DFT Journal of Humanities and Social Sciences
Quality. Integrity. Open Access Publishing.
Volume 1, Issue 1, 2026 | https://dftjournal.online

OPEN ACCESS
CC BY 4.0
Review Article

RECEIVED	REVISED	ACCEPTED	PUBLISHED ONLINE
19 November 2026	Not Applicable	19 November 2026	20 November 2026

Vol. 1, No. 1, pp. [1–33]
© 2026 Meshkat Ebnay Mozahid & Md. Anis Mahmud Sakib. This article is published under the Creative Commons Attribution 4.0 International licence (CC BY 4.0): https://creativecommons.org/licenses/by/4.0/

Daily Generative AI Use and AI Dependency-Related Academic Vulnerability Among University Students in Bangladesh: A Multi-Institutional Cross-Sectional Study

Meshkat Ebnay Mozahid[1*], meshkat@cou.ac.bd, https://orcid.org/0009-0004-7086-9171

Md. Anis Mahmud Sakib[2], anismahmud-2018423208@soc.du.ac.bd

[1] *Department of Statistics, Comilla University, Comilla, Bangladesh*

[2] *Department of Sociology, University of Dhaka, Dhaka, Bangladesh*

Corresponding author (*): Meshkat Ebnay Mozahid, meshkat@cou.ac.bd

Abstract

Background: Generative artificial intelligence (AI) tools are now routine in higher education, yet little is known about how the intensity of daily use relates to students' dependency and perceived academic vulnerability in low- and middle-income settings. We examined this association among Bangladeshi university students.

Methods: We analyzed a publicly available cross-sectional online survey of 1,104 students from 30 universities in Bangladesh (February-June 2026). A nine-item AI Dependency and Academic Vulnerability (AIDAV) index (0-100) was evaluated using reliability analysis and exploratory and confirmatory factor analysis. Associations were estimated using linear regression with university-clustered standard errors, random-intercept models, modified Poisson regression, item-level ordinal logistic regression, and interaction terms, with sensitivity analyses excluding duplicate records and the largest institution.

Results: The index was unidimensional and reliable (α = 0.936; CFI = 0.985; RMSEA = 0.062). The mean AIDAV score was 58.67 (SD = 23.62), and 67.21% of students had high dependency-vulnerability. Compared with students using AI for less than 1 h/day, adjusted AIDAV scores were 31.36 points (95% CI: 28.07, 34.65) higher among those using AI for 1-3 h/day and 40.21 points (95% CI: 37.62, 42.80) higher among those using it for more than 3 h/day; the corresponding adjusted prevalence ratios for high dependency were 3.88 (2.92, 5.16) and 4.36 (3.16, 6.02). Male students scored 3.81 points (2.52, 5.09) higher, although the sex gap narrowed at higher usage levels (interaction p = .005). Estimates were robust across sensitivity analyses. The AIDAV index was not associated with self-reported CGPA (OR per SD = 1.01; 0.86, 1.19).

Conclusions: Longer daily AI use was associated with greater self-reported dependency and academic vulnerability in a graded, dose-response pattern. Universities should integrate AI literacy, verification practices, and assessment designs that preserve independent learning. Longitudinal studies are needed to establish temporality.

Keywords: generative artificial intelligence; AI dependency; academic vulnerability; cognitive offloading; higher education; Bangladesh

1. Introduction

Generative artificial intelligence (AI) tools such as ChatGPT and Google Gemini entered everyday academic life within months of the public release of ChatGPT in late 2022 (Kasneci et al., 2023; Lo, 2023). Students now use these tools to explain difficult concepts, draft and edit text, summarize readings, write code, and prepare for examinations (Chan & Hu, 2023; Kasneci et al., 2023). The same capabilities, however, raise concerns about academic integrity (Cotton et al., 2024), the reliability of machine-generated content, and, more fundamentally, whether habitual reliance on AI erodes the cognitive work that learning requires. International guidance has therefore called for a human-centered approach that protects learner agency and the development of higher-order thinking (Miao & Holmes, 2023).

A growing body of evidence suggests that unrestricted use of generative AI can improve immediate task performance while weakening learning. In a field experiment with nearly 1,000 high school students, access to a standard GPT-4 interface improved practice performance, yet students performed worse than controls once access was removed (Bastani et al., 2025). Laboratory studies with university students report that AI assistance reduces mental effort but also the depth of reasoning (Stadler et al., 2024), fosters “metacognitive laziness” (Fan et al., 2025), and encourages students to rely on rather than learn from AI feedback (Darvishi et al., 2024). Survey studies link frequent AI use to cognitive offloading and lower critical thinking (Gerlich, 2025), to procrastination, memory loss, and poorer academic performance (Abbas et al., 2024), and to a recognizable pattern of “AI dependency” (Zhang et al., 2024). A systematic review concluded that over-reliance on AI dialogue systems can impair critical thinking, decision-making, and analytical reasoning (Zhai et al., 2024).

Most of this evidence comes from China, Europe, North America, and Pakistan. Bangladesh, with one of the largest tertiary student populations in South Asia and a mixed public-private university sector, has received comparatively little attention. Available studies are

descriptive surveys of single institutions (Naher, 2023), technology-acceptance studies (Rahman et al., 2025), or qualitative interview studies (Hasib & Islam, 2025). These studies document rapid adoption and emerging concerns about over-reliance and plagiarism, but they do not quantify how the intensity of everyday AI use relates to students' dependency and perceived academic vulnerability, nor do they account for the clustering of students within universities.

Three gaps follow. First, it is unclear whether the association between AI use and dependency-related vulnerability follows a graded, dose-response pattern across levels of daily use. Second, little is known about whether this association differs by sex, academic level, or institutional type, all of which shape technology adoption (Møgelvang et al., 2024; Stöhr et al., 2024). Third, the relationship between self-reported dependency and academic performance remains contested (Abbas et al., 2024; Bastani et al., 2025).

This study addresses these gaps using a multi-institutional survey of 1,104 Bangladeshi university students. It asks four research questions: (RQ1) Do nine questionnaire items on AI reliance and academic consequences form a reliable and valid composite measure? (RQ2) Is daily AI usage time associated with AI dependency-related academic vulnerability, independent of sociodemographic and academic characteristics, and is the association graded? (RQ3) Does the association differ by sex, academic level, or university type? (RQ4) Is AI dependency-related academic vulnerability associated with self-reported cumulative grade point average (CGPA)? The article contributes a psychometrically evaluated composite index, cluster-robust estimates of a dose-response association, and evidence on effect modification from an under-studied South Asian context. Section 2 reviews the literature and sets out the conceptual framework, Section 3 describes the methods, Section 4 presents the results, Section 5 discusses the findings, and Section 6 concludes.

2. Literature review and conceptual framework

2.1 Generative AI in higher education: adoption and its correlates

Research on AI in higher education predates generative AI but has long focused on institutional applications such as profiling, assessment, and intelligent tutoring, with limited attention to students' own practices (Zawacki-Richter et al., 2019). After 2022, attention shifted to student adoption. Technology-acceptance studies show that habit, performance expectancy, and hedonic motivation drive students' intention to use ChatGPT (Strzelecki, 2024). Large

surveys reveal systematic differences in use: in Sweden, male students and students in technology and engineering reported more frequent use and more positive attitudes, whereas female students expressed more concern (Stöhr et al., 2024); a Norwegian study similarly documented gender differences in the breadth and frequency of chatbot use (Møgelvang et al., 2024). Students themselves describe generative AI as useful for personalized support while also voicing concerns about accuracy, over-reliance, and the value of their own learning (Chan & Hu, 2023). In Bangladesh, students report using ChatGPT mainly for assignments and examination preparation (Naher, 2023), peer influence shapes adoption (Rahman et al., 2025), and interviewees acknowledge over-reliance and unattributed use (Hasib & Islam, 2025).

2.2 Cognitive offloading, over-reliance, and academic vulnerability

The theoretical starting point is cognitive offloading, the use of external resources to reduce the internal cognitive demands of a task (Risko & Gilbert, 2016). Offloading can be adaptive, consistent with the view that external tools can become part of a coupled cognitive system (Clark & Chalmers, 1998). When information is expected to remain accessible, however, people are less likely to remember it themselves (Sparrow et al., 2011). Generative AI extends offloading from information storage to reasoning and composition. From the perspective of cognitive load theory, removing productive effort can reduce the germane processing that builds durable knowledge (Sweller et al., 2019). Experimental work confirms that large language models lower perceived effort but yield shallower reasoning (Stadler et al., 2024), and neurophysiological evidence suggests lower neural engagement and weaker recall of one's own AI-assisted writing (Kosmyna et al., 2025).

A second mechanism concerns trust and verification. Research on automation shows that reliable-seeming systems invite complacency and reduced monitoring (Parasuraman & Manzey, 2010). Among knowledge workers, greater confidence in generative AI is associated with less critical thinking, and critical effort shifts toward verifying and integrating AI output (Lee et al., 2025). AI may also create illusions of understanding, in which users believe they understand more than they do (Messeri & Crockett, 2024). Such dynamics can produce a cluster of vulnerabilities: turning to AI as a first resort, accepting outputs without modification or verification, prioritizing speed over understanding, reduced confidence when working without AI, and submitting work that one does not fully understand. Studies of "AI dependency" describe

this cluster and link it to academic stress, performance expectations, and reduced independent thinking (Zhang et al., 2024), while longitudinal evidence among adolescents indicates that prior mental health problems can predict later AI dependence, underscoring the possibility of bidirectional relationships (Huang et al., 2024).

2.3 Conceptual framework

Building on these strands, the study treats daily AI usage time as an indicator of exposure intensity. Greater exposure is expected to increase opportunities for offloading and habitual reliance, expressed in behaviors such as using AI as a first instinct, accepting unmodified output, and not verifying information. These behaviors are, in turn, expected to coincide with perceived academic vulnerabilities: weaker recall, reduced confidence in one's writing and analytical skills, perceived weakening of critical thinking, and submission of work that is not fully understood. Because the available items measure both reliance behaviors and perceived consequences, and because these are expected to co-occur, they are modeled as indicators of one latent construct, termed AI dependency-related academic vulnerability (AIDAV). Sex, academic level, and university type are treated as potential effect modifiers, and CGPA as a downstream academic outcome. Given the cross-sectional design, the framework describes expected associations rather than causal pathways. Three hypotheses were specified: H1, AIDAV increases with daily usage time in a graded manner; H2, the gradient differs by sex and academic level; and H3 (exploratory), higher AIDAV is associated with lower CGPA.

3. Methodology

3.1 Study design and data source

This was a secondary analysis of a cross-sectional survey dataset, "AI Dependency and Academic Vulnerability in Higher Education: A Survey Dataset from Bangladesh," deposited in Mendeley Data under a CC BY 4.0 license (Islam et al., 2026). Reporting follows the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guideline for cross-sectional studies (von Elm et al., 2007).

3.2 Context, participants, and sampling

According to the data descriptor, a structured questionnaire adapted from previously validated instruments was administered through Google Forms between February 10 and June 15, 2026, and distributed via WhatsApp, Facebook, and LinkedIn. Participation was voluntary; incomplete and inconsistent responses were removed by the original investigators, and personally identifiable information was excluded before publication (Islam et al., 2026). The dataset contains 1,104 respondents. Daffodil International University contributed 420 respondents (38.04%). Sampling was non-probabilistic, and no survey weights, strata, or primary sampling units were provided; all estimates are therefore unweighted, and the clustering of students within universities was addressed analytically.

Data screening identified no missing values except for one respondent who selected "Prefer not to say" for gender, who was excluded from adjusted models (analytic N = 1,103). Exact duplication of all 15 variables was found for 162 records (14.67%) beyond their first occurrence. Because the questionnaire consists entirely of categorical items, identical response patterns are possible among genuine respondents, and the data descriptor states that inconsistent responses had already been removed. Duplicate records were therefore retained in the primary analysis and excluded in a sensitivity analysis (analytic n = 941).

3.3 Measures

Outcome. Nine five-point Likert items (1 = strongly disagree to 5 = strongly agree) captured AI reliance behaviors and perceived academic consequences (Appendix Table A1). Two protective items, verifying AI-generated information (Q1) and being able to study effectively without AI (Q2), were reverse-scored so that higher values indicate greater vulnerability. The AIDAV index was computed as the mean of the nine items and linearly rescaled to 0-100 as $[(\text{mean} - 1)/4] \times 100$. A binary indicator of high AI dependency-vulnerability was defined as a mean item score greater than 3, that is, above the neutral midpoint.

Exposure. Daily AI usage time was recorded in three categories: less than 1 hour, 1-3 hours, and more than 3 hours per day.

Covariates. Covariates were sex (female, male), age group (23 years or younger, 24-26, above 26; the "Under 20" category, n = 28, was merged with "20-23"), academic level (undergraduate, graduate, master's), current CGPA (3.00 or below, 3.01-3.50, above 3.50; the

"Below 2.50" category, n = 8, was merged with "2.51-3.00"), and university type (public, private). University type was classified by the authors from the institution names

3.4 Data analysis

Psychometric evaluation. Internal consistency was assessed with Cronbach's α and corrected item-rest correlations. Sampling adequacy was assessed with the Kaiser-Meyer-Olkin (KMO) measure, and dimensionality with principal-component factoring, the Kaiser criterion, and the scree plot. A one-factor confirmatory factor analysis (CFA) was estimated by maximum likelihood, with fit judged by the comparative fit index (CFI) and Tucker-Lewis index (TLI) ≥ 0.95, root mean square error of approximation (RMSEA) ≤ 0.06-0.08, and standardized root mean square residual (SRMR) ≤ 0.08 (Hu & Bentler, 1999). Guided by modification indices, a second model allowed correlated residuals between the two reverse-worded items to account for a wording method effect.

Descriptive and bivariate analyses. Characteristics were compared across usage groups with Pearson's χ^2 tests. AIDAV scores were compared with one-way analysis of variance (ANOVA) with Bonferroni-adjusted pairwise comparisons and, because Bartlett's test indicated unequal variances, with Kruskal-Wallis and Wilcoxon rank-sum tests. The intraclass correlation coefficient (ICC) for universities was estimated from an empty random-intercept model.

Main model. AIDAV (0-100) was regressed on daily usage time, sex, age group, academic level, CGPA, and university type using ordinary least squares (OLS) with standard errors clustered by university (30 clusters) (Cameron & Miller, 2015). A graded association was assessed with orthogonal polynomial contrasts, and adjusted means were obtained as predictive margins. Diagnostics included variance inflation factors (VIF), the Breusch-Pagan and White tests, the Ramsey RESET test, the link test, the Shapiro-Wilk test of residuals, standardized residuals, leverage, and Cook's distance.

Sensitivity analyses. The main model was re-estimated (a) with HC3 heteroskedasticity-consistent standard errors (Long & Ervin, 2000), (b) excluding duplicate records, (c) excluding the dominant institution, (d) as a random-intercept linear mixed model with restricted maximum likelihood, and (e) as a fractional logit model for the bounded outcome rescaled to 0-1, reporting average marginal effects (AMEs) (Papke & Wooldridge, 1996). A wild cluster bootstrap was

planned for the small number of clusters but could not be executed because the available routine was incompatible with Stata 17; the HC3 and mixed models were used as alternatives.

Binary and item-level outcomes. Because high AIDAV was common (67.21%), adjusted prevalence ratios (aPRs) were estimated with modified Poisson regression with a log link and cluster-robust standard errors (Zou, 2004); adjusted odds ratios (aORs) from logistic regression are reported for comparison, with model performance assessed using the Hosmer-Lemeshow test, the area under the receiver operating characteristic curve (AUC), classification accuracy, and the link test. Each item was analyzed with ordinal logistic regression; the parallel-lines assumption was assessed with the Brant and related tests (Brant, 1990) and compared using information criteria, and partial proportional odds models were fitted where appropriate (Williams, 2006).

Effect modification and CGPA. Interactions of usage time with sex, academic level, and university type were tested with joint Wald F tests. Finally, CGPA was regressed on the standardized AIDAV score and covariates using ordinal logistic regression, with a partial proportional odds model as a robustness check. All tests were two-sided with $\alpha = .05$. Analyses were conducted in Stata 17 (StataCorp, College Station, TX), using the user-written commands estout, oparallel, and gologit2.

3.5 Ethical considerations

This study used fully anonymized, publicly available secondary data released under a CC BY 4.0 license. The original investigators report that participation was voluntary, that respondents were informed of the academic purpose of the study, and that personally identifiable information was removed before publication (Islam et al., 2026). The authors had no access to identifying information and no contact with participants.

4. Results

4.1 Sample characteristics

Of the 1,104 respondents, 545 (49.41%) were female and 558 (50.59%) male; 541 (49.00%) were aged 20-23 years and 457 (41.39%) 24-26 years; 776 (70.29%) were undergraduates; 649 (58.79%) reported a CGPA of 3.01-3.50; and 683 (61.87%) attended private universities (Table 1). Daily AI use was less than 1 hour for 355 students (32.16%), 1-3 hours for

291 (26.36%), and more than 3 hours for 458 (41.49%). Usage time differed by sex ($\chi^2(2) = 10.920$, $p = .004$), age group ($\chi^2(6) = 42.967$, $p < .001$), academic level ($\chi^2(4) = 28.340$, $p < .001$), CGPA ($\chi^2(4) = 13.630$, $p = .009$), and university type ($\chi^2(2) = 15.576$, $p < .001$). Heavy users (more than 3 h/day) were more often male (52.5% vs. 43.7% among light users), master's students (26.4% vs. 14.1%), students with a CGPA above 3.50 (26.6% vs. 17.2%), and students at public universities (45.0% vs. 33.0%).

Table 1

Characteristics of respondents by daily AI usage time, n (%)

Characteristic	< 1 h/day (n = 355)	1-3 h/day (n = 291)	> 3 h/day (n = 458)	Total (N = 1,104)	χ^2 (df); p
Sex					10.920 (2); .004
Female	200 (56.3)	128 (44.0)	217 (47.5)	545 (49.4)	
Male	155 (43.7)	163 (56.0)	240 (52.5)	558 (50.6)	
Age group (years)					42.967 (6); < .001
Under 20	11 (3.1)	12 (4.1)	5 (1.1)	28 (2.5)	
20-23	215 (60.6)	131 (45.0)	195 (42.6)	541 (49.0)	
24-26	114 (32.1)	131 (45.0)	212 (46.3)	457 (41.4)	
Above 26	15 (4.2)	17 (5.8)	46 (10.0)	78 (7.1)	
Academic level					28.340 (4); < .001
Undergraduate	282 (79.4)	197 (67.7)	297 (64.8)	776 (70.3)	
Graduate	23 (6.5)	37 (12.7)	40 (8.7)	100 (9.1)	
Master's	50 (14.1)	57 (19.6)	121 (26.4)	228 (20.7)	
Current CGPA					13.630 (4); .009[a]
Below 2.50	3 (0.8)	2 (0.7)	3 (0.7)	8 (0.7)	
2.51-3.00	67 (18.9)	64 (22.0)	73 (15.9)	204 (18.5)	
3.01-3.50	224 (63.1)	165 (56.7)	260 (56.8)	649 (58.8)	
Above 3.50	61 (17.2)	60 (20.6)	122 (26.6)	243 (22.0)	
University type					15.576 (2); < .001
Private	238 (67.0)	193 (66.3)	252 (55.0)	683 (61.9)	
Public	117 (33.0)	98 (33.7)	206 (45.0)	421 (38.1)	
High AIDAV (mean item score > 3)					512.624 (2); < .001
No	280 (78.9)	50 (17.2)	32 (7.0)	362 (32.8)	
Yes	75 (21.1)	241 (82.8)	426 (93.0)	742 (67.2)	

Note. Column percentages. Sex was missing for one respondent ("Prefer not to say"). AIDAV = AI Dependency and Academic Vulnerability index. [a] Test performed after merging "Below 2.50" with "2.51-3.00" because of sparse cells.

4.2 Psychometric properties of the AIDAV index

Inter-item correlations ranged from 0.532 to 0.730, all $p < .05$. Internal consistency was high (Cronbach's $\alpha = 0.936$; standardized $\alpha = 0.936$), corrected item-rest correlations ranged from 0.693 (Q3) to 0.831 (Q9), and no item deletion would have raised α (Appendix Table A1). Sampling adequacy was excellent (overall KMO = 0.946; item KMO 0.911-0.967). A single component had an eigenvalue above 1 (5.97, explaining 66.28% of the variance), followed by 0.68 for the second component (Appendix Figure A1), with loadings from 0.756 to 0.875. The one-factor CFA fit acceptably on incremental indices but not on RMSEA ($\chi^2(27) = 366.27$; CFI = 0.952; TLI = 0.936; RMSEA = 0.107, 90% CI: 0.097, 0.117; SRMR = 0.037). The largest modification index (223.26) involved the residuals of the two reverse-worded items. Allowing this residual correlation (r = 0.456) improved fit substantially ($\chi^2(26) = 135.16$; CFI = 0.985; TLI = 0.979; RMSEA = 0.062, 90% CI: 0.052, 0.072; SRMR = 0.023; likelihood-ratio $\chi^2(1) = 231.10$, $p < .001$), with standardized loadings from 0.701 to 0.876. These results support treating the nine items as one construct (RQ1).

Item responses are shown in Figure 1. Most students agreed or strongly agreed with each reliance and consequence item (54.17% to 59.78% for Q3-Q9). About half disagreed that they verify AI-generated information (50.27%) or that they can study effectively without AI (49.73%). Submission of work without fully understanding AI-generated content (Q9) showed the most polarized distribution, with 28.35% strongly agreeing and 14.76% strongly disagreeing.

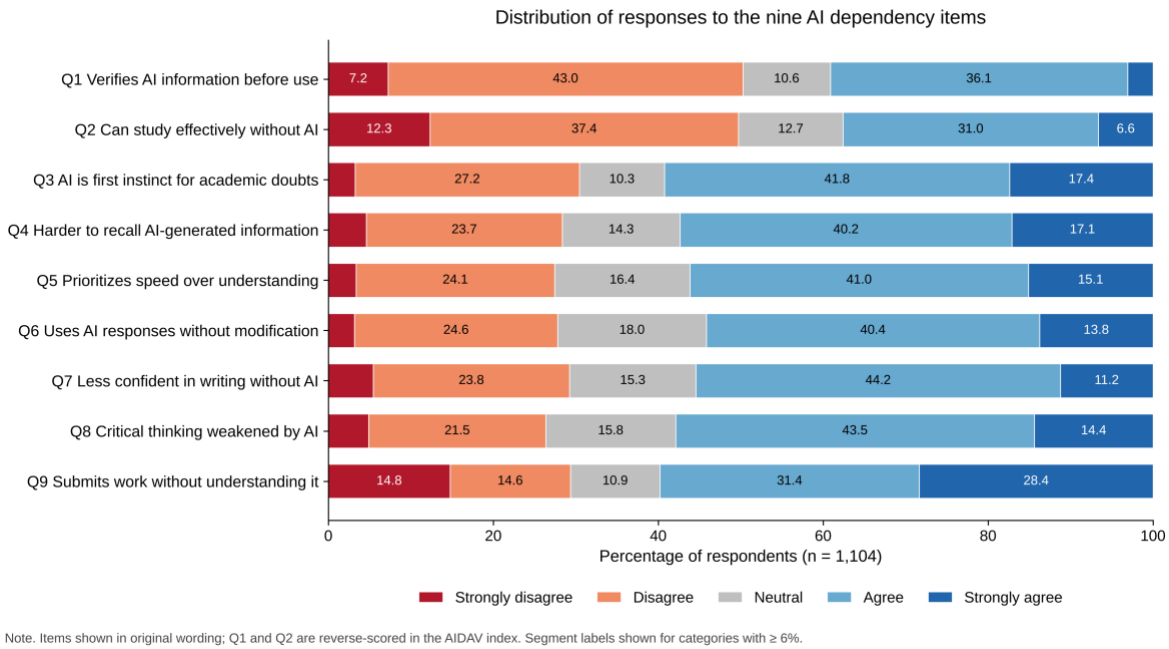


Figure 1. Distribution of responses to the nine AI dependency and academic vulnerability items (N = 1,104). Items are shown in their original wording; Q1 and Q2 are reverse-scored in the AIDAV index.

4.3 Distribution of the AIDAV index and bivariate associations

The mean AIDAV score was 58.67 (SD = 23.62; median = 66.67; interquartile range 33.33-77.78), and 742 students (67.21%) had high dependency-vulnerability. Scores rose steeply with daily usage time, from a mean of 33.43 (SD = 19.41) among students using AI for less than 1 h/day to 65.20 (SD = 14.30) among those using it for 1-3 h/day and 74.08 (SD = 12.84) among those using it for more than 3 h/day ($F(2, 1101) = 713.38$, $p < .001$; Kruskal-Wallis $\chi^2(2) = 532.72$, $p < .001$; Table 2, Figure 2). All Bonferroni-adjusted pairwise differences were significant (31.77, 40.66, and 8.89 points; all $p < .001$). The distribution among light users was concentrated around 20-30 with a long upper tail, whereas heavy users clustered around 75-80 (Figure 2; Appendix Figure A2). The prevalence of high dependency-vulnerability was 21.1%, 82.8%, and 93.0% across the three groups ($\chi^2(2) = 512.62$, $p < .001$). Male students scored higher than female students (61.94 vs. 55.27; $z = -4.79$, $p < .001$), whereas differences across CGPA categories were not significant ($\chi^2(2) = 4.66$, $p = .097$). The empty random-intercept model indicated modest clustering by university (ICC = 0.035, 95% CI: 0.015, 0.080; likelihood-ratio test $\chi^2(01) = 61.37$, $p < .001$).

Table 2

Bivariate distribution of the AIDAV index (0-100) by respondent characteristics

Characteristic	n	Mean (SD)	Median	Test statistic	p
Daily AI usage time				$F(2, 1101) = 713.38$[a]	< .001
< 1 h/day	355	33.43 (19.41)	25.00	KW $\chi^2(2) = 532.72$[b]	< .001
1-3 h/day	291	65.20 (14.30)	66.67		
> 3 h/day	458	74.08 (12.84)	77.78		
Sex				$z = -4.79$[c]	< .001
Female	545	55.27 (24.11)	61.11		
Male	558	61.94 (22.68)	69.44		
Academic level				—	
Undergraduate	776	56.80 (24.66)	66.67		
Graduate	100	62.94 (19.37)	66.67		
Master's	228	63.16 (20.75)	69.44		
Current CGPA				KW $\chi^2(2) = 4.66$[b]	.097
3.00 or below	212	57.59 (22.40)	63.89		
3.01-3.50	649	57.99 (24.35)	66.67		
Above 3.50	243	61.42 (22.56)	66.67		
University type				-	
Private	683	56.65 (23.14)	63.89		
Public	421	61.94 (24.05)	72.22		
Total	1,104	58.67 (23.62)	66.67		

Note. [a] One-way ANOVA (Bartlett's test of equal variances $\chi^2(2) = 73.83$, $p < .001$). [b] Kruskal-Wallis test with ties. [c] Wilcoxon rank-sum test. - = formal test not performed; descriptive statistics only. KW = Kruskal-Wallis.

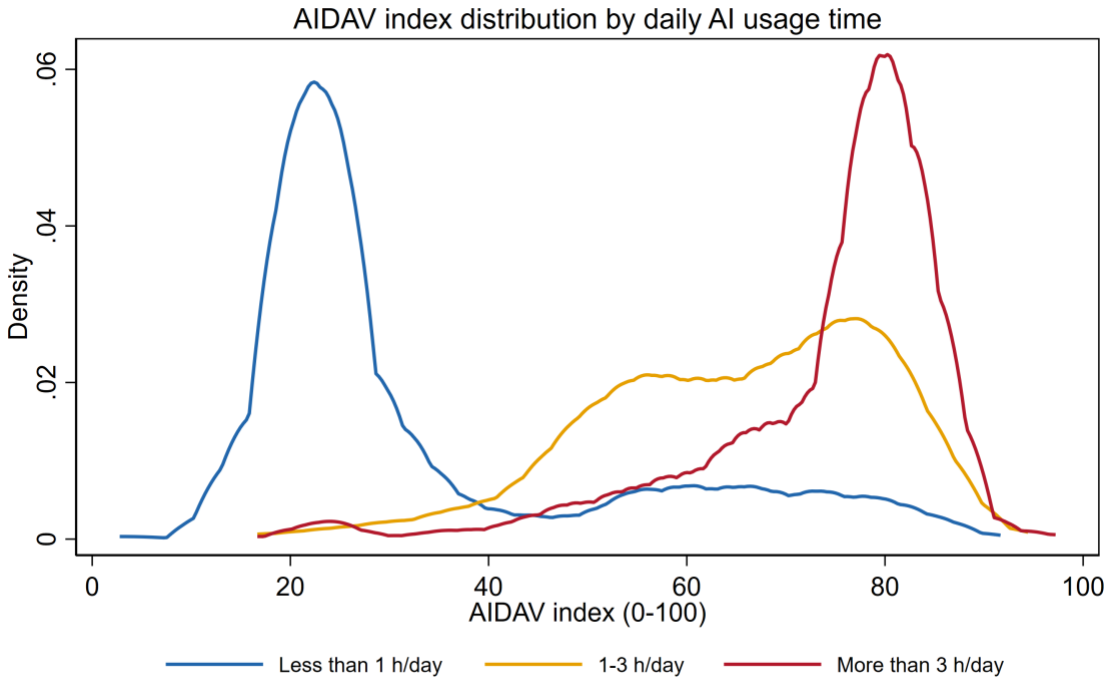


Figure 2. Kernel density of the AIDAV index (0-100) by daily AI usage time.

4.4 Multivariable association between daily AI use and AIDAV

In the fully adjusted model with university-clustered standard errors (Table 3, Model 1), students using AI for 1-3 h/day scored 31.36 points higher (95% CI: 28.07, 34.65; $p < .001$) and those using AI for more than 3 h/day 40.21 points higher (95% CI: 37.62, 42.80; $p < .001$) than students using AI for less than 1 h/day (joint $F(2, 29) = 530.71$, $p < .001$). Orthogonal contrasts confirmed a significant linear trend ($F(1, 29) = 1005.88$, $p < .001$) and a significant quadratic component ($F(1, 29) = 88.45$, $p < .001$), indicating a graded but decelerating association: the increase from less than 1 h/day to 1-3 h/day was larger than the further increase to more than 3 h/day. Adjusted mean AIDAV scores were 33.71 (95% CI: 31.36, 36.06), 65.07 (61.88, 68.22), and 73.92 (71.54, 76.31), respectively (Figure 3). Male students scored 3.81 points higher than female students (95% CI: 2.52, 5.09; $p < .001$). Age group, academic level, CGPA, and university type were not significantly associated with AIDAV. The model explained 57.27% of the variance ($R^2 = 0.5727$). The random-intercept model produced nearly identical estimates (Model 2: 30.18 and 39.47 points; male 3.26), with a small residual ICC of 0.032 (95% CI: 0.010, 0.094).

Table 3

Multivariable linear models of the AIDAV index (0-100)

Variable	Model 1: OLS, cluster-robust β (95% CI)	p	Model 2: random intercept β (95% CI)	p
Daily AI usage (ref. < 1 h/day)				
1-3 h/day	31.36 (28.07, 34.65)	< .001	30.18 (27.73, 32.64)	< .001
> 3 h/day	40.21 (37.62, 42.80)	< .001	39.47 (37.27, 41.67)	< .001
Sex (ref. female)				
Male	3.81 (2.52, 5.09)	< .001	3.26 (1.34, 5.19)	.001
Age group (ref. 23 or younger)				
24-26	−1.12 (−3.72, 1.49)	.389	−0.96 (−3.14, 1.22)	.388
Above 26	−2.15 (−4.76, 0.46)	.103	−1.67 (−5.60, 2.26)	.405
Academic level (ref. undergraduate)				
Graduate	1.87 (−1.09, 4.83)	.208	1.59 (−1.71, 4.90)	.344
Master's	1.03 (−2.40, 4.47)	.543	1.91 (−0.81, 4.62)	.168
CGPA (ref. 3.00 or below)				
3.01-3.50	1.26 (−1.33, 3.85)	.329	1.35 (−1.07, 3.76)	.275
Above 3.50	0.32 (−2.42, 3.05)	.814	−0.06 (−2.93, 2.81)	.968
University type (ref. private)				
Public	2.07 (−1.31, 5.44)	.220	0.12 (−3.19, 3.43)	.942
Intercept	30.42 (26.14, 34.70)	< .001	33.10 (29.54, 36.66)	< .001
Model statistics				
Observations / clusters	1,103 / 30		1,103 / 30	
Overall test	$F(10, 29) = 235.94$	< .001	Wald $\chi^2(10) = 1365.17$	< .001
R^2	0.5727		—	
Variance: university / residual	—		7.72 / 233.15	
Residual ICC (95% CI)	—		0.032 (0.010, 0.094)	

Note. Model 1: ordinary least squares with standard errors clustered by university (30 clusters). Model 2: linear mixed model with a university random intercept, estimated by restricted maximum likelihood. β = unstandardized coefficient. ICC = intraclass correlation coefficient. Sensitivity analyses are reported in Appendix Table A2.

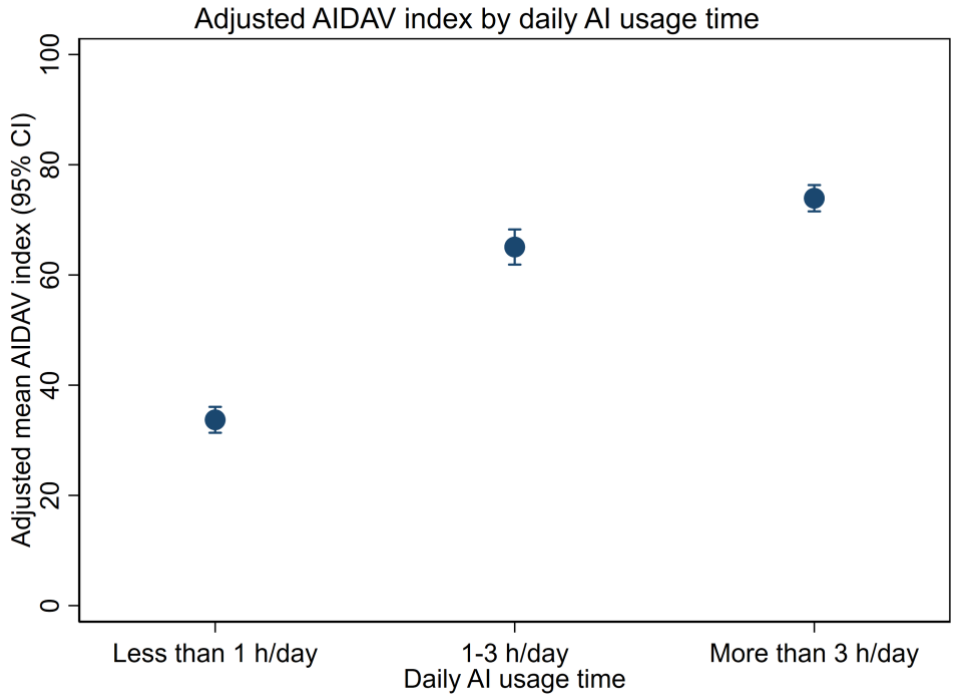


Figure 3. Adjusted mean AIDAV index (95% CI) by daily AI usage time, from the cluster-robust linear model adjusted for sex, age group, academic level, CGPA, and university type.

Diagnostics did not indicate problematic collinearity (mean VIF = 1.33; maximum 1.72) or misspecification (RESET $F(3, 1089) = 0.99$, $p = .396$; link test squared-prediction term $p = .454$). Residuals were heteroskedastic (Breusch-Pagan $\chi^2(1) = 69.27$, $p < .001$; White $\chi^2(51) = 122.70$, $p < .001$) and departed from normality (Shapiro-Wilk $W = 0.970$, $p < .001$), supporting the use of robust standard errors and the fractional logit sensitivity model. Twenty-two observations had absolute standardized residuals above 3, but no observation was unduly influential (maximum Cook's distance = 0.020; maximum leverage = 0.030; Appendix Figure A3).

4.5 Sensitivity analyses

The usage gradient was robust (Appendix Table A2). With HC3 standard errors, the confidence intervals for 1-3 h/day (28.69, 34.03) and more than 3 h/day (37.77, 42.65) remained narrow; university type reached nominal significance only under this specification ($\beta = 2.07$; 95% CI: 0.02, 4.11; $p = .048$). Excluding duplicate records attenuated but did not alter the estimates (28.07 and 36.62 points; n = 941; $R^2 = 0.4730$), and excluding Daffodil International University produced similar estimates (30.70 and 39.91 points; n = 684). In the fractional logit

model, the AMEs on the 0-1 scale were 0.313 (95% CI: 0.282, 0.345) for 1-3 h/day and 0.402 (0.377, 0.428) for more than 3 h/day. The male-female difference was significant in every specification. The only other change was a negative association for students older than 26 in the analysis excluding duplicates (β = −2.67; 95% CI: −4.77, −0.58; p = .014).

4.6 High AI dependency-vulnerability

In modified Poisson regression (Table 4), the adjusted prevalence of high dependency-vulnerability was 3.88 times higher (95% CI: 2.92, 5.16) among students using AI for 1-3 h/day and 4.36 times higher (95% CI: 3.16, 6.02) among those using it for more than 3 h/day than among light users. Male sex (aPR = 1.09; 95% CI: 1.05, 1.13) and graduate-level study (aPR = 1.10; 95% CI: 1.01, 1.20; p = .021) were also associated with a higher prevalence. The corresponding odds ratios (18.11 and 51.10) were far larger than the prevalence ratios, illustrating how odds ratios overstate relative risk when an outcome is common. Adjusted predicted probabilities from the logistic model were 0.217 (95% CI: 0.163, 0.271), 0.825 (0.749, 0.901), and 0.929 (0.895, 0.964) across the three usage groups. The logistic model discriminated well (AUC = 0.877; 85.77% correctly classified; sensitivity 89.88%; specificity 77.35%; Appendix Figure A4) and showed no link misspecification (p = .339), but the Hosmer-Lemeshow test indicated some miscalibration ($\chi^2(8) = 24.56$, p = .002).

Table 4

Factors associated with high AI dependency-vulnerability (mean item score > 3; prevalence 67.21%)

Variable	aPR (95% CI)	p	aOR (95% CI)	p
Daily AI usage (ref. < 1 h/day)				
1-3 h/day	3.88 (2.92, 5.16)	< .001	18.11 (11.80, 27.78)	< .001
> 3 h/day	4.36 (3.16, 6.02)	< .001	51.10 (31.05, 84.11)	< .001
Sex (ref. female)				
Male	1.09 (1.05, 1.13)	< .001	1.57 (1.27, 1.95)	< .001
Age group (ref. 23 or younger)				
24-26	0.96 (0.90, 1.03)	.308	0.76 (0.48, 1.20)	.237
Above 26	0.93 (0.85, 1.03)	.153	0.63 (0.33, 1.23)	.178
Academic level (ref. undergraduate)				
Graduate	1.10 (1.01, 1.20)	.021	1.89 (1.26, 2.85)	.002
Master's	1.06 (0.97, 1.16)	.217	1.43 (0.80, 2.55)	.231
CGPA (ref. 3.00 or below)				
3.01-3.50	1.05 (0.95, 1.16)	.386	1.27 (0.71, 2.30)	.420
Above 3.50	1.00 (0.92, 1.09)	.921	0.97 (0.59, 1.58)	.895
University type (ref. private)				
Public	1.04 (0.94, 1.15)	.405	1.30 (0.73, 2.30)	.378
Model statistics				
Observations / clusters	1,103 / 30		1,103 / 30	
Fit	—		Pseudo R^2 = 0.390; AUC = 0.877	

Note. aPR = adjusted prevalence ratio from modified Poisson regression (log link); aOR = adjusted odds ratio from logistic regression. Both models use standard errors clustered by university and adjust for all variables shown. AUC = area under the receiver operating characteristic curve.

4.7 Item-level associations

Daily usage time was associated with higher odds of agreement on every item (Table 5). Adjusted odds ratios ranged from 9.73 to 13.82 for 1-3 h/day and from 17.20 to 37.02 for more than 3 h/day. The strongest associations among heavy users involved the two reverse-scored protective items, not being able to study effectively without AI (aOR = 37.02; 95% CI: 24.24, 56.54) and not verifying AI-generated information (aOR = 32.30; 95% CI: 20.83, 50.10), followed by submitting work without understanding it (aOR = 28.83) and perceived weakening of critical thinking (aOR = 26.08). The weakest, though still large, association concerned using AI as a first instinct for academic doubts (aOR = 17.20). Male students had higher odds on seven of the nine items, but not for Q3 or Q7.

The parallel-lines assumption was rejected for all estimable items (Brant $\chi^2(30)$ from 89.85 to 148.5; all $p < .001$), which is expected given the large sample, and could not be tested for Q5 because of perfect prediction in the generalized model. The Bayesian information criterion nevertheless favored the proportional odds model for every estimable item. In a partial proportional odds model for Q9, the usage-outcome association varied across thresholds (for more than 3 h/day, aOR = 62.88 at the lowest threshold and 7.40 at the highest) but remained well above 1 at every threshold, so the substantive conclusion was unchanged.

Table 5

Adjusted odds ratios from item-level ordinal logistic regressions

Item (higher = more vulnerable)	1-3 h/day aOR (95% CI)	> 3 h/day aOR (95% CI)	Male aOR (95% CI)	Brant $\chi^2(30)$
Q1 Does not verify AI information (R)	11.87 (8.10, 17.39)	32.30 (20.83, 50.10)	1.44 (1.15, 1.80)	135.3
Q2 Cannot study effectively without AI (R)	12.25 (8.13, 18.46)	37.02 (24.24, 56.54)	1.36 (1.06, 1.75)	130.7
Q3 AI is first instinct for academic doubts	10.46 (8.06, 13.57)	17.20 (12.28, 24.09)	1.11 (0.81, 1.51)	89.85
Q4 Harder to recall AI-generated information	10.02 (7.06, 14.20)	21.12 (17.33, 25.73)	1.45 (1.20, 1.76)	117.8
Q5 Prioritizes speed over understanding	10.27 (6.59, 16.00)	21.60 (15.53, 30.05)	1.33 (1.10, 1.61)	n.e.
Q6 Uses AI responses without modification	9.73 (5.98, 15.84)	20.11 (13.30, 30.41)	1.42 (1.23, 1.65)	108.2
Q7 Less confident in writing without AI	9.93 (6.23, 15.84)	21.99 (14.55, 33.22)	1.25 (0.92, 1.70)	99.88
Q8 Critical thinking weakened by AI	12.50 (8.67, 18.01)	26.08 (16.81, 40.47)	1.54 (1.29, 1.82)	129.0
Q9 Submits work without understanding it	13.82 (9.65, 19.79)	28.83 (17.13, 48.53)	1.37 (1.11, 1.68)	148.5

Note. Each row is a separate ordinal logistic regression with university-clustered standard errors, adjusted for age group, academic level, CGPA, and university type (reference: < 1 h/day; female). (R) = reverse-scored. All Brant tests $p < .001$. n.e. = not estimable because of perfect prediction in the generalized model.

4.8 Effect modification

The usage gradient differed by sex (joint $F(2, 29) = 6.30$, $p = .005$). Among light users, male students scored substantially higher than female students (adjusted means 37.66 vs. 30.11), but the gap narrowed among moderate users (66.40 vs. 63.83) and heavy users (74.88 vs. 73.09), with interaction coefficients of −4.98 (95% CI: −8.96, −0.99) and −5.76 (95% CI: −10.39, −1.13) (Figure 4). The gradient also differed by academic level (joint $F(4, 29) = 6.43$, $p < .001$): among light users, master's students scored higher than undergraduates ($\beta = 10.43$; 95% CI: 2.60, 18.27; $p = .011$), but their usage gradient was flatter (interaction $\beta = -13.26$ for 1-3 h/day and −11.88 for more than 3 h/day; $p < .001$ and $p = .012$). There was no evidence of effect modification by university type ($F(2, 29) = 0.63$, $p = .542$).

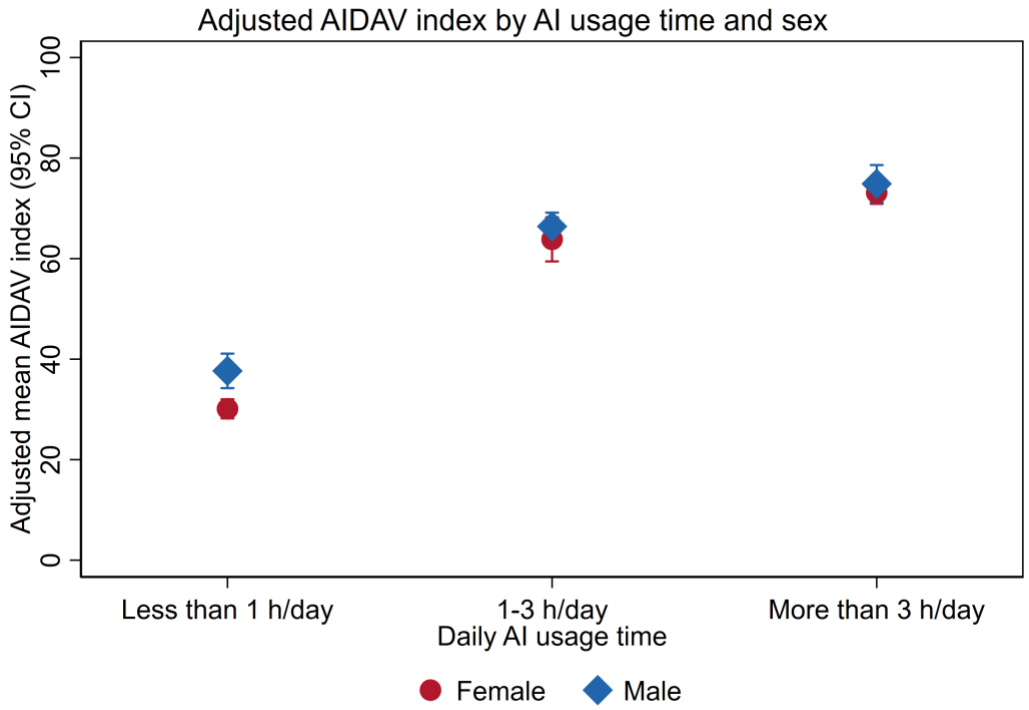


Figure 4. Adjusted mean AIDAV index (95% CI) by daily AI usage time and sex, from the cluster-robust linear model including a usage × sex interaction and adjusted for age group, academic level, CGPA, and university type.

4.9 AIDAV and self-reported CGPA

In bivariate analysis, the share of heavy AI users was higher among students with a CGPA above 3.50 (50.21%) than among those with a CGPA of 3.00 or below (35.85%; $\chi^2(4) = 13.63$, $p = .009$). After adjustment, however, the AIDAV index was not associated with CGPA (aOR per SD = 1.01; 95% CI: 0.86, 1.19; $p = .881$), nor was usage time (more than 3 h/day: aOR = 1.31; 95% CI: 0.92, 1.88; $p = .135$; Appendix Table A3). The predicted probability of a CGPA above 3.50 changed only from 0.217 to 0.225 across −2 to +2 SD of AIDAV. Master's study (aOR = 1.51; 95% CI: 1.14, 1.99) and age 24-26 years (aOR = 1.32; 95% CI: 1.04, 1.69) were associated with a higher CGPA. Although the parallel-lines assumption was violated (Brant $\chi^2(9) = 43.65$, $p < .001$), the partial proportional odds model, in which the AIDAV coefficient was constrained to be parallel, gave the same result (aOR = 1.02; 95% CI: 0.87, 1.19; $p = .803$). H3 was therefore not supported.

5. Discussion

This multi-institutional study of 1,104 Bangladeshi university students found that a nine-item measure of AI reliance behaviors and perceived academic consequences formed a coherent and reliable construct, and that AI dependency-related academic vulnerability was strongly and incrementally associated with daily AI usage time. Students using AI for more than 3 hours per day scored about 40 points higher on a 0-100 scale than those using it for less than 1 hour, and their prevalence of high dependency-vulnerability was more than four times higher. The association held after adjustment for sociodemographic and academic characteristics, under alternative variance estimators, in a multilevel model, in a fractional response model, and after excluding duplicate records and the largest institution. These findings support H1 and, with qualifications, H2.

The magnitude and direction of the usage gradient are consistent with prior work. Gerlich (2025) reported a significant negative correlation between frequency of AI tool use and critical thinking among 666 participants, mediated by cognitive offloading; in our data, the odds of agreeing that one's critical thinking had weakened were 26.08 times higher among heavy users. Abbas et al. (2024), using a three-wave design with 494 students, found that ChatGPT use predicted memory loss and procrastination; our item on difficulty recalling AI-generated information showed an aOR of 21.12 among heavy users. Zhang et al. (2024) identified reduced independent and critical thinking among the most frequently reported negative effects of AI dependency, and the systematic review by Zhai et al. (2024) reached a similar conclusion. The present study extends these findings by showing a graded association across three exposure levels in a South Asian sample, with the steepest increase between light and moderate use.

The item-level results suggest which behaviors carry the strongest association. The largest odds ratios concerned not being able to study without AI (aOR = 37.02) and not verifying AI-generated information (aOR = 32.30). This pattern is consistent with automation research showing that reliance on seemingly competent systems reduces monitoring (Parasuraman & Manzey, 2010), and with evidence that greater confidence in generative AI is associated with less critical thinking and that critical effort shifts toward verification (Lee et al., 2025). Experimental studies show that students tend to rely on rather than learn from AI assistance and that performance declines when assistance is removed (Bastani et al., 2025; Darvishi et al.,

2024). The strong association with not being able to study without AI mirrors these findings in a naturalistic setting.

Several mechanisms may explain these associations. Cognitive offloading reduces internal processing when an external resource is reliably available (Risko & Gilbert, 2016), and expected future access lowers memory for the information itself (Sparrow et al., 2011). Generative AI extends offloading to reasoning and composition, potentially removing the productive effort that cognitive load theory associates with durable learning (Sweller et al., 2019). Students using large language models report lower cognitive load but produce weaker justifications (Stadler et al., 2024), show signs of metacognitive laziness (Fan et al., 2025), and exhibit reduced neural engagement and weaker recall of their own AI-assisted essays (Kosmyna et al., 2025). AI outputs may also create illusions of understanding (Messeri & Crockett, 2024), which could explain why 59.78% of students in our sample agreed that they submit work without fully understanding AI-generated content. The quadratic pattern, in which the largest increase occurred between less than 1 h/day and 1-3 h/day, suggests that the shift from occasional to routine use may matter more than additional hours beyond routine use. Because the data are cross-sectional, reverse causation is plausible: students who already feel less confident or more pressured may turn to AI more often. Longitudinal evidence that prior mental health problems predict later AI dependence (Huang et al., 2024) and that academic stress mediates the link between self-efficacy and AI dependency (Zhang et al., 2024) supports this interpretation.

The sex and academic-level interactions add nuance. Male students reported higher vulnerability overall, consistent with evidence that men adopt and use AI chatbots more frequently and view them more positively (Møgelvang et al., 2024; Stöhr et al., 2024). The male-female gap, however, was largest among light users (37.66 vs. 30.11) and almost disappeared among heavy users (74.88 vs. 73.09). Sex differences may therefore reflect attitudes and patterns of use at low intensity, whereas intensive use is associated with similarly high vulnerability regardless of sex. Master's students scored higher at low use but had a flatter gradient, possibly because advanced students use AI more selectively or have more established study routines; Stöhr et al. (2024) likewise reported higher use among master's students. The absence of effect modification by university type suggests that the association is not confined to the private sector, which dominates the sample.

Contrary to H3, the AIDAV index was not associated with self-reported CGPA. This contrasts with Abbas et al. (2024), who reported negative effects of ChatGPT use on academic performance. Several explanations are possible. CGPA is cumulative and banded, and much of it may have been earned before intensive AI use. Grades may also reward the products that AI helps produce, in line with experimental evidence that AI assistance improves task scores during use without corresponding learning gains (Bastani et al., 2025; Fan et al., 2025). If so, grades may be an insensitive indicator of the learning deficits that students themselves perceive. The positive bivariate association between heavy use and a high CGPA, which disappeared after adjustment, probably reflects confounding by academic level and age.

In the Bangladeshi context, the findings quantify concerns raised in qualitative and acceptance studies. Hasib and Islam (2025) described over-reliance and unattributed use among university students, and Rahman et al. (2025) showed that peer influence drives student adoption. The present results indicate that these concerns are widespread: two-thirds of students met the threshold for high dependency-vulnerability, and about half did not verify AI outputs.

5.1 Strengths and limitations

The study has several strengths: a large sample from 30 institutions, a psychometrically evaluated composite outcome supported by both exploratory and confirmatory factor analysis, explicit modeling of clustering by university, prevalence ratios for a common binary outcome, item-level and partial proportional odds analyses, tests of effect modification, and multiple sensitivity analyses that addressed duplicate records, a dominant institution, heteroskedasticity, and the bounded outcome.

Several limitations should be considered. First, the cross-sectional design precludes causal inference and the determination of temporal order. Second, all measures were self-reported, and common-method variance may have inflated associations between usage time and perceived vulnerability (Podsakoff et al., 2003). Third, recruitment through social media produced a non-probability sample in which one private university contributed 38.04% of respondents; no survey weights were available, and results may not generalize to all Bangladeshi students. Fourth, 14.67% of records were exact duplicates; although the estimates were robust to their exclusion, their origin could not be verified. Fifth, usage time was measured in three broad categories, which limits dose-response modeling, and CGPA was self-reported in bands. Sixth,

important confounders, including field of study, academic workload, self-efficacy, socioeconomic status, and mental health, were not available. Seventh, the nine items were developed by the original investigators, and the AIDAV index combines reliance behaviors with perceived consequences; measurement invariance across groups was not tested, and the two reverse-worded items shared method variance. Eighth, inference relied on 30 unbalanced clusters, and a wild cluster bootstrap could not be executed, although the HC3 and multilevel analyses gave consistent results. Finally, the logistic model showed some miscalibration, and the parallel-lines assumption was violated in the ordinal models, although the conclusions were unchanged in partial proportional odds models.

5.2 Implications and directions for future research

For practice, universities in Bangladesh should integrate AI literacy into curricula, with explicit training in verifying, critiquing, and acknowledging AI outputs, in line with international guidance (Miao & Holmes, 2023). Assessment designs that make the learning process visible, such as staged submissions, oral defenses, and in-class reasoning tasks, may counter the tendency to submit work that is not understood (Cotton et al., 2024). Institutions can also favor AI tutors with pedagogical guardrails, which have been shown to mitigate learning losses (Bastani et al., 2025). Because the steepest increase occurred between occasional and routine use, guidance on purposeful and bounded use may be more realistic than prohibition. For policy, national bodies overseeing higher education could develop sector-wide guidelines that apply to both public and private universities.

Future research should use longitudinal or experimental designs to establish temporal order; measure AI use objectively, for example with usage logs; include validated scales of critical thinking, self-efficacy, and academic stress; test measurement invariance of the AIDAV index across sex, academic level, and institution type; and link dependency to objective learning outcomes rather than cumulative grades alone. Probability-based samples covering public, private, and national universities in Bangladesh would also strengthen generalizability.

6. Conclusion

Among Bangladeshi university students, longer daily use of generative AI was strongly associated with greater self-reported dependency and academic vulnerability. The association was graded, robust to alternative specifications, and present in both public and private

universities, and it was strongest for not being able to study without AI and not verifying AI outputs. Dependency-vulnerability was not associated with self-reported CGPA, suggesting that grades may not capture the learning deficits students perceive. These findings support AI-literacy education, verification-focused practices, and assessment redesign that preserves independent learning. Longitudinal and experimental research is needed to establish whether intensive AI use leads to these vulnerabilities or whether vulnerable students turn to AI more often.

Declarations

Author contributions (CRediT). Meshkat Ebnay Mozahid: Conceptualization; Methodology; Investigation; Data curation; Formal analysis; Software; Visualization; Writing - original draft; Writing - review & editing. Md. Anis Mahmud Sakib:

Funding. This research received no external funding.

Conflict of interest. The authors declare no conflict of interest.

Ethics approval. Not applicable. This study is a secondary analysis of anonymized, publicly available data (Mendeley Data, CC BY 4.0) and involved no contact with participants or access to identifying information.

Consent to participate. Not applicable to the present secondary analysis. According to the data descriptor, participation in the original survey was voluntary, and respondents were informed of the academic purpose of the study before completing the questionnaire (Islam et al., 2026).

Consent for publication. Not applicable. The manuscript contains no individual-level identifiable data.

Data availability. The data analyzed in this study are publicly available in Mendeley Data: Islam, Sharmin, and Tuhin (2026), “AI Dependency and Academic Vulnerability in Higher Education A Survey Dataset from Bangladesh,” Version 1, https://doi.org/10.17632/6s6nvmpgbb.1. The Stata 17 do-file used for the analysis is available from the corresponding author on reasonable request.

Use of generative AI. During manuscript preparation, generative AI tools were used only for language editing and improving readability. These tools were not used for data analysis, interpretation, or generation of scientific content. The authors take full responsibility for the content and integrity of the manuscript.

Acknowledgements. The authors thank the creators of the dataset for making it publicly available.

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*, Article 10. https://doi.org/10.1186/s41239-024-00444-7

Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. *Proceedings of the National Academy of Sciences*, *122*(26), Article e2422633122. https://doi.org/10.1073/pnas.2422633122

Brant, R. (1990). Assessing proportionality in the proportional odds model for ordinal logistic regression. *Biometrics*, *46*(4), 1171-1178. https://doi.org/10.2307/2532457

Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, *50*(2), 317-372. https://doi.org/10.3368/jhr.50.2.317

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, *20*, Article 43. https://doi.org/10.1186/s41239-023-00411-8

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, *58*(1), 7-19. https://doi.org/10.1093/analys/58.1.7

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, *61*(2), 228-239. https://doi.org/10.1080/14703297.2023.2190148

Darvishi, A., Khosravi, H., Sadiq, S., Gašević, D., & Siemens, G. (2024). Impact of AI assistance on student agency. *Computers & Education*, *210*, Article 104967. https://doi.org/10.1016/j.compedu.2023.104967

Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., Shen, Y., Li, X., & Gašević, D. (2025). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. *British Journal of Educational Technology*, *56*(2), 489-530. https://doi.org/10.1111/bjet.13544

Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies*, *15*(1), Article 6. https://doi.org/10.3390/soc15010006

Hasib, M., & Islam, M. S. (2025). How university students in Bangladesh engage with ChatGPT: A qualitative study. *PLOS ONE*, *20*(9), Article e0333089. https://doi.org/10.1371/journal.pone.0333089

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Huang, S., Lai, X., Ke, L., Li, Y., Wang, H., Zhao, X., Dai, X., & Wang, Y. (2024). AI technology panic-Is AI dependence bad for mental health? A cross-lagged panel model and the mediating roles of motivations for AI use among adolescents. *Psychology Research and Behavior Management*, *17*, 1087-1102. https://doi.org/10.2147/PRBM.S440889

Islam, M. J., Sharmin, S., & Tuhin, H. H. (2026). *AI dependency and academic vulnerability in higher education: A survey dataset from Bangladesh* (Version 1) [Data set]. Mendeley Data. https://doi.org/10.17632/6s6nvmpgbb.1

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I., & Maes, P. (2025). *Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2506.08872

Lee, H.-P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (Article 1121). Association for Computing Machinery. https://doi.org/10.1145/3706598.3713778

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, *13*(4), Article 410. https://doi.org/10.3390/educsci13040410

Long, J. S., & Ervin, L. H. (2000). Using heteroscedasticity consistent standard errors in the linear regression model. *The American Statistician*, *54*(3), 217-224. https://doi.org/10.1080/00031305.2000.10474549

Messeri, L., & Crockett, M. J. (2024). Artificial intelligence and illusions of understanding in scientific research. *Nature*, *627*(8002), 49-58. https://doi.org/10.1038/s41586-024-07146-0

Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO. https://doi.org/10.54675/EWZM9535

Møgelvang, A., Bjelland, C., Grassini, S., & Ludvigsen, K. (2024). Gender differences in the use of generative artificial intelligence chatbots in higher education: Characteristics and consequences. *Education Sciences*, *14*(12), Article 1363. https://doi.org/10.3390/educsci14121363

Naher, J. (2023). A survey to understand the experience of ChatGPT usage among engineering university students in Bangladesh. *Technium*, *15*(1), 1-8. https://doi.org/10.47577/technium.v15i.9660

Papke, L. E., & Wooldridge, J. M. (1996). Econometric methods for fractional response variables with an application to 401(k) plan participation rates. *Journal of Applied Econometrics*, *11*(6), 619-632. https://doi.org/10.1002/(SICI)1099-1255(199611)11:6<619::AID-JAE418>3.0.CO;2-1

Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors*, *52*(3), 381-410. https://doi.org/10.1177/0018720810376055

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879-903. https://doi.org/10.1037/0021-9010.88.5.879

Rahman, A., Islam, M. K., Al-Mamun, A., & Islam, M. S. (2025). Teachers' and students' use of ChatGPT at social science faculty in the public and private universities of Bangladesh [version 2; peer review: 2 approved]. *F1000Research*, *14*, Article 269. https://doi.org/10.12688/f1000research.161351.2

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

Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google effects on memory: Cognitive consequences of having information at our fingertips. *Science*, *333*(6043), 776-778. https://doi.org/10.1126/science.1207745

Stadler, M., Bannert, M., & Sailer, M. (2024). Cognitive ease at a cost: LLMs reduce mental effort but compromise depth in student scientific inquiry. *Computers in Human Behavior*, *160*, Article 108386. https://doi.org/10.1016/j.chb.2024.108386

Stöhr, C., Ou, A. W., & Malmström, H. (2024). Perceptions and usage of AI chatbots among students in higher education across genders, academic levels and fields of study. *Computers and Education: Artificial Intelligence*, *7*, Article 100259. https://doi.org/10.1016/j.caeai.2024.100259

Strzelecki, A. (2024). Students' acceptance of ChatGPT in higher education: An extended unified theory of acceptance and use of technology. *Innovative Higher Education*, *49*, 223-245. https://doi.org/10.1007/s10755-023-09686-1

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, *31*(2), 261-292. https://doi.org/10.1007/s10648-019-09465-5

von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., & Vandenbroucke, J. P. (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *The Lancet*, *370*(9596), 1453-1457. https://doi.org/10.1016/S0140-6736(07)61602-X

Williams, R. (2006). Generalized ordered logit/partial proportional odds models for ordinal dependent variables. *The Stata Journal*, *6*(1), 58-82. https://doi.org/10.1177/1536867X0600600104

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education - where are the educators? *International Journal of Educational Technology in Higher Education*, *16*, Article 39. https://doi.org/10.1186/s41239-019-0171-0

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. *Smart Learning Environments*, *11*, Article 28. https://doi.org/10.1186/s40561-024-00316-7

Zhang, S., Zhao, X., Zhou, T., & Kim, J. H. (2024). Do you have AI dependency? The roles of academic self-efficacy, academic stress, and performance expectations on problematic AI usage behavior. *International Journal of Educational Technology in Higher Education*, *21*, Article 34. https://doi.org/10.1186/s41239-024-00467-0

Zou, G. (2004). A modified Poisson regression approach to prospective studies with binary data. *American Journal of Epidemiology*, *159*(7), 702-706. https://doi.org/10.1093/aje/kwh090

Appendix A. Supplementary tables and figures

Table A1

Item wording, distribution, and psychometric properties of the AIDAV items (N = 1,104)

Item wording (original direction)	Mean (SD)	Agree %	Disagree %	Item-rest r	α if deleted	PCF loading	CFA λ
Q1 I verify AI-generated information before using it in academic work. (R)	2.85 (1.09)	39.13	50.27	0.716	0.931	0.775	0.701
Q2 I can study and complete tasks effectively without any AI support. (R)	2.82 (1.19)	37.59	49.73	0.733	0.930	0.791	0.719
Q3 My first instinct for any academic doubt is to ask an AI rather than checking books or notes.	3.43 (1.15)	59.24	30.43	0.693	0.932	0.756	0.710
Q4 I find it harder to recall information that I have previously generated using AI.	3.41 (1.16)	57.34	28.35	0.791	0.926	0.840	0.810
Q5 When using AI, I prioritize completing assignments quickly rather than fully understanding the subject matter.	3.40 (1.11)	56.16	27.45	0.752	0.929	0.808	0.789
Q6 I often use AI-generated responses without making any major modifications.	3.37 (1.09)	54.17	27.81	0.751	0.929	0.807	0.789
Q7 I feel less confident in my writing and analytical skills when I am not using AI.	3.32 (1.12)	55.43	29.26	0.771	0.927	0.824	0.812
Q8 My critical thinking ability has weakened due to reliance on AI tools.	3.41 (1.12)	57.88	26.36	0.795	0.926	0.845	0.837
Q9 I often submit academic work without fully understanding the content generated by AI.	3.44 (1.41)	59.78	29.35	0.831	0.924	0.875	0.876

Note. Mean (SD) and agree/disagree percentages refer to the original item direction; agree = agree or strongly agree; disagree = disagree or strongly disagree. Item-rest correlations, α if item deleted, principal-component factor (PCF) loadings, and standardized CFA loadings (λ; model with correlated residuals for Q1-Q2) refer to the scored direction, with Q1 and Q2 reverse-scored (R). Overall α = 0.936; KMO = 0.946; first eigenvalue = 5.97 (66.28% of variance).

Table A2

Sensitivity analyses of the association between daily AI use and the AIDAV index

Variable	HC3 OLS β (95% CI)	Excluding duplicates β (95% CI)	Excluding DIU β (95% CI)	Fractional logit AME (95% CI)
1-3 h/day	31.36 (28.69, 34.03)	28.07 (22.69, 33.44)	30.70 (25.24, 36.16)	0.313 (0.282, 0.345)
> 3 h/day	40.21 (37.77, 42.65)	36.62 (30.82, 42.42)	39.91 (35.25, 44.58)	0.402 (0.377, 0.428)
Male	3.81 (1.81, 5.80)	3.01 (1.51, 4.50)	3.26 (1.10, 5.42)**	0.038 (0.026, 0.051)
Age 24-26	−1.12 (−3.48, 1.24)	−2.11 (−4.62, 0.39)	−2.12 (−5.50, 1.26)	−0.011 (−0.036, 0.014)
Age above 26	−2.15 (−6.54, 2.24)	−2.67 (−4.77, −0.58)*	−3.07 (−6.63, 0.49)	−0.021 (−0.047, 0.004)
Graduate	1.87 (−1.45, 5.18)	1.38 (−1.55, 4.30)	0.83 (−2.91, 4.58)	0.019 (−0.010, 0.047)
Master's	1.03 (−1.81, 3.88)	0.61 (−2.65, 3.87)	−0.66 (−3.95, 2.62)	0.010 (−0.023, 0.044)
CGPA 3.01-3.50	1.26 (−1.30, 3.81)	2.04 (−0.78, 4.85)	1.81 (−2.00, 5.62)	0.013 (−0.012, 0.037)
CGPA above 3.50	0.32 (−2.69, 3.33)	0.61 (−1.86, 3.09)	0.41 (−3.52, 4.35)	0.003 (−0.023, 0.029)
Public university	2.07 (0.02, 4.11)*	0.76 (−2.63, 4.15)	−1.05 (−4.84, 2.74)	0.021 (−0.012, 0.053)
N (clusters)	1,103	941 (30)	684 (29)	1,103 (30)
R^2 / pseudo R^2	0.5727	0.4730	0.5261	0.0981

Note. References as in Table 3. Columns 2-4: linear models of AIDAV (0-100); column 2 uses HC3 standard errors, and columns 3 and 4 use university-clustered standard errors. Column 5: average marginal effects from fractional logit regression of AIDAV rescaled to 0-1, with university-clustered standard errors. DIU = Daffodil International University. ∗ $p < .05$; ∗∗ $p < .01$; ∗∗∗ $p < .001$.

Table A3

Ordinal logistic regression of CGPA category (3.00 or below; 3.01-3.50; above 3.50) on the AIDAV index

Variable	aOR (95% CI)	p
AIDAV index (per SD)	1.01 (0.86, 1.19)	.881
Daily AI use 1-3 h/day (ref. < 1 h/day)	0.94 (0.59, 1.52)	.814
Daily AI use > 3 h/day	1.31 (0.92, 1.88)	.135
Male (ref. female)	0.80 (0.60, 1.06)	.123
Age 24-26 (ref. 23 or younger)	1.32 (1.04, 1.69)	.025
Age above 26	1.53 (0.97, 2.43)	.069
Graduate (ref. undergraduate)	1.22 (0.73, 2.04)	.457
Master's	1.51 (1.14, 1.99)	.004
Public university (ref. private)	0.97 (0.76, 1.25)	.837
Partial proportional odds model: AIDAV (per SD)	1.02 (0.87, 1.19)	.803

Note. N = 1,103; standard errors clustered by university (30 clusters); pseudo R^2 = 0.0146. Brant test of parallel lines $\chi^2(9)$ = 43.65, $p < .001$; the partial proportional odds model (gologit2, autofit) constrained the AIDAV coefficient to be parallel across thresholds.

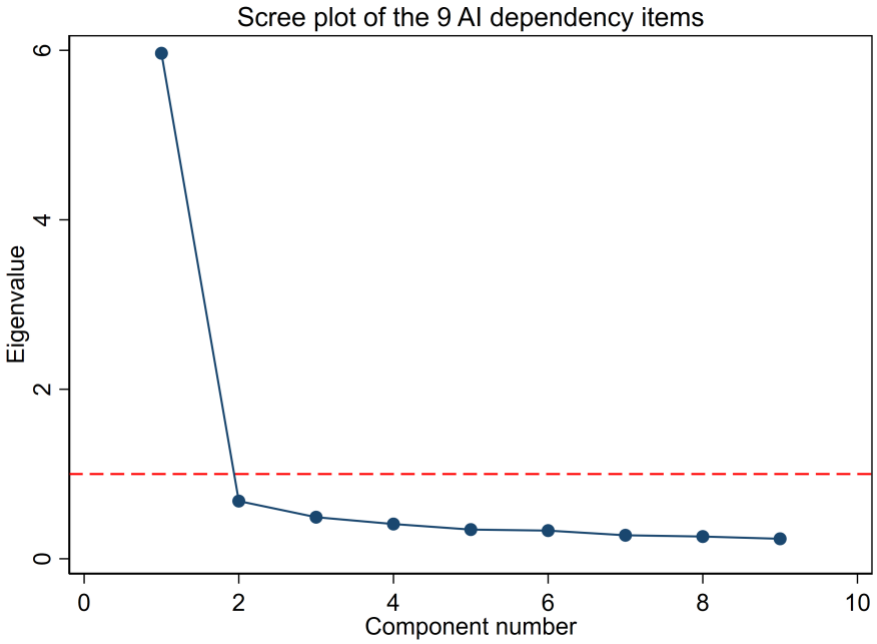


Figure A1. Scree plot of the nine AIDAV items. The dashed line marks an eigenvalue of 1.

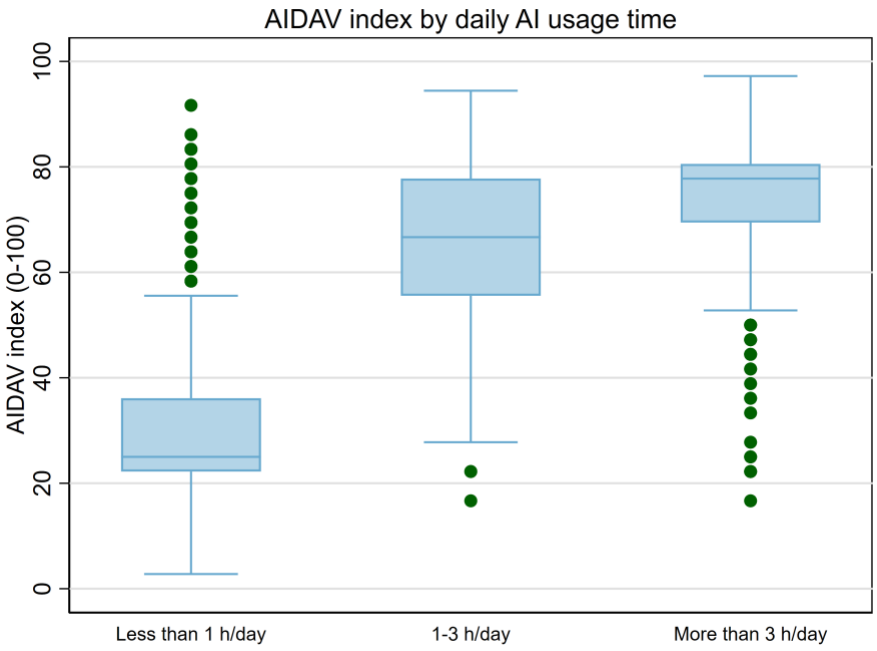


Figure A2. Box plots of the AIDAV index by daily AI usage time.

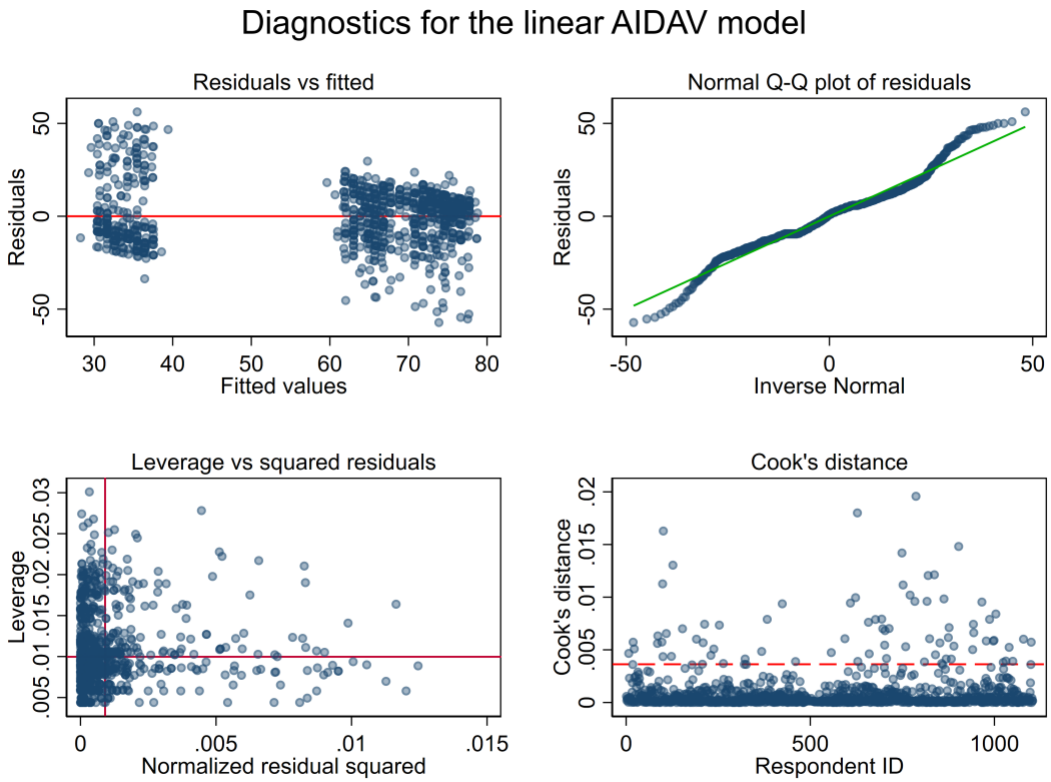


Figure A3. Regression diagnostics for the linear AIDAV model: residuals versus fitted values, normal Q-Q plot, leverage versus squared residuals, and Cook's distance (dashed line = 4/N). The banded residual pattern reflects the categorical predictors.

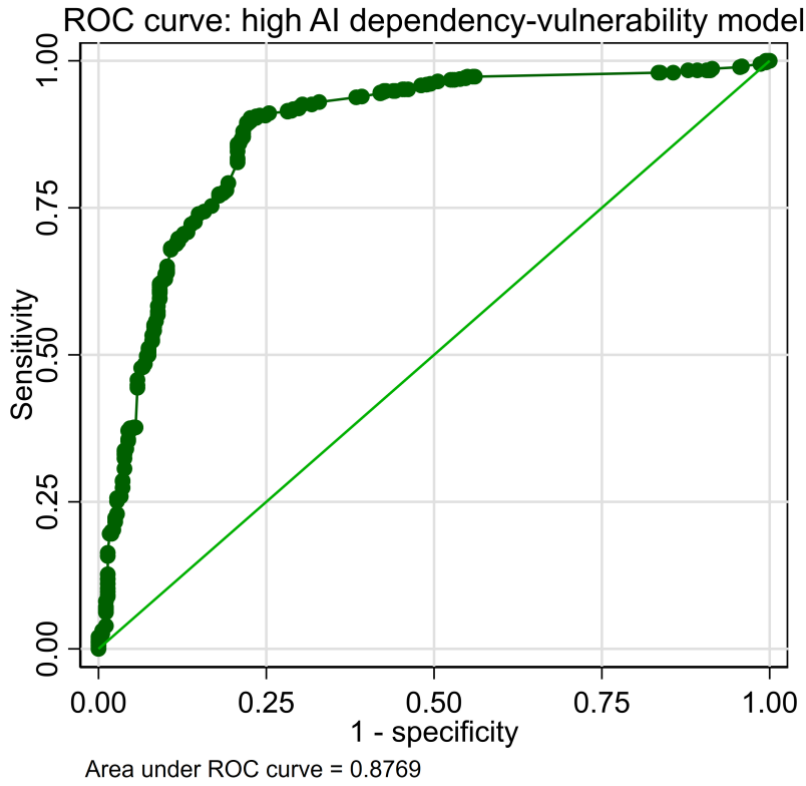


Figure A4. Receiver operating characteristic curve for the logistic model of high AI dependency-vulnerability (AUC = 0.8769).