https://doi.org/10.62663/pjpprp.v16i2.256

Development and Validation of ChatGPT Reliance Scale

Syeda Farhana Kazmi, Owais Ahmad, & Shamsa Siddique
Department of Psychology, Hazara University, Mansehra, Pakistan

As ChatGPT becomes more common in universities, many people are concerned about its effects on students' learning, research quality, originality, and even procrastination habits. In response, this study developed and validated a detailed ChatGPT Reliance Scale for university students. The goal was to provide a reliable and accurate tool for capturing the complex ways students use ChatGPT in higher education. To this end, we conducted a cross-sectional survey of 1,000 students from various universities and postgraduate colleges in Abbottabad and Peshawar. An exploratory factor analysis (EFA) was performed on the first segment of the dataset, comprising 409 participants. The EFA indicated a three-factor structure that explained 61.99% of the total variance. A confirmatory factor analysis (CFA) was then conducted on the remaining 400 participants, supporting the three-factor model and yielding acceptable fit indices (χ^2 (235) = 1074, $p < .001$; CMIN/DF = 4.57; CFI = 0.91; GFI = 0.91; RMSEA = 0.06). All factor loadings were statistically significant ($p < .001$) and ranged from 0.54 to 0.85. Overall, the scale demonstrated strong internal consistency (Cronbach's α = 0.92, McDonald's ω = 0.93), suggesting it is a reliable instrument for assessing how and to what extent university students rely on ChatGPT.

Keywords: ChatGPT, Higher Education, Reliance, Novelty, Procrastination, Artificial Intelligence.

Generative artificial intelligence (AI) is becoming part of everyday life, reshaping how people think, work, and learn. One of the best-known tools is OpenAI's ChatGPT, widely used to generate human-like text, solve problems, summarize information, and provide interactive support across a range of domains (Zhou & Lee, 2024). As professionals and college students increasingly use ChatGPT in their academic and professional lives, it is increasingly important to understand how it affects their thoughts, feelings, and behavior.

ChatGPT provides university students with quick access to resources that can enhance efficiency and learning, including summaries, writing assistance, explanations, and coding support. However, this ease of use may lead to overreliance when students substitute AI for cognitive participation (Rudolph et al., 2024). In a similar spirit, employees, particularly knowledge workers, use ChatGPT to compose emails, prepare reports, give presentations, and brainstorm ideas. Workflows may be streamlined, but over time, opportunities for critical thinking and autonomous problem-solving may diminish. This growing reliance raises important concerns about users' relationships with generative AI with respect to productivity, autonomy, and self-regulation.

Reliance is a degree of dependence, and reliance on ChatGPT reflects the extent to which people rely on it to perform assigned tasks. Reliance on ChatGPT is fundamentally distinct from the frequency of use. It concerns the importance of ChatGPT in helping students complete an assignment. One student may routinely use ChatGPT alongside a literature search to develop assignment ideas, while another may use it only occasionally. However, for the latter student, ChatGPT may play a more important role in idea generation. Reliance on ChatGPT thus concerns the extent to which students regard it as fundamental to their learning activities (Stojanov et al., 2024). ChatGPT reliance entails regularly relying on ChatGPT for tasks you could do on your own. Although this behavior is shaped by ChatGPT's conversational and

[1]Correspondence concerning this article should be addressed to Mr. Owais Ahmad, Department of Psychology, Hazara University, Mansehra, Pakistan. Email: owaisak55@gmail.com

https://doi.org/10.62663/pjpprp.v16i2.256

generative design, it still aligns with what we know about cognitive offloading—using technology to think or remember for us (Risko & Gilbert, 2016; Rudolph et al., 2024).

Novelty, or the extent to which users are attracted to ChatGPT because of its perceived originality or distinctiveness, is a second dimension. This novelty can lead to heightened, sometimes shallow, use behaviors, especially among professionals and students using AI tools for the first time (Shin, 2021). A third and crucial element is procrastination, which occurs when users engage with ChatGPT excessively, either as a distraction or to avoid work or decision-making, thereby creating the appearance of productivity (Meier et al., 2023). According to one study, when people have a heavy workload, they are more likely to adopt and use technology. When faced with a heavy workload, students typically seek ways to cope. As a result, people use simple strategies or shortcuts (e.g., ChatGPT) to manage such stressful circumstances (i.e., excessive workload) (Abbas et al., 2024).

Currently, there is no well-established psychometric tool that systematically examines how university students and working adults use ChatGPT across these three areas. This gap persists despite growing concerns about the use of AI in academic and professional contexts. The present study aims to address this need by developing a multidimensional scale comprising three subscales: Reliance, Novelty, and Procrastination. The scale is intentionally designed to fit academic and workplace contexts and draws on ideas from digital behavior, cognitive psychology, and human–AI interaction.

This scale has practical value as well as research significance. In educational settings, it can help teachers and administrators understand whether students are using AI to replace their own effort and understanding or to support and extend their learning. When students exhibit high levels of dependence or procrastination, the scale can help identify them, enabling more targeted guidance or interventions. In professional environments, the scale can inform organizational decisions about productivity, digital well-being, and the responsible use of AI. Understanding the extent to which employees rely on AI can help managers distinguish between situations in which AI genuinely supports productivity and those in which it may undermine independent thinking or sound judgment.

From a research perspective, this tool provides a starting point for studying how the use of generative AI relates to outcomes such as self-regulation, digital literacy, learning performance, and job satisfaction. As AI systems become increasingly embedded in academic and professional routines, a reliable, well-validated measure of how people engage with ChatGPT is important. Such a measure can support efforts to promote ethical, effective, and balanced forms of collaboration between humans and AI.

Extended Mind Theory

This idea suggests that our thinking does not occur solely in our heads. It can also include the tools and technologies we use in our everyday environment. In this sense, ChatGPT can function as an extension of the mind, helping us generate new ideas by serving as an external semantic network or a brainstorming partner.

Runco and Jaeger (2012) describe creativity as the production of ideas that are both new and valuable. According to Extended Mind Theory, when someone uses a system such as ChatGPT as a thinking partner—by offloading memories, exploring possibilities, or seeking unexpected links—the interaction becomes part of that person's own cognitive process. Because ChatGPT draws on a large and varied body of training data, it can offer ideas that spark connections users might not reach on their own. Recent work also suggests that using AI

https://doi.org/10.62663/pjpprp.v16i2.256

language models can enhance creative self-perception and support idea generation in tasks such as writing and problem-solving (Yeo, 2023).

Self-Determination Theory adds another layer to this discussion. It proposes that people are more motivated when they see themselves as competent and effective. ChatGPT can affect this in two different ways. On the one hand, it can help users understand complex concepts and complete tasks more efficiently, which may strengthen their sense of capability. On the other hand, if individuals routinely avoid effortful learning and lean too heavily on AI-produced answers, their skills may develop more slowly, and their competence may become fragile and dependent on access to AI (Deci & Ryan, 2017). In such cases, individuals may begin to credit the AI rather than their own abilities for successful outcomes, thereby undermining their sense of self-efficacy.

Dual-Process Theory (Thinking in System 1 and System 2)

According to Evans and Stanovich's (2013) Dual-Process Theory of Cognition, human thought processes are mediated by two systems, i.e., System 1: Fast, intuitive, and automated thought; and System 2: Analytical, laborious, and slow thinking.

System 1 is particularly well served by ChatGPT, which provides quick, easily assimilated outputs with minimal cognitive strain. Regular usage of ChatGPT may eventually train users—both professionals and students—to rely on fast AI-assisted responses rather than System 2's laborious cognitive processes. As users increasingly delegate problem-solving to AI, they become less engaged in deeper learning or analytical reasoning, resulting in greater reliance on AI (Rudolph et al., 2024).

Extension of the Technology Acceptance Model (TAM)

Davis's (1989) Technology Acceptance Model identifies perceived usefulness and perceived usability as the primary determinants of new technology acceptance. Reliance on ChatGPT is fueled by perceived usefulness ("ChatGPT helps me get things done more efficiently").

Engagement is increased by interactive novelty and ease of use, especially in the early phases of exposure.

To account for novelty-driven usage, newer versions of TAM, such as TAM 3 (Venkatesh & Bala, 2008), also consider hedonic motivation, curiosity, and habit. Because of its captivating design, ChatGPT's uniqueness as a state-of-the-art conversational and generative tool increases initial attraction and sustained use (Shin, 2021).

Theory of Cognitive Offloading

Risko and Gilbert (2016) describe cognitive offloading as the act of reducing the mental effort we expend internally by turning to external tools, including digital technologies. In many ways, ChatGPT functions like an always-available external brain: users employ it to assist with tasks such as reading, generating summaries, brainstorming ideas, and drafting emails or written responses.

While this kind of offloading can make work feel easier and more efficient, it may also lead to growing dependence and habitual use. When people rely on ChatGPT too often rather than engaging directly with their tasks, this can encourage procrastination and a form of digital avoidance (Meier et al., 2023). To examine these patterns more systematically, the ChatGPT Reliance Scale (CRS) has been developed as a psychometric tool that captures how and to what

https://doi.org/10.62663/pjpprp.v16i2.256

extent users rely on ChatGPT across cognitive, behavioral, and emotional domains. In this context, reliance refers to the extent to which students perceive ChatGPT as necessary for their learning activities. Their level of knowledge about the tool and their attitudes toward it not only shape whether they adopt it but also how frequently they use it. More broadly, in the field of artificial intelligence (AI), students' AI literacy and their views about ChatGPT help predict how they use it and, in turn, how strongly they come to depend on it for tasks (Uppal & Hajian, 2024).

The CRS, therefore, offers a way to assess patterns of reliance on ChatGPT, and these patterns have important consequences. In educational settings, higher scores on this scale may be linked to weaker self-regulation, lower academic engagement, and less robust learning outcomes. In professional settings, heavy reliance on ChatGPT for routine or instrumental tasks may boost productivity in the short term, but over time, it could limit opportunities for skill development and independent decision-making.

Method

A cross-sectional survey methodology was used to verify the ChatGPT reliance scale. The participants in the study were from higher education institutions, such as universities and postgraduate colleges. Data were collected from nearly 1000 individuals across various universities and doctoral colleges in Abbottabad and Peshawar. This resulted in a response rate of 22.22%, which surpassed the minimum of 10 participants per item required for scale development and validation. After removing incomplete and missing data, the final sample included 809 participants- 500 females (61.8%) and 309 males (38.2%).

The participants' mean age was 21.96 years (SD = 2.21; range = 18-26). The study required that participants be enrolled in a higher-degree program, such as a bachelor's or master's program. Furthermore, students from schools and intermediate colleges were excluded from the study because they do not have regular access to ChatGPT for their studies and assignments.

Measures

The ChatGPT Reliance scale, designed and validated exclusively for this study, served as the study's primary instrument. The Scale comprises 45 items in total, assessing three aspects of students: reliance, novelty, and procrastination. Items were developed based on different theories mentioned in the introduction, experts in the fields had formal discussions, and those discussions led to an item pool with each subscale consisting of 15 items. All items were scored on a 5-point Likert scale, with responses ranging from strongly disagree to strongly agree. Scores are interpreted as follows: higher scores indicate greater reliance, novelty, and procrastination.

Procedure

Data were collected from participants within their university departments. Hard copies of questionnaires were delivered to participants, who were informed that there were no correct or incorrect responses. The researchers also advised them that participation in the study is entirely voluntary and that any participant may withdraw at any time. Furthermore, they were instructed to complete the questionnaire regarding their use of ChatGPT. They were also taught that it is the participants' ethical and moral obligation to fill out the questionnaire honestly and completely. The data-gathering procedure took more than two months because it was difficult to travel, obtain official authorization, and collect data from the appropriate pupils.

https://doi.org/10.62663/pjpprp.v16i2.256

Furthermore, after collecting data from several institutions in Abbottabad and Peshawar, the data were cleaned, and incomplete, pattern-filled questionnaires were excluded, resulting in 809 participants' data being considered accurate. The research reported a 19% attrition rate.

Data Analysis

Following data clarification, the data were analyzed using SPSS version 30, and AMOS 23 was used to conduct the CFA. The data analysis procedure consisted of three major phases.

The first phase involved checking the data's normality, analyzing its frequency distributions, and calculating the scale's overall reliability. Similarly, the overall convergent validity of the scale was assessed using two distinct measures of the same construct. Convergent validity for each subscale was also evaluated using multiple scales that measure the same components. Furthermore, the discriminant validity of the research measure was evaluated using two independent scales assessing the opposite construct. All reliabilities and validities were estimated to be at or above 0.85.

In the next step, the data were divided into two segments of 409 and 400 participants for EFA and CFA, respectively. Exploratory factor analysis (EFA) was used to determine the factorial validity of the 45-item scale. Cross-loading and items with poor loading are permitted. Principal component analysis with varimax rotation was employed. After the EFA, items with good factor loadings and a discrimination index of more than .35 were included in the final version of the scale, which became a 24-item scale in total, comprising 8 items per subscale.

Following EFA, the next stage was Confirmatory factor analysis (CFA) to assess model fit and the psychometric properties of the final 24-item scale. The sample comprised 400 participants; segment two was separated from the total sample of 809 participants. The goodness-of-fit indices chi-square, CFI, GFI, TLI, RMSEA, and SRMR were examined. Acceptable model fit criteria were CFI, GFI, and TLI > 0.90 and RMSEA < 0.08.

Results

The original 45-item ChatGPT Reliance Scale underwent an exploratory factor analysis (EFA). Based on the outcomes of this inquiry, the scale was modified to its final 24-item format. The 24-item scale was then further investigated. Table 3 presents the factor loadings, means, standard deviations, and item-total correlations for the 24 items of the scale. The EFA, which included principal component analysis and Varimax rotation, revealed a three-factor structure accounting for 61.99% of the total variance.

Table 1

Demographic Features of Sample Measuring ChatGPT Reliance (N = 809).

Demographics	Frequency	%	M	SD
Age			21.96	2.21
Gender				
Male	309	38.2		
Female	500	61.8		
Current Semester of students				
1st	98	12.11		
2nd	80	9.88		

https://doi.org/10.62663/pjpprp.v16i2.256

3rd	67	8.28
4th	106	13.10
5th	125	15.45
6th	153	18.91
7th	101	12.48
8th	79	9.76

https://doi.org/10.62663/pjpprp.v16i2.256

Table 2

Communalities Estimates of Extraction Order Via Principal Factors Analysis of ChatGPT Reliance-24 (N = 409).

Item numbers	Value
1	.69
2	.68
3	.50
4	.63
5	.47
6	.54
7	.57
8	.56
9	.80
10	.65
11	.59
12	.66
13	.65
14	.60
15	.65
16	.63
17	.81
18	.78
19	.73
20	.69
21	.76
22	.69
23	.61
24	.63

https://doi.org/10.62663/pjpprp.v16i2.256

Table 3

Factor Loading of 24-Item ChatGPT Reliance Scale (N= 409)

Items	Reliance	Novelty	Procrastination	M	SD	Item-total correlation
I use ChatGPT for tasks that I can complete independently.	.72			3.31	1.14	.54**
I frequently use ChatGPT to solve difficulties rather than attempting to figure out myself.	.69			3.19	1.10	.48**
I trust ChatGPT's recommendations more than traditional sources.	.63			3.11	1.17	.43**
I usually double-check my work using ChatGPT more frequently than required.	.62			3.11	1.07	.57**
I prefer ChatGPT's ideas over my thinking.	.60			3.02	1.12	.39**
I frequently seek ChatGPT's assistance with tasks that I am capable of completing on my own.	.60			3.04	1.09	.52**
I rely on ChatGPT to confirm information even when I already know the answer.	.59			2.98	1.18	.50**
I feel hesitant about making decisions without reaching out to ChatGPT.	.57			2.94	1.17	.48**
ChatGPT gives me to new viewpoints that I had not explored before.		.80		3.64	1.02	.57**
I believe that ChatGPT improves my capacity to think beyond the box.		.77		3.59	1.01	.51**
ChatGPT enables me to find connections between concepts that I hadn't recognized previously.		.68		3.57	.95	.49**
ChatGPT allows me to develop new and creative content more effectively.		.60		3.67	.98	.49**
I feel more productive after using ChatGPT to brainstorm or produce ideas.		.59		3.58	1.00	.42**
ChatGPT frequently surprises me with		.58		3.67	1.06	.47**

[1]Correspondence concerning this article should be addressed to Mr. Owais Ahmad, Department of Psychology, Hazara University, Mansehra, Pakistan. Email: owaisak55@gmail.com

fresh ideas.						
I utilize ChatGPT to help me build new problem-solving strategies.		.57		3.65	.89	.54**
I have used ChatGPT to experiment with various forms or frameworks for my tasks.		.54		3.57	1.00	.47**
Rather than focusing on my work, I usually examine unrelated things on ChatGPT.			.85	2.87	1.16	.55**
Sometimes, I spend more time engaging with ChatGPT than doing my actual tasks.			.84	2.88	1.13	.56**
I use ChatGPT to postpone starting necessary assignments.			.79	2.85	1.16	.58**
I occasionally tell myself that I'm being productive while using ChatGPT, even when I'm not.			.78	2.94	1.14	.48**
I rely on ChatGPT to produce replies, even when I could write them myself.			.75	2.91	1.16	.46**
I feel less motivated to engage in independent academic tasks because chatGPT provides quick answers.			.73	3.01	1.17	.42**
I find myself asking ChatGPT irrelevant questions instead of doing my work.			.59	2.81	1.15	.51**
My use of ChatGPT has caused delays in completing assignments and tasks.			.55	2.96	1.12	.46**
% Variance	25.08					
Cumulative Variance	61.99					
Kaiser-Meyer-Olkin measure	.87					
Bartlett’s test of Sphericity	18463.83<.001					

https://doi.org/10.62663/pjpprp.v16i2.256

Table 4

Ideal Fit Indices Intended for ChatGPT Reliance Scale-24 (N = 400)

			Goodness of Fit				
Representations	X^2(df)	X^2/df	GFI	AGFI	CFI	IFI	RMSEA
ChatGPT Reliance (Three factors)	1074(235)	4.57	.90	.91	.91	.91	.06

Table 4 presents the CFA results for the goodness-of-fit index (GFI) for the scale's three subscales. Appropriate figures for GFI (goodness-of-fit index), AGFI (adjusted goodness-of-fit index), CFI (comparative fit index), and IFI (incremental fit index) are .90 or higher, and the present analysis gained scores that are higher than .90. As a result, the present research provides substantial results. Cronbach's α, composite reliability, and convergent validity for the three factors are all excellent, as shown in Table 5.

Table 5

The CFA Reliability and Validity Outcomes for the Concluding Model of ChatGPT Reliance Scale (N = 409).

Variables	items	Factor loading	α (>.7)	CR (>0.6)	AVE (>0.5)
Factor 1	1	.72	.84	.83	.39
	2	.69			
	3	.63			
	4	.62			
	5	.60			
	6	.60			
	7	.59			
	8	.57			
Factor 2	9	.80	.86	.84	.41
	10	.77			
	11	.68			
	12	.60			
	13	.59			
	14	.58			
	15	.57			
	16	.54			
Factor 3	17	.85	.90	.90	.62
	18	.84			
	19	.79			
	20	.78			
	21	.75			
	22	.73			
	23	.59			
	24	.55			

https://doi.org/10.62663/pjpprp.v16i2.256

Table 6

Correlation among Subscales of ChatGPT Reliance Scale (N = 809).

Scales	Reliance	Novelty	Procrastination	Total
Reliance		.52**	.48**	.83**
Novelty			.36**	.78**
Procrastination				.78**

Note. ** = Correlations are significant at the 0.01 level.

Table 7

Gender Differences on ChatGPT Reliance Scale (N= 809)

	Boys (n = 309)		Girls (n = 500)				
Variable	*M*	*SD*	*M*	*SD*	*t* (807)	p	Cohen's d
ChatGPT Reliance	155.33	17.48	144.08	25.17	5.62	.00**	.47

Note. M = Mean, SD = Standard Deviation.

Table 8

Convergent Validity of ChatGPT Reliance with ChatGPT Usage, Positive Attitudes Towards Artificial Intelligence, Creativity and Procrastination Scales (N = 220)

Scales	ChatGPT reliance	ChatGPT usage	Positive attitudes towards AI	Creativity	Procrastination
ChatGPT reliance		.85**	.48**	.77**	.53**
ChatGPT usage			.57**	.85**	.56**
Positive attitudes towards AI				.26**	.41**
Creativity					.44**
Procrastination					

Note. ** = Correlations are significant at the 0.01 level.

Table 9

Discriminant Validity of ChatGPT Reliance with Digital Well-Being and Negative Attitudes Towards Artificial Intelligence Scales (N = 220)

Scales	ChatGPT reliance	Negative attitudes towards AI	Digital Well-Being
ChatGPT reliance		-.82**	-.85**
Negative attitudes towards AI			-.80**
Digital Well-Being			

Note. ** = Correlations are significant at the 0.01 level.

Item–total correlations were calculated to assess the extent to which each item was associated with the overall scale score. For items for which this information was available, correlations ranged from .39** to .58**, indicating moderate to strong relationships between individual items and the total scale. This pattern suggests that the items align well with the construct the scale is intended to measure.

The internal consistency of the scale was excellent, with Cronbach's alpha of .92 and McDonald's omega of .93. Data were gathered on several occasions and across multiple related measures from more than 220 participants to examine convergent and discriminant validity

https://doi.org/10.62663/pjpprp.v16i2.256

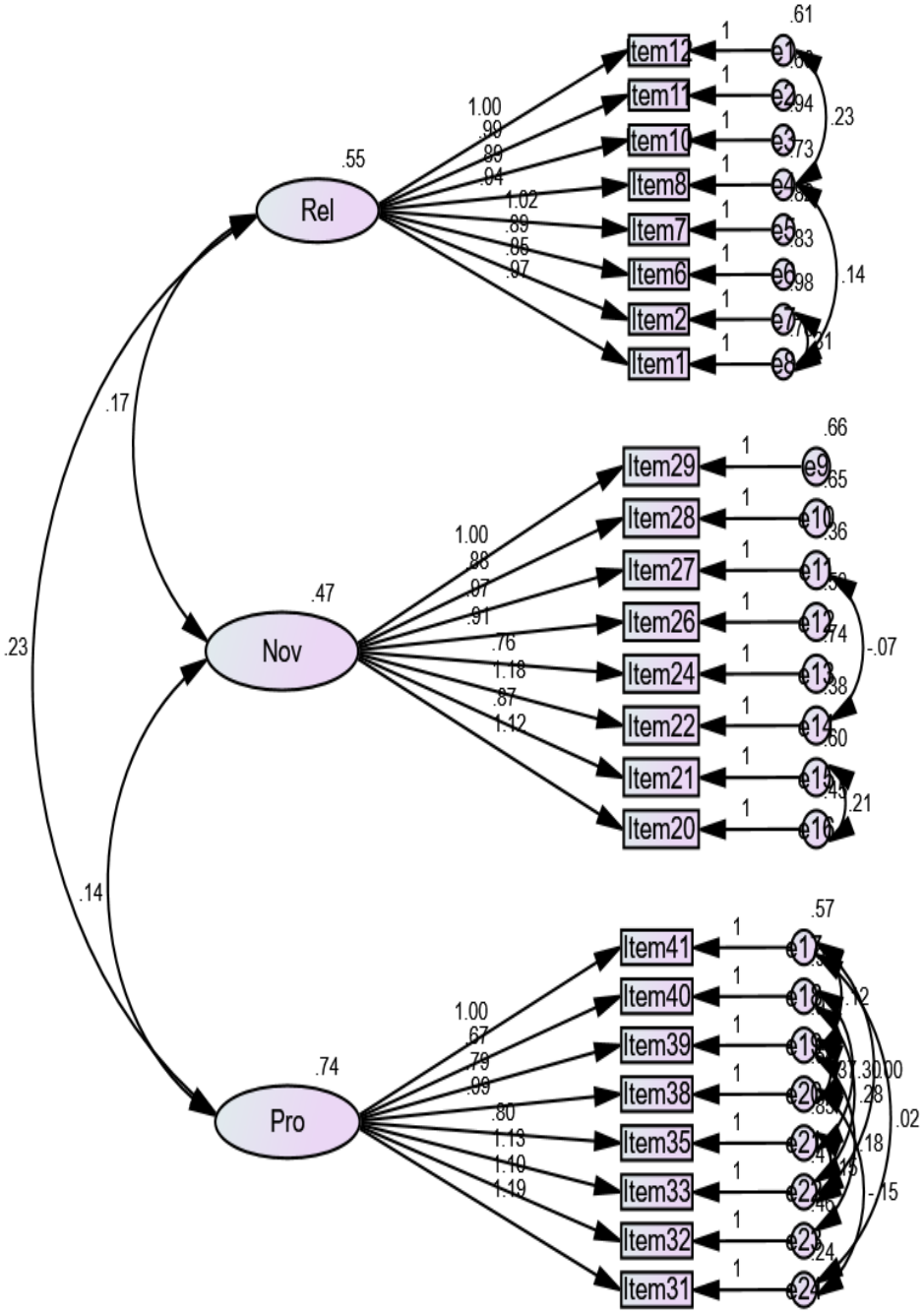
thoroughly. Convergent validity with the ChatGPT Usage Scale was strong (r = .85**). Convergent validity was also assessed using the General Attitude Towards Artificial Intelligence Scale, which includes two subscales: Positive Attitudes and Negative Attitudes. The ChatGPT Reliance Scale correlated positively with Positive Attitudes (r = .48**, p < .001) and negatively with Negative Attitudes (r = −.82**, p < .001), providing clear support for both convergent and discriminant validity.

Convergent validity was then examined at the subscale level. The Reliance subscale showed a positive correlation of .56** with a related construct. The Novelty subscale was positively associated with a creativity measure (r = .59**). The Procrastination subscale was positively related to an established student procrastination scale (r = .48**). Discriminant validity was further supported by a strong negative correlation with the Digital Well-Being Scale (r = −.85**). Overall, the reliability and validity indices were strong and comfortably exceeded commonly accepted benchmarks.

The three-factor structure identified in the exploratory factor analysis (EFA) was further tested using confirmatory factor analysis (CFA). In this model, the three components—Reliance (REL), Novelty (NOV), and Procrastination (PRO)—were specified as latent variables, and their respective items were treated as observed indicators (Fig. 1). The CFA results indicated a satisfactory model fit: $\chi 2(235) = 1074$, $p < .001$; CMIN/DF = 4.57; CFI = .91; TLI = .91; RMSEA = .06 (90% CI: .05–.066). All standardized factor loadings were statistically significant ($p < .001$).

Taken together, the EFA and CFA findings support a three-factor structure for the ChatGPT Reliance Scale—Reliance, Novelty, and Procrastination—and show that the scale has strong psychometric qualities, including good model fit, substantial factor loadings, and high levels of reliability and validity.

https://doi.org/10.62663/pjpprp.v16i2.256



https://doi.org/10.62663/pjpprp.v16i2.256

Discussion

This study aimed to develop and test a scale to assess how university students use, perceive, and experience ChatGPT. The analyses support a three-part structure for the ChatGPT Usage Scale, reflecting dependence, novelty, and procrastination. Overall, the scale shows solid psychometric qualities: the model fits the data well, factor loadings are strong, and reliability is high. Together, these findings suggest that the scale is a valuable tool for examining students' engagement with ChatGPT.

The first factor, Reliance, reflects the extent to which students rely on ChatGPT and their perceptions of this technology. It includes items assessing the extent to which they trust ChatGPT's answers, how often they use it for feedback, and their awareness of its limitations. This aligns with prior work on technology use and confidence in automated systems. In this context, reliance refers to the extent to which students see ChatGPT as central to their learning tasks (Stojanov et al., 2024). Although shaped by ChatGPT's interactive and generative nature, it also aligns with research on cognitive offloading, in which people shift mental work to external tools (Risko & Gilbert, 2016; Rudolph et al., 2024). Factor loadings for Reliance ranged from .57 to .72, indicating that many university students lean on ChatGPT for assignments and projects.

The second factor, Novelty, captures the extent to which working with ChatGPT encourages new ideas, perspectives, or approaches. Novelty tends to be higher when people use this type of tool for the first time, among both students and professionals (Shin, 2022). Items in this dimension assess whether ChatGPT helps students think in new ways or produce work that appears more original. Factor loadings for Novelty ranged from .54 to .80. These results are consistent with a preregistered study in which 77 university students completed ill-defined tasks: students with access to ChatGPT produced answers that were rated as more detailed (d = 0.61), higher in quality (d = 0.69), and more original (d = 0.55) than those in the control group (Urban et al., 2024).

Other work has examined creativity over longer periods. Liu et al. (2024) followed participants for 7 days and conducted a 30-day follow-up. Performance initially improved when ChatGPT was available but returned to baseline after its removal. The study also noted that ideas became increasingly similar over time—a pattern that persisted even after use ceased (Liu et al., 2024). In another study, Toma and Yáñez-Pérez (2024) tracked 31 students over 10 weeks. A little over half (53.6%) showed gains in creativity, 25% showed declines, and 21.4% showed no change. Taken together, these findings suggest that using ChatGPT does not necessarily harm creativity overall, though its impact may vary from person to person.

The third factor, Procrastination, addresses situations in which students turn to ChatGPT not to advance their work but to avoid it. In such cases, they may spend time exploring the tool, asking unnecessary questions, or repeatedly refining prompts, while feeling productive. This pattern aligns with earlier research showing that people under heavy workload are more likely to adopt technologies as a coping mechanism, sometimes by seeking shortcuts rather than tackling tasks directly (Meier et al., 2023; Reinecke et al., 2021; Abbas et al., 2024). In this study, the Procrastination factor was designed to capture the extent to which students delay their work because ChatGPT offers quick answers and ready-made solutions. As students become accustomed to instant help, they may delay initiating or completing tasks, assuming they can complete them later with the tool's support. Factor loadings for this dimension ranged from .55 to .85. These results are consistent with theoretical arguments that the ease and speed of such tools can foster a ‘‘shortcut’’ mentality, in which students routinely postpone work because they believe they can complete it quickly later (Yilmaz, 2023).

https://doi.org/10.62663/pjpprp.v16i2.256

Evidence from Pakistan reinforces this concern. An extensive survey of 494 participants, reported in Fast Company, found that heavy users of ChatGPT reported more procrastination, memory difficulties, and even lower GPAs. Another study with 113 university students in Pakistan found a significant positive association ($r = .241$, $p = .010$) between reliance on AI-based helpers (measured with the DAI scale) and self-reported procrastination. Mukhtar et al. (2025) therefore argue that overreliance on such technologies can lead to increased delays in work.

The three-factor model showed a good overall fit, as indicated by confirmatory factor analysis (CFI = 0.91; TLI = 0.91; RMSEA = 0.06). These indicators support the construct validity of the scale. In addition, reliability estimates were strong: composite reliability was 0.92, and McDonald's omega (ω) was 0.93, suggesting that the items within each factor measure closely related aspects of the same underlying constructs. When combined with evidence of content validity from the scale development process and factorial validity from both exploratory and confirmatory factor analyses, these findings support the ChatGPT Usage Scale as a sound and practical instrument for future research.

At the same time, several limitations need to be acknowledged. First, the sample was drawn from university and postgraduate students in Abbottabad and Peshawar. This focus may limit the extent to which the results can be generalized to other regions, cultures, or educational systems. Future research should test and refine the scale with students from a broader range of institutions and backgrounds to establish its suitability across contexts. Second, the study relied on self-report, which is always vulnerable to social desirability and other response biases—especially in an area where ethical questions about the use of AI in academic work are still actively debated. Later studies could address this by adding behavioral indicators, such as actual usage logs, or by combining self-reports with performance-based measures.

Finally, tools like ChatGPT are developing rapidly. New versions bring new features, interaction styles, and use cases. For this reason, the scale should not be viewed as fixed. It will need to be updated and revalidated regularly to ensure that it continues to reflect how students use these tools in their studies, and to enable researchers to track changes in patterns of dependence, novelty seeking, and procrastination over time.

References

Abbas, M., Jam, F. A., & Khan, T. I. (2024). Is It Harmful or helpful? Examining the Causes and Consequences of Generative AI Usage among University Students. *International Journal of Educational Technology in Higher Education*, *21*(1). https://doi.org/10.1186/s41239-024-00444-7

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319–340. https://www.jstor.org/stable/249008

Deci, E. L., & Ryan, R. M. (2017). Self-Determination theory in work organizations: The state of a science. *Annual Review of Organizational Psychology and Organizational Behavior*, *4*(1), 19–43. https://doi.org/10.1146/annurev-orgpsych-032516-113108

Evans, J. St. B. T., & Stanovich, K. E. (2013). Dual-Process Theories of Higher Cognition: Advancing the debate. *Perspectives on Psychological Science*, *8*(3), 223–241. https://doi.org/10.1177/1745691612460685

Liu, Q., Zhou, Y., Huang, J., & Li, G. (2024). When ChatGPT is gone, creativity reverts and homogeneity persists. *ArXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2401.06816

Meier, A., Ine Beyens, S., T., J. Loes Pouwels, & Valkenburg, P. M. (2023). Habitual social media and smartphone use is associated with task delay among some, but not all, adolescents. *Journal of Computer-Mediated Communication*, *28*(3). https://doi.org/10.1093/jcmc/zmad008

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, *20*(9), 676–688. https://doi.org/10.1016/j.tics.2016.07.002

Rudolph, J., Mohamed, M., & Popenici, S. (2024). Higher Education's Generative Artificial Intelligence Paradox: The Meaning of Chatbot Mania. *Journal of University Teaching & Learning Practice*, *21*(06). https://doi.org/10.53761/54fs5e77

Runco, M. A., & Jaeger, G. J. (2012). The Standard Definition of Creativity. *Creativity Research Journal*, *24*(1), 92–96. https://www.tandfonline.com/doi/abs/10.1080/10400419.2012.650092

Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies*, *146*(146), 102551. https://doi.org/10.1016/j.ijhcs.2020.102551

Stojanov, A., Liu, Q., & Hwee, J. (2024). University students' self-reported reliance on ChatGPT for learning: A latent profile analysis. *Computers and Education. Artificial Intelligence*, *6*, 100243–100243. https://doi.org/10.1016/j.caeai.2024.100243

Toma, R. B., & Yánez-Pérez, I. (2024). Effects of ChatGPT use on undergraduate students' creativity: a threat to creative thinking? *Discover Artificial Intelligence*, *4*(1). https://doi.org/10.1007/s44163-024-00172-x

Uppal, K., & Hajian, S. (2024). Students' Perceptions of ChatGPT in Higher Education: A Study of Academic Enhancement, Procrastination, and Ethical Concerns. *European Journal of Educational Research*, *14*(1), 199–213. https://doi.org/10.12973/eu-jer.14.1.199

Urban, M., Filip Děchtěrenko, Jiří Lukavský, Hrabalová, V., Filip Svacha, Brom, C., & Urban, K. (2024). ChatGPT improves creative problem-solving performance in university students: An experimental study. *Computers & Education*, *215*, 105031–105031. https://doi.org/10.1016/j.compedu.2024.105031

Venkatesh, V., & Bala, H. (2008). Technology Acceptance Model 3 and a research agenda for interventions. *Decision Sciences*, *39*(2), 273–315. https://doi.org/10.1111/j.1540-5915.2008.00192.x

Yeo, M. A. (2023). Academic integrity in the age of Artificial Intelligence (AI) authoring apps. *TESOL Journal*, *14*(3), 1–12. https://doi.org/10.1002/tesj.716

Yilmaz, R., & Yilmaz, F. G. K. (2023a). Augmented intelligence in programming learning: Examining student views on the use of ChatGPT for programming learning. *Computers in Human Behavior: Artificial Humans*, *1*(2), 100005. https://doi.org/10.1016/j.chbah.2023.100005

Yilmaz, R., & Yilmaz, F. G. K. (2023b). The effect of generative artificial intelligence (AI)-based tool use on students' computational thinking skills, programming self-efficacy, and motivation. *Computers and Education: Artificial Intelligence*, *4*, 100147. https://doi.org/10.1016/j.caeai.2023.100147

Zhou, E., & Lee, D. (2024). Generative artificial intelligence, human creativity, and art. *PNAS Nexus*, *3*(3). https://doi.org/10.1093/pnasnexus/pgae052

Received 19 Oct 2025
Revision Received 14 Jan 2026
Published 02 Feb 2026