review, 74*(5), 1–15.

Zhai, X., Chu, X., & Chai, C. S. (2021). Integrating artificial intelligence into education: Trends and challenges. *Educational Technology Research and Development, 69*(4), 1–23.

Zhou, J., Li, X., & Zhang, H. (2023). Examining ChatGPT adoption using TAM and TPB. *Journal of Educational Computing Research, 61*(7), 1583–1601