Journal of Curriculum Development Studies
Volume: 5 Issue: 4 Year: 2026
ISSN: 2957-4161



KMF Publishers
www.kmf-publishers.com/jocds/

DOI: https://doi.org/10.64907/xkmf.v5i4.jocds.1

Journal of Curriculum Development Studies

RESEARCH ARTICLE

OPEN ACCESS
Freely available online

Received: 20 September 2026
Revised: 25 October 2026
Accepted: 28 October 2026
Published: 30 October 2026

Corresponding author:
Sifa Sultana
sifasarker0@gmail.com

[1]Department of Law
[2]Department of Business Administration
Shanto-Mariam University of Creative Technology
Dhaka, Bangladesh

Institutional AI Policy Awareness and Responsible AI Use in Legal Education in Bangladesh

Sifa Sultana[1*]; Mst. Sadia Rahman Arpita[1]; Parvas Mia[1]; Kazi Abdul Mannan[2]

Abstract: Institutional guidance on generative artificial intelligence is often evaluated through its existence or students' general support for regulation. Neither establishes whether awareness of rules corresponds with responsible use. This exploratory study separates perceived knowledge of institutional rule status, anticipated policy benefits and behaviour-specific self-reports in a Bangladesh legal-education database. The complete database contains N = 385 records; its n = 245 current law students form the primary analytical subsample, with all records retained for sensitivity analysis. The principal outcome is endorsement of using AI only as permitted by course or institutional rules. Item-level logistic models, robust intervals, respondent bootstrap standardisation and repeated stratified cross-validation examine association and incremental predictive value. Policy awareness is endorsed by 151 students (61.6%); permitted-use compliance by 150 (61.2%). The student association is positive but uncertain (adjusted odds ratio 1.33; 95% confidence interval 0.95-1.86); the complete-sample estimate is likewise uncertain (1.17; 0.89-1.53). Standardised endorsement probabilities at awareness scores four and three differ by 6.7 percentage points (95% bootstrap interval –0.9-14.7). Associations with independent-work boundaries and avoiding wholly personal attribution are also uncertain. Adding awareness changes the mean held-out Brier score from 0.24308 to 0.24268, while an intercept-only benchmark performs better at 0.23740. Support for training and consistent enforcement is substantially more common than awareness endorsement. Low internal consistency precludes treating these items as one reflective policy construct. The codebook labels identifiers and comments synthetic, leaving numerical provenance unresolved; comments are excluded. The findings identify limits in using perceived awareness as a surrogate for compliance and motivate assessment-specific communication and prospective evaluation. They establish neither institutional policy effectiveness nor misconduct prevalence.

Keywords: institutional AI policy; legal education; academic integrity; policy awareness; permitted use; predictive validation; Bangladesh

Citation information
Cite this article as Sultana, S., Arpita, M. S. R., Mia, P., & Mannan, K. A. (2026). Institutional AI Policy Awareness and Responsible AI Use in Legal Education in Bangladesh. *Journal of Curriculum Development Studies, 5*(4). DOI: https://doi.org/10.64907/xkmf.v5i4.jocds.1

2026 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.

1. Introduction

Institutional AI policy enters a law student's education through concrete decisions about assignments. A learner may use a language model to understand a concept, organise a legal argument, translate a passage or draft an answer. Whether that assistance is acceptable depends partly on the educational purpose and the instructions governing the task. Universities therefore need more than a general position on artificial intelligence. They need a way to communicate permitted assistance that students can interpret and apply. Research on generative AI in education documents both opportunities for support and concerns about independent learning and assessment integrity.[1]

A central difficulty is deciding what counts as evidence that institutional guidance works. A published policy demonstrates that a document exists. Agreement that clearer rules would help demonstrates an evaluation of a possible arrangement. A student's statement that they know whether rules exist demonstrates perceived awareness of rule status. These observations concern different stages of institutional governance. None, on its own, establishes that the student understands a particular assessment instruction or follows it in practice. Policy research becomes more useful when it separates those stages rather than treating all favourable responses as evidence of successful implementation.

Legal education heightens the importance of this distinction. Students learn to identify applicable authorities, reason from sources and accept responsibility for legal propositions. Generated language can support those activities while also supplying plausible but unreliable material. Evidence of legal hallucinations provides a reason for careful source checking, but it does not determine the permissible amount of assistance in a particular assignment. An institution must communicate the boundary between support and substitution as well as the standards of legal reasoning expected from the student.[2]

Recent university-policy studies have mapped guidance, identified recurring themes and proposed educational frameworks. Chan connects institutional policy with pedagogical and operational considerations. An, Yu and James examine guidance across university activities, while McDonald and colleagues analyse institutional policy documents. Such work clarifies what universities choose to communicate. It also leaves an empirical question about the relationship between a student's perceived knowledge of rules and

[1] Enkelejda Kasneci and others, 'ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education' (2023) 103 Learning and Individual Differences 102274; Debby RE Cotton, Peter A Cotton and J Reuben Shipway, 'Chatting and Cheating: Ensuring Academic Integrity in the Era of ChatGPT' (2024) 61 Innovations in Education and Teaching International 228; Mike Perkins, 'Academic Integrity Considerations of AI Large Language Models in the Post-Pandemic Era: ChatGPT and Beyond' (2023) 20(2) Journal of University Teaching and Learning Practice, DOI: 10.53761/1.20.02.07.

[2] Matthew Dahl and others, 'Large Legal Fictions: Profiling Legal Hallucinations in Large Language Models' (2024) 16 Journal of Legal Analysis 64.

specific reported practices. Document analysis and student-level behavioural analysis offer complementary evidence; they cannot substitute for one another.[3]

The Group 4 research plan assigns Sifa Sultana the topic of institutional AI policy awareness and responsible AI use in legal education in Bangladesh. The associated questionnaire contains one item directly addressing knowledge of institutional rules and several items addressing different integrity practices. It also contains statements about clearer rules, training, confidentiality restrictions and enforcement. This configuration permits a focused analysis, provided the statements are not combined into a single policy-effectiveness score. The present article treats their different wording as a substantive feature of the research design.

The main question is whether greater perceived awareness corresponds with endorsement of using AI only as permitted by the course or institution. Two additional outcomes address recognising independent-work boundaries and avoiding the representation of substantially generated work as entirely one's own. These outcomes keep the analysis within the relationship between institutional permission and personal academic responsibility. Disclosure and legal-citation verification are not the primary outcomes, preserving a distinct contribution within the shared Group 4 project.

The analysis further asks whether awareness adds useful information when predicting responses outside the observations used to estimate a model. An association estimated in the full sample does not automatically imply useful prediction. Conversely, weak predictive performance cannot settle whether better communication would change behaviour. Separating these questions is important when administrators consider using awareness surveys as a monitoring instrument. The study therefore reports effect estimates and their uncertainty alongside repeated held-out performance against a simple benchmark.[4]

The contextual claim is deliberately bounded. Bangladesh supplies the educational setting described in the research materials, but the workbook does not establish a national probability sample. Its 385 records include graduates, legal professionals and other respondents as well as students. The total database sample is N = 385, and the primary analytical subsample comprises n = 245 current law students (63.6%). The remaining 140 records (36.4%) are retained in a separate complete-sample sensitivity analysis; they are not classified as current students. Hasib and Islam's qualitative research provides relevant evidence about university students'

[3] Cecilia Ka Yuk Chan, 'A Comprehensive AI Policy Education Framework for University Teaching and Learning' (2023) 20 International Journal of Educational Technology in Higher Education 38; Yunjo An, Ji Hyun Yu and Shadarra James, 'Investigating the Higher Education Institutions' Guidelines and Policies Regarding the Use of Generative AI in Teaching, Learning, Research, and Administration' (2025) 22 International Journal of Educational Technology in Higher Education 10; Nora McDonald and others, 'Generative Artificial Intelligence in Higher Education: Evidence from an Analysis of Institutional Policies and Guidelines' (2025) 3 Computers in Human Behavior: Artificial Humans 100121.

[4] Galit Shmueli, 'To Explain or to Predict?' (2010) 25 Statistical Science 289.

engagement with ChatGPT in Bangladesh, without establishing the prevalence of the outcomes studied here.[5]

Three contributions follow. First, the article distinguishes knowledge of rule status from support for hypothetical institutional arrangements. Second, it estimates awareness associations separately for three specified practices, avoiding a latent integrity construct unsupported by the available responses. Third, it examines incremental predictive performance rather than treating an in-sample coefficient as sufficient evidence for institutional monitoring. These contributions concern analytical precision and the limits of inference; the design cannot identify the causal effectiveness of a university policy.

A further boundary concerns provenance. The codebook describes respondent identifiers and optional contextual comments as synthetic. That description does not establish whether the numerical responses are synthetic, but it prevents treating their empirical origin as verified. The article excludes comments and reports calculations conditionally on the supplied records. Original collection documentation, author declarations and applicable ethics arrangements must be verified before empirical submission. Preserving this distinction is part of research quality, not a reason to replace uncertain findings with stronger language. The following sections develop the framework, explain the analysis, present the results and identify the institutional evidence that a prospective study should collect.

2. Literature Review and Theoretical Foundation

The theoretical foundation connects institutional communication with behaviour-specific integrity reasoning. It draws on the theory of planned behaviour and academic integrity scholarship while distinguishing the constructs those traditions require from the items actually available. The objective is an interpretable account of policy awareness, reported practice and predictive information, rather than a comprehensive model of institutional effectiveness.

2.1 Institutional policy as communication and educational design

AI guidance serves several purposes. It can describe acceptable assistance, specify reporting expectations, protect sensitive information and explain how assessment decisions will be made. These purposes require coordination between university guidance and the instructions attached to an individual task. A general permission statement may leave uncertainty when an assignment requires independent analysis, while a broad prohibition may leave ordinary language support unclear. The educational challenge is therefore to translate principles into decisions students can make in the context of their work.

Chan's framework treats AI policy as a matter that spans educational practice and institutional operation. Its value for this study is the recognition that policy cannot be reduced to a single prohibition. Moore and

[5] Mir Hasib and Md Shariful Islam, 'How University Students in Bangladesh Engage with ChatGPT: A Qualitative Study' (2025) 20 PLOS ONE e0333089.

Lookadoo similarly foreground clear communication in policy development. Their work supports examining the relationship between communication and interpretation, while it does not establish that a particular awareness item measures students' actual comprehension of institutional guidance.[6]

An, Yu and James analyse guidance across teaching, learning, research and administration. Wang and colleagues examine policies, resources and guidelines in universities, and McDonald and colleagues identify recurring concerns in institutional documents. Together these studies provide evidence that AI governance involves multiple organisational activities. Their document-based designs primarily reveal the content and organisation of guidance, rather than how a particular student responds to an assignment. The present analysis addresses that complementary student-level question without assuming that the institutions represented in the workbook have comparable policies.[7]

A further distinction concerns the level at which rules are recognised. Jiang, Xie and Cao investigate institutional and classroom regulation and report differences in how students and faculty perceive their clarity and consequences. Their findings make it plausible that students can recognise course instructions while remaining uncertain about university guidance. The outcome of reducing AI use is also different from the outcome of using it responsibly: permitted use can increase when a policy makes legitimate assistance clearer. Accordingly, the present study does not equate frequent use with non-compliance.[8]

The primary awareness statement asks whether respondents know whether their institution has rules or guidance. A respondent could agree because they know that no institutional guidance exists. Agreement therefore cannot be interpreted as confirmation of policy presence. Nor does it establish recall of policy wording, confidence in applying it, access to an instructor or the legitimacy of enforcement. This wording sets the boundary of the exposure: perceived knowledge of rule status. Its relationship with permitted-use endorsement is worth examining, but its meaning should remain narrower than effective policy implementation.

[6] Chan (n 3); Sarah Moore and Kathryn Lookadoo, 'Communicating Clear Guidance: Advice for Generative AI Policy Development in Higher Education' (2024) 87 Business and Professional Communication Quarterly 610.

[7] An, Yu and James (n 3); Hui Wang and others, 'Generative AI in Higher Education: Seeing ChatGPT through Universities' Policies, Resources, and Guidelines' (2024) 7 Computers and Education: Artificial Intelligence 100326; McDonald and others (n 3).

[8] Yingying Jiang, Lindai Xie and Xiaoyu Cao, 'Exploring the Effectiveness of Institutional Policies and Regulations for Generative AI Usage in Higher Education' (2025) 79 Higher Education Quarterly e70054.

2.2 Planned behaviour and the place of awareness

The theory of planned behaviour explains a specified behaviour through attitudes, subjective norms, perceived behavioural control and intentions. Ajzen's formulation is relevant because policy compliance is an action taken under perceived expectations and practical conditions. Beck and Ajzen's work on dishonest actions illustrates the applicability of the framework to integrity-related decisions. The theory also requires careful specification of the behaviour under study; an intention to use AI responsibly in general is less precise than an intention to follow the permission rules for a particular task.[9]

Awareness can plausibly provide information needed to evaluate an action, but information is not equivalent to any one planned-behaviour construct. Knowing that guidance exists differs from approving its content. Knowing an instruction differs from perceiving that peers follow it. Understanding a rule differs from having sufficient time, language support or access to approved tools to comply. A policy-awareness coefficient consequently cannot be read as a test of attitude, subjective norm or perceived control.

The available questionnaire does not separately measure these constructs with validated multi-item scales or assess intentions before subsequent behaviour. The study therefore uses planned-behaviour reasoning to specify outcomes and interpret possible mechanisms, rather than claiming to test the full theory. Training attendance and AI-use frequency enter as measured covariates. They do not become objective competence or behavioural control merely because those concepts would be theoretically useful.

A positive awareness association is nevertheless a reasonable exploratory expectation. Students who perceive themselves as informed about the institutional position may be better able to match their use to what they believe is permitted. The same pattern could arise because conscientious students seek guidance, because instructors communicate more clearly in some courses or because respondents give similarly favourable answers to both items. Cross-sectional self-report cannot distinguish these mechanisms. The hypothesis therefore concerns an observed association, not the direction or size of an intervention effect.

[9] Icek Ajzen, 'The Theory of Planned Behavior' (1991) 50 Organizational Behavior and Human Decision Processes 179; Lisa Beck and Icek Ajzen, 'Predicting Dishonest Actions Using the Theory of Planned Behavior' (1991) 25 Journal of Research in Personality 285.

2.3 Academic integrity and specified practices

Academic integrity provides the normative foundation for interpreting permitted use. The International Center for Academic Integrity identifies values including honesty, fairness and responsibility. These values concern how students represent their work and participate in assessment, rather than whether a particular technology is inherently acceptable. In an AI-assisted assignment, responsibility includes understanding the task, recognising what must remain independent and avoiding a misleading claim to wholly personal production.[10]

Earlier integrity research places behaviour within institutional and peer environments. McCabe and Treviño examine contextual influences associated with academic dishonesty, and later work reviews the accumulated evidence. Whitley's review likewise indicates that integrity behaviour is associated with multiple factors. These findings caution against treating a single awareness response as a stable personal characteristic or an exhaustive explanation of conduct. Institutional design and student judgement can both matter without being separately identifiable in this workbook.[11]

Contemporary AI integrity scholarship argues for clearer expectations and assessment design adapted to available tools. Perkins discusses the integrity implications of language models, while ENAI recommendations situate ethical AI use within educational responsibility. The AI Assessment Scale offers a practical framework for differentiating permitted levels of assistance. Its pilot implementation illustrates how assessment expectations can be made more explicit. These are relevant educational approaches, but the Group 4 questionnaire does not measure participation in their implementation or permit an evaluation of their effectiveness.[12]

The two secondary outcomes address related but distinct practices. Recognising the boundary between assistance and independent work concerns interpreting the task. Avoiding the submission of substantially generated work as entirely one's own concerns the representation of authorship. A student could endorse one statement without endorsing the other. Johnston and colleagues' student-perspective study and Gruenhagen and colleagues' research on assessment assistance support examining such distinctions rather than applying a universal classification to all AI use.[13]

[10] International Center for Academic Integrity, The Fundamental Values of Academic Integrity (3rd edn, ICAI 2021).

[11] Donald L McCabe and Linda Klebe Treviño, 'Academic Dishonesty: Honor Codes and Other Contextual Influences' (1993) 64 Journal of Higher Education 522; Donald L McCabe, Linda Klebe Treviño and Kenneth D Butterfield, 'Cheating in Academic Institutions: A Decade of Research' (2001) 11 Ethics & Behavior 219; Bernard E Whitley Jr, 'Factors Associated with Cheating among College Students: A Review' (1998) 39 Research in Higher Education 235.

[12] Perkins, 'Academic Integrity Considerations' (n 1); Tomas Foltynek and others, 'ENAI Recommendations on the Ethical Use of Artificial Intelligence in Education' (2023) 19 International Journal for Educational Integrity 12; Mike Perkins and others, 'The Artificial Intelligence Assessment Scale (AIAS): A Framework for Ethical Integration of Generative AI in Educational Assessment' (2024) 21(6) Journal of University Teaching and Learning Practice, DOI: 10.53761/q3azde36; Leon Furze and others, 'The AI Assessment Scale (AIAS) in Action: A Pilot Implementation of GenAI-Supported Assessment' (2024) 40 Australasian Journal of Educational Technology 38.

[13] Heather Johnston and others, 'Student Perspectives on the Use of Generative Artificial Intelligence Technologies in Higher Education' (2024) 20 International Journal for Educational Integrity 2; Jan Henrik Gruenhagen and others, 'The Rapid Rise of Generative AI and Its Implications for Academic Integrity: Students' Perceptions and Use of Chatbots for Assistance with Assessments' (2024) 7 Computers and Education: Artificial Intelligence 100273.

2.4 Why policy expectations should remain separate

The remaining institutional items ask whether clearer rules would help, training would improve compliance, confidentiality restrictions would be followed, or consistent enforcement would encourage responsibility. These statements do not share the same temporal or behavioural status. Some evaluate an anticipated benefit; another expresses a conditional intention. They are not direct observations of current institutional practice. High endorsement may identify receptiveness to a proposed arrangement, but it cannot demonstrate that the arrangement exists or would produce the expected result.

Combining these statements with current awareness would obscure the research problem. A respondent can reasonably know little about existing rules while strongly supporting clearer guidance. Treating both answers as one latent policy-support score would remove precisely the distinction the study seeks to examine. Low internal consistency provides an additional empirical reason for retaining individual items, although internal consistency alone cannot determine a scale's theoretical validity.[14]

AI literacy presents a similar measurement issue. Long and Magerko describe competencies that include critical evaluation of AI systems. Awareness of inaccuracy and awareness of bias capture limited self-reported orientations within that wider field. They do not establish the ability to interpret a university regulation or assess the quality of legal research. Their adjustment role is consequently modest: they account for two recorded orientations, not comprehensive AI literacy.[15]

2.5 Association and predictive information

Institutional monitoring often assumes that an awareness survey can identify students who need intervention. That is a predictive claim, distinct from the claim that awareness and compliance are associated in observed data. Shmueli explains why explanatory and predictive objectives require separate attention. A model can produce an interpretable coefficient while making weak predictions in new observations, particularly when the signal is small or the available measures are noisy.[16]

The present analysis examines overall probability error through the Brier score and log loss, alongside discrimination through the area under the receiver operating characteristic curve. Proper scoring rules assess the quality of probabilistic predictions rather than rewarding a model simply for assigning the majority class. Steyerberg and colleagues explain the complementary roles of overall performance and discrimination. Here the purpose is a restrained check on

[14] Lee J Cronbach, 'Coefficient Alpha and the Internal Structure of Tests' (1951) 16 Psychometrika 297; Klaas Sijtsma, 'On the Use, the Misuse, and the Very Limited Usefulness of Cronbach's Alpha' (2009) 74 Psychometrika 107.

[15] Duri Long and Brian Magerko, 'What Is AI Literacy? Competencies and Design Considerations' in Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (ACM 2020).

[16] Shmueli (n 4).

incremental information, not development of a deployable screening system.[17]

The framework yields one primary hypothesis and three research questions. H1 anticipates a positive association between perceived rule-status awareness and permitted-use endorsement. RQ1 asks how awareness differs descriptively from support for prospective institutional arrangements. RQ2 asks whether the awareness association extends to independent-work boundaries and avoiding wholly personal attribution. RQ3 asks whether awareness improves held-out probability prediction beyond the recorded covariates and an intercept-only benchmark. All are exploratory because the analysis was developed after receipt of the database and was not preregistered.

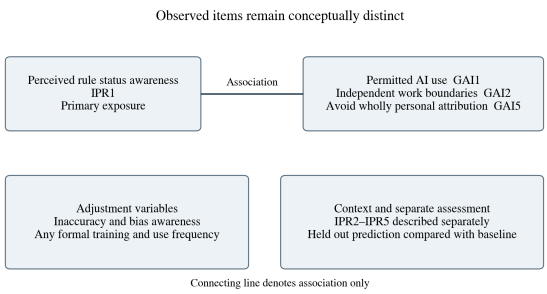

Figure 1. Analytical framework separating exposure, outcomes, contextual evaluations and prediction. The connecting line denotes association, not a causal effect.

Figure 1 summarises the analytical roles. The policy-expectation items provide context rather than mediators, and the secondary outcomes remain separate. The framework makes no claim that the available data identify a causal pathway from policy provision through awareness to integrity. That pathway requires observations of institutional guidance, its communication and subsequent conduct that are absent here.

3. Methodology

This section explains how the complete database (N = 385) is distinguished from the law-student analytical subsample (n = 245), how policy awareness and responsible-use responses are operationalised, and how association and predictive performance are evaluated. The student analysis addresses the legal-education question directly; the complete-sample analysis assesses whether its interpretation changes when graduates, legal professionals and other respondents are included. Data checks, uncertainty estimation and ethical safeguards are specified to make both denominators and the limits of inference transparent.

3.1 Design and data provenance

The study is a cross-sectional secondary analysis of the supplied Group-4-Law-SDG database, interpreted alongside the common questionnaire and Sifa Sultana's research-plan entry. The workbook contains 385 records and a codebook. The research plan describes a quantitative survey and a preferred probability-sampling approach. It does not establish that probability sampling was implemented. Collection dates, recruitment invitations, institution identifiers and a sampling frame are not documented in the materials available for this analysis.

The plan's sample target follows the conventional large-population calculation

[17] Tilmann Gneiting and Adrian E Raftery, 'Strictly Proper Scoring Rules, Prediction, and Estimation' (2007) 102 Journal of the American Statistical Association 359; Ewout W Steyerberg and others, 'Assessing the Performance of Prediction Models: A Framework for Traditional and Novel Measures' (2010) 21 Epidemiology 128.

with 95% confidence, 5% nominal precision and an assumed proportion of one half, yielding approximately 385. This calculation describes a planning target, not a guarantee of national representativeness or the precision of the student subgroup. The achieved inclusion probabilities and response rate cannot be reconstructed. No sampling weights are created.[18]

The codebook describes respondent identifiers as a synthetic sequence and optional comments as synthetic contextual comments. The numerical response-generation process is not documented. Consequently, the analysis reports properties of the supplied records, and its interpretation as evidence about human respondents remains conditional on provenance verification. Synthetic comments are excluded from all interpretation. STROBE-informed reporting is used to make design and evidentiary limitations explicit; it does not certify the original collection process.[19]

3.2 Analytical population and data checks

The complete database sample is N = 385: 245 current law students (63.6%), 71 law graduates (18.4%), 37 advocates or legal professionals (9.6%) and 32 other relevant respondents (8.3%). All 385 records are retained in the study. The primary student-specific analysis uses n = 245 because its outcomes concern course permissions and independent academic work; the other 140 records represent different current statuses, rather than missing or invalid student responses. The complete-sample sensitivity model uses all N = 385 records to assess the association across the broader respondent group. Nine of the 245 students report never using AI; a further sensitivity analysis excludes only these nine and therefore uses n = 236. No substantive-item missingness reduces any of these denominators.

The substantive items use a five-point agreement scale and have no missing values in the supplied records. All recorded eligibility and consent indicators are affirmative. Fourteen missing entries concern only the AI-tool field, which is not used in the models. Checks find no duplicate respondent identifiers, exact duplicate twenty-item substantive response vectors or records giving one identical response across all twenty items. These checks establish internal features of the file; they cannot verify recruitment, consent administration or the authenticity of the numeric data.

[18] William G Cochran, Sampling Techniques (3rd edn, John Wiley & Sons 1977).

[19] Erik von Elm and others, ‘The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for Reporting Observational Studies’ (2007) 4 PLOS Medicine e296.

Table 1. Measures and analytical roles

Item	Content	Analytical role
IPR1	Perceived knowledge of rule status	Primary exposure per score point
GAI1	Use only as permitted	Primary endorsement outcome
GAI2 / GAI5	Independent boundaries / personal attribution	Separate secondary outcomes
IPR2-IPR5	Prospective or conditional policy evaluations	Descriptive context only
AIL1 / AIL3	Inaccuracy / bias awareness	Separate score covariates
D8 / D7	Any formal training / use frequency	Binary / ordered covariates

Source: supplied Q4 questionnaire and codebook. Endorsement is a score of four or five. Item content does not establish validated latent constructs or observed misconduct.

Table 1 distinguishes the exposure, outcomes and contextual statements before modelling. None of the contextual policy items is treated as an observed institutional intervention. The covariates account for recorded orientations and experience, while the design leaves considerable unmeasured variation in courses, institutions and individual circumstances.

3.3 Measures and outcome definitions

IPR1 is the primary exposure, entered as an ordered score from one to five and centred at three. The main outcome is endorsement of GAI1, defined as agreement or strong agreement, scores four or five. Scores one through three are grouped as non-endorsement. This grouping includes neutrality and cannot be described as misconduct, rule violation or an admission of prohibited use. GAI2 and GAI5 use the same threshold as separate secondary outcomes. The exact wording is reproduced in Appendix A.

Four covariates are selected for their conceptual relevance and availability: awareness that AI can generate plausible inaccuracies, awareness of bias, any formal AI training and frequency of AI use. Training combines short orientation and course or certified training versus none. Frequency is ordered from never, coded zero, to daily or almost daily, coded four. The two awareness covariates are centred at three and frequency at two. AIL2 is excluded because its verification wording overlaps behavioural content. Demographic screening and stepwise variable selection are not used.

Within the student subsample (n = 245), Cronbach's alpha is calculated for the four five-item blocks as a diagnostic of the proposed scale structure. The values are low: 0.218 for AI literacy, 0.198 for integrity practice, 0.309 for responsibility and 0.162 for institutional policy. These heterogeneous statements are therefore retained at item level. Neither reflective structural equation

modelling nor an aggregate responsible-use score is fitted. Alpha is not interpreted as proof that each individual item is valid or invalid.[20]

3.4 Estimation and uncertainty

Unpenalised logistic regression estimates endorsement odds. The primary model includes IPR1 and the four covariates; an unadjusted awareness model provides a reference. Maximum likelihood is solved numerically, and HC1 sandwich covariance provides robust standard errors with the finite-sample multiplier n divided by n minus the number of coefficients. Reported intervals are 95% Wald intervals. Robust covariance addresses a limited variance assumption, not measurement error, unmeasured confounding or unknown clustering by institution.

The primary awareness test is reported separately from the two secondary outcome tests. Holm adjustment is applied to the awareness coefficients in the two secondary models. The family excludes the primary test and descriptive analyses, a distinction stated before interpreting the results. Effect estimates, intervals and substantive meaning receive priority over a binary significance label.[21]

Marginal standardisation translates the primary fitted association into mean predicted endorsement at IPR1 scores three and four, holding the distribution of the other covariates at the observed student distribution. A respondent bootstrap resamples 245 records with replacement, refits the model and standardises within each resample. Two thousand successful replicates provide pointwise percentile intervals; the random seed is 20261020. These are model-based adjusted comparisons, not intervention effects.[22]

Sensitivity analyses use strong agreement alone as the primary outcome, exclude never-users and include all 385 records. A quadratic awareness term checks the linear-score specification against a nested alternative through a one-degree-of-freedom likelihood-ratio comparison. Sparse extreme awareness categories preclude confident category-specific modelling. No interactions, mediation effects or institution-level effects are estimated.

3.5 Held-out assessment and ethics

Within the student subsample (n = 245), predictive assessment compares an intercept-only model, a model with the four covariates and that model plus IPR1. Twenty repetitions of stratified five-fold cross-validation use identical folds for all three models within each repetition. Every validation prediction is produced without fitting the model on that record. Each repetition yields pooled held-out Brier score, log loss and discrimination; reported summaries average the twenty repetitions. The intercept benchmark uses training-fold outcome prevalence. Ranges across repetitions describe partition sensitivity and are not confidence intervals or independent replications.

[20] Sijtsma (n 14).

[21] Sture Holm, ‘A Simple Sequentially Rejective Multiple Test Procedure’ (1979) 6 Scandinavian Journal of Statistics 65; Ronald L Wasserstein and Nicole A Lazar, ‘The ASA Statement on p-Values: Context, Process, and Purpose’ (2016) 70 American Statistician 129.

[22] Bradley Efron and Robert J Tibshirani, An Introduction to the Bootstrap (Chapman & Hall 1993).

The secondary analysis involves no new recruitment or contact. The supplied questionnaire specifies voluntary participation, anonymity, no collection of names and the ability to stop before completing the survey. No formal ethics approval document is supplied. The applicable review arrangements and implementation of the original consent process cannot be verified from the workbook. An absent approval document is not presented as an exemption. Data minimisation, exclusion of synthetic comments and aggregate reporting protect confidentiality in this analysis; no attempt is made to re-identify records or use the fitted models to classify students. Original collectors must confirm the applicable ethics pathway before empirical submission.

4. Results

Results are reported for the complete database (N = 385) and the law-student analytical subsample (n = 245), with each denominator stated explicitly. Student-specific descriptive, regression, bootstrap and cross-validation results use n = 245; the complete-sample sensitivity model uses all N = 385 records. These complementary analyses separate policy-response patterns, associations with specified practices and held-out predictive performance. Estimates describe recorded endorsement rather than verified conduct or national prevalence.

4.1 Student profile and policy response patterns

The complete database comprises N = 385 respondents, including the n = 245 current law students profiled in Table 2. Across all 385 records, policy awareness is endorsed by 233 respondents (60.5%), permitted use by 241 (62.6%), independent-work boundaries by 248 (64.4%) and avoiding wholly personal attribution by 256 (66.5%). Within the student subsample, 171 report weekly or daily AI use, nine report never using AI and 116 report some formal training. Table 2 describes this student subsample; its counts are not the distribution of the complete database or a representative national student population.

Table 2. Profile of law-student records (n = 245)

Profile	Category order	Counts in order
Age	18-20 / 21-23 / 24-26	37 / 102 / 57
Age	27-30 / 31+	24 / 25
Gender	Male / female / other / prefer not to say	131 / 110 / 2 / 2
Education	Undergraduate / graduate	139 / 62
Education	Postgraduate / professional	15 / 29
Location	Dhaka / other city corporation / district town	92 / 47 / 64
Location	Upazila-rural / other	36 / 6
Use frequency	Never / less than monthly / monthly	9 / 24 / 41
Use frequency	Weekly / daily or almost daily	85 / 86
Training	None / short orientation / course or certified	129 / 78 / 38

Denominator: n = 245 current law students drawn from the complete N = 385 database. The other 140 records comprise graduates, legal professionals and other relevant respondents and enter the complete-sample sensitivity analysis. Counts are unweighted;

no institution-level distribution is available.

The diversity of recorded statuses within the original workbook supports the decision to define a student-only primary sample. Within that sample, frequency and training vary sufficiently to enter the model as recorded covariates. The file lacks institution identifiers, so neither the profile nor the subsequent analysis can distinguish variation within universities from variation between them.

Policy awareness is endorsed by 151 students, or 61.6%. Endorsement is more common for the anticipated benefits of training and consistent enforcement: 84.5% and 86.9%, respectively. Table 3 presents these items separately because their different wording precludes treating the percentages as successive stages of a validated implementation scale.

Table 3. Distinct institutional policy responses among law students (n = 245)

Statement abbreviated	Mean	SD	Endorsement n (%)
IPR1 Rule-status awareness	3.71	0.82	151 (61.6%)
IPR2 Clearer rules would help	3.79	0.77	159 (64.9%)
IPR3 Training would improve compliance	4.21	0.71	207 (84.5%)
IPR4 Would follow confidentiality restrictions	3.85	0.74	172 (70.2%)
IPR5 Consistent enforcement would help	4.31	0.70	213 (86.9%)

All items range from one to five. IPR1 concerns perceived knowledge of rule status; IPR2-IPR5 concern anticipated or conditional arrangements. Their percentages are not a validated implementation sequence.

Among the 94 students who do not endorse current rule-status awareness, 81 endorse the training-benefit statement and 84 endorse the enforcement-benefit statement. The latter represents 89.4% of that subgroup. These responses show that limited awareness endorsement can coexist with receptiveness to prospective institutional arrangements. They do not establish dissatisfaction, exposure to inconsistent enforcement or the effectiveness of training. Rank correlations between IPR1 and IPR2-IPR5 range from approximately –0.045 to 0.043, reinforcing their weak empirical correspondence in this file.

4.2 Permitted use across awareness categories

Permitted-use endorsement is recorded for 150 students, or 61.2%. Independent-work boundary recognition is endorsed by 160, or 65.3%, and avoiding wholly personal attribution by 168, or 68.6%. These are separate marginal percentages; they do not indicate that the same students consistently endorse every practice.

The unadjusted permitted-use rate rises across grouped awareness responses. It is 46.7% among the fifteen students disagreeing with awareness, 55.7% among the 79 neutral students and 65.6% among the 151 endorsing awareness. Table 4 retains group counts so the uncertainty inherent in the small disagreement group remains visible.

Table 4. Permitted-use endorsement by grouped policy awareness (n = 245)

Awareness response	Group n	GAI1 endorsed n	Endorsed %
Disagreement (1-2)	15	7	46.7
Neutral (3)	79	44	55.7
Agreement (4-5)	151	99	65.6

Grouped categories are descriptive; the primary regression uses the original ordered awareness score. Non-endorsement includes neutrality and cannot be interpreted as misconduct.

The grouped pattern is descriptive. At the original five levels, only one student selects the lowest awareness response, making a detailed interpretation of that category uninformative. The rank association between the original awareness and permitted-use scores is weakly positive, Spearman's rho 0.110, with p = 0.087. The unadjusted logistic odds ratio is 1.36 per awareness point, with 95% interval 0.99-1.87. Both summaries are consistent with a modest positive association while leaving considerable uncertainty.

4.3 Adjusted association and standardised probabilities

After adjustment, the primary awareness odds ratio is 1.33 per agreement point, with 95% interval 0.95-1.86 and p = 0.096. Table 5 reports all non-intercept coefficients so that the exposure estimate can be assessed alongside the measured covariates. Each interval includes one. The results therefore do not establish a reliable independent association for awareness, training, use frequency or either recorded AI-awareness orientation.

Table 5. Adjusted model for permitted-use endorsement (n = 245)

Predictor	OR	95% CI	p
IPR1 policy-rule awareness	1.33	0.95-1.86	0.096
AIL1 inaccuracy awareness	1.20	0.86-1.69	0.285
AIL3 bias awareness	0.85	0.57-1.27	0.438
Any formal AI training	1.40	0.82-2.38	0.219
AI-use frequency step	1.13	0.89-1.43	0.331

Outcome: GAI1 score four or five. All predictors enter simultaneously. Item scores and frequency are per one-point or one-category increment; training compares any orientation/course with none. Intervals use HC1 sandwich covariance; intercept omitted.

The positive awareness estimate is compatible with higher permitted-use endorsement, but the interval also allows a small association in the opposite direction. It cannot establish a policy effect. Adjustment changes the point estimate only modestly from the unadjusted model, yet the absence of major attenuation does not show that confounding has been removed. Course expectations, access to guidance and conscientiousness remain unmeasured.

Standardisation yields a predicted endorsement probability of 56.5% at awareness score three and 63.3% at score four. The adjusted difference is 6.7 percentage points, with 95% bootstrap interval −0.9-14.7. All 2,000 bootstrap fits produce usable estimates. The difference offers an accessible scale for the same uncertain relationship; it does not transform the observational comparison into the effect of improving awareness. Score-specific intervals are pointwise and are not simultaneous coverage statements.

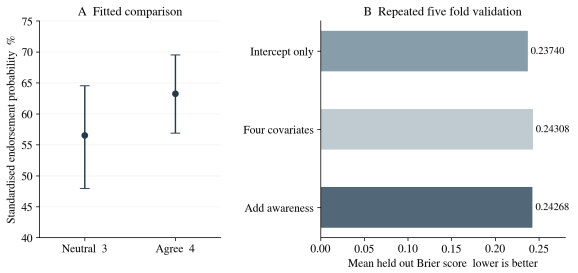

Figure 2. Student-subsample results (n = 245) for the adjusted probability comparison and held-out performance. Panel A shows pointwise 95% respondent-bootstrap intervals at awareness scores three and four; panel B shows mean cross-validated Brier scores, without inferential error bars.

Figure 2 places the adjusted probabilities beside the held-out prediction results. The two panels answer different questions. The left panel describes the fitted conditional comparison standardised over the observed covariates. The right panel assesses prediction in records held out from estimation. Neither identifies the consequences of introducing or changing institutional rules.

4.4 Secondary outcomes and sensitivity checks

The awareness estimate does not extend consistently across the secondary practices. For independent-work boundary endorsement, the adjusted odds ratio is 0.93, with 95% interval 0.67-1.29. For avoiding wholly personal attribution, it is 0.86, with interval 0.62-1.20. Both Holm-adjusted p-values are 0.742. These estimates do not support a uniform positive association between perceived rule-status awareness and a general responsible-use disposition.

Table 6. Awareness associations in secondary and sensitivity models

Outcome or sample	n	OR	95% CI	Raw p / Holm p
Independent-work boundaries	245	0.93	0.67-1.29	0.668 / 0.742
Avoid wholly personal attribution	245	0.86	0.62-1.20	0.371 / 0.742
Strong agreement only	245	1.27	0.86-1.88	0.223 / —
Current AI users	236	1.41	0.9994-1.9885	0.0504 / —
Complete database	385	1.17	0.89-1.53	0.262 / —

The first four rows use the student subsample: n = 245 for the two secondary outcomes and strict endorsement, and n = 236 for current users. The final row includes every record in the complete database (N = 385). All models adjust for the same four covariates; Holm correction applies only to the first two secondary outcomes. Extra precision is retained for the current-user estimate because its interval is close to one.

Within the student subsample, strict endorsement yields an awareness odds ratio of 1.27 (95% CI 0.86-1.88; n = 245). Excluding the nine never-users gives an odds ratio of 1.41 (95% CI 0.9994-1.9885; p = 0.0504; n = 236), retaining statistical uncertainty. Including every record produces an adjusted odds ratio of 1.17 (95% CI 0.89-1.53; p = 0.262; N = 385). Thus, the complete-database estimate and the student-specific estimate are both compatible with the absence of an association, while permitting a positive association of varying size. The complete-sample result concerns a broader, mixed-status group and cannot be presented as an estimate for 385 current law students.

The analyses use overlapping records and are not independent replications. The student quadratic-term check yields a likelihood-ratio statistic of 0.321 (one degree of freedom; p = 0.571), without establishing linearity.

4.5 Incremental predictive performance

Table 7 compares the three predictive models using the same held-out partitions. Lower Brier score and log loss indicate better probability predictions. The awareness-augmented model improves these metrics only slightly relative to the four-covariate model. Its mean discrimination is also modest, with area under the curve approximately 0.540.

Table 7. Mean held-out performance over twenty five-fold repetitions (n = 245)

Model	Brier score	Log loss	ROC area
Intercept only	0.23740	0.66773	0.500
Four recorded covariates	0.24308	0.68022	0.505
Covariates plus awareness	0.24268	0.67998	0.540

Lower Brier score and log loss are better; higher ROC area indicates stronger discrimination. Repetitions share the same records and do not provide independent replications. Benchmark probabilities are training-fold prevalence; model comparisons use identical folds.

The mean Brier change from adding awareness is –0.00040. Across twenty repetitions it ranges from –0.00269 to 0.00325, improving in thirteen repetitions and worsening in seven. These partition results are dependent and are not a significance test. More importantly, neither covariate model matches the intercept-only Brier score of 0.23740. Log loss gives the same ordering. Within the n = 245 student subsample, awareness alone has an apparent, in-sample rank-discrimination value of 0.564, which should not be confused with externally validated performance. The observed predictive information is insufficient to justify individual student screening.

5. Discussion

The discussion interprets the findings in relation to institutional communication, planned behaviour and academic integrity. It considers the primary law-student analysis (n = 245) alongside the complete-database sensitivity analysis (N = 385), then distinguishes an uncertain awareness-practice association from the limited predictive information available for individual monitoring. The implications concern task-specific understanding and the evaluation of guidance, with conclusions bounded by the composition, measurement and provenance of the source data.

5.1 What perceived awareness can establish

The central finding is a limited and uncertain relationship between perceived rule-status awareness and self-reported permitted use. The adjusted point estimate favours a positive association, but its interval includes the null and a small negative association. The standardised student difference likewise includes zero. The complete-database model (N = 385) yields an adjusted odds ratio of 1.17 (95% CI 0.89-1.53), so including all respondent statuses does not provide clear evidence of an awareness association. This broader estimate addresses a mixed-status group; it does not enlarge the number of current students. These findings neither demonstrate that awareness is irrelevant nor establish that raising it would improve conduct. They identify the magnitude and uncertainty of an association within the supplied student records. The distinction follows established cautions about interpreting non-significant results as evidence of no effect.[23]

The exposure's wording helps explain why a simple awareness model may provide limited information. Knowing whether institutional guidance exists is a comparatively thin form of policy knowledge. Students may know the answer while remaining uncertain about the details of an assignment. They may also follow course instructions without recognising them as part of an institutional policy. Neither possibility is directly measured, but both show why a stronger inference would require richer exposure measurement. Awareness, access, comprehension and application should be observed separately in future research.

This interpretation complements institutional document studies. An, Yu and James, Wang and colleagues, and McDonald and colleagues describe the content of university guidance and resources. The present analysis cannot evaluate whether similar guidance was supplied to these students. Its contribution is to demonstrate that an individual awareness response and a favourable view of policy arrangements need not behave like one coherent construct. An institutional document can be well designed while poorly communicated; a student can approve regulation while knowing little about the current rules. Those possibilities require evidence at both organisational and individual levels.[24]

The higher endorsement of prospective training and enforcement benefits should therefore be interpreted as receptiveness. In the recorded responses, even students without awareness endorsement frequently support those statements. This pattern may make educational communication worth evaluating, but it does not show unmet demand of a specified magnitude or prove that enforcement is currently inconsistent. The questionnaire asks respondents to imagine an improvement. Their predictions about its consequences are not observations of that improvement's effects.

[23] Douglas G Altman and J Martin Bland, 'Absence of Evidence Is Not Evidence of Absence' (1995) 311 BMJ 485; Wasserstein and Lazar (n 21).
[24] An, Yu and James (n 3); Wang and others (n 7); McDonald and others (n 3).

5.2 Behavioural specificity and the theoretical contribution

The secondary outcomes caution against assigning all integrity practices the same antecedents. Awareness has a positive point estimate for permitted use but slightly negative, uncertain estimates for the two secondary practices. These differences do not establish opposite mechanisms. Their practical meaning is that the data offer no clear basis for treating awareness as a common positive predictor across a general integrity domain. The small alpha for the practice block supports retaining that behavioural specificity.

Planned-behaviour reasoning clarifies why this matters. Following a permission rule, identifying an independent-work requirement and representing authorship accurately are related decisions with different information needs. A policy may explain one well and leave another implicit. The framework therefore contributes a specification of what must be measured before a fuller behavioural model is tested. It does not add an unmeasured institutional construct to the theory or claim that awareness mediates subjective norms.[25]

Academic integrity scholarship supplies the normative connection between those decisions. Honesty and responsibility require students to understand what they are being asked to demonstrate and to represent their contribution accurately. However, ethical commitments and administrative knowledge are not identical. The contextual integrity literature already indicates that conduct depends on more than a student's knowledge of rules. The present results fit that broader caution without identifying which omitted contextual factors account for the observed uncertainty.[26]

A future theoretical model could examine policy communication as an antecedent to task-specific comprehension, perceived legitimacy and control, followed by intention and observed practice. That model would require temporal ordering and direct measures of each stage. It should also permit the possibility that clearer permission increases legitimate AI use. A reduction in overall use is not inherently the desired outcome if the institutional objective is responsible educational assistance rather than avoidance of all technology.

5.3 The value and limits of prediction

The held-out comparison provides a distinctive check on the policy-awareness narrative. Adding IPR1 improves the covariate model's mean Brier score by approximately four ten-thousandths. Yet the resulting model remains worse than a prevalence-only benchmark. A modest improvement over a weak model cannot establish practical value. Reporting only the incremental change would omit the more informative comparison with the simplest available prediction rule.

The result also illustrates why coefficient interpretation and prediction should remain

[25] Ajzen (n 9); Beck and Ajzen (n 9).
[26] ICAI (n 10); McCabe, Treviño and Butterfield (n 11).

separate. An estimated odds ratio of 1.33 can coexist with little useful discrimination when responses overlap substantially and the sample is limited. That coexistence is not a contradiction. The coefficient describes a conditional association under the model; held-out scores assess the model's probability assignments when records are excluded from fitting. Neither establishes causal effectiveness, and weak prediction does not imply that a well-designed educational intervention would fail.[27]

No operational prediction system is proposed. The analysis does not define an action threshold, assess the consequences of false classification or validate across institutions. Unknown clustering also means that random record-level folds may share institutional environments between training and validation. External validation at the institution or cohort level could perform differently. The appropriate conclusion is that these recorded variables do not justify individual screening, rather than a broader judgement about what all future institutional data could predict.

5.4 Measurement and contextual boundaries

Self-report creates a further limit. Respondents may interpret permission differently, overlook assistance they regard as ordinary or answer in socially desirable ways. The outcome is therefore endorsement of a statement about permitted use. It is not a verified count of compliant submissions. Neutrality can reflect uncertainty rather than prohibited behaviour, and disagreement cannot be assigned a disciplinary meaning without task-specific evidence. Consistent terminology protects students as well as the validity of the study.

Research on AI detection reinforces the need for such care. Weber-Wulff and colleagues document limitations in detection tools, while Liang and colleagues identify bias against non-native English writing in examined detectors. These findings do not directly test the Group 4 responses. They do show why a weak awareness indicator should not be combined with an automated-text score and presented as evidence of misconduct. Educational support and evidentiary adjudication require different standards.[28]

The Bangladesh context warrants further investigation, but the present file cannot establish national prevalence, compare universities or identify differences across institutional policy regimes. Hasib and Islam provide a qualitative account of university engagement with ChatGPT that helps situate the topic. The present analysis offers a separate quantitative description conditional

[27] Shmueli (n 4); Steyerberg and others (n 17).

[28] Debora Weber-Wulff and others, 'Testing of Detection Tools for AI-Generated Text' (2023) 19 International Journal for Educational Integrity 26; Weixin Liang and others, 'GPT Detectors Are Biased against Non-Native English Writers' (2023) 4 Patterns 100779.

on the workbook's provenance. Its strongest contribution is a reproducible analytical distinction, with empirical generalisation dependent on authentic collection records and a defensible sampling account.[29]

6. Theoretical and Practical Implications

The implications follow from the separation of constructs and from the limited predictive information observed. They concern how institutions communicate expectations and how researchers evaluate that communication. They are proposals for development and testing rather than interventions shown effective by this survey.

6.1 Implications for institutional communication

Universities should distinguish making guidance available from establishing that students can apply it. A practical communication process would link the university's general position to the instructions for each assessment. Students need to identify which forms of assistance are permitted, which intellectual tasks must remain independent and whom to consult when the boundary is unclear. Chan's policy framework and guidance from UNESCO provide relevant educational starting points, while assessment-specific frameworks illustrate how permissions can be differentiated.[30]

For law assignments, examples can make these boundaries concrete. An instruction might allow language editing while requiring the student to construct the legal argument independently. Another assignment might permit a generated outline but require an explanation of how authorities were selected and checked. The institution should choose the relevant rule and communicate it before submission. The examples are proposed applications, not descriptions of policies found in the database.

Comprehension should be evaluated through scenarios rather than familiarity alone. A student could identify the existence of guidance yet be unable to explain whether a particular use is permitted. Short scenarios can ask the student to apply the actual course instruction and explain the decision. Such an approach would generate a more direct indicator of actionable understanding than the current awareness item. It should be accessible to students with different levels of English proficiency and prior experience.

6.2 Fairness and implementation

High support for training and consistent enforcement offers a reason to investigate acceptability through student consultation. It should not be treated as consent to intrusive monitoring or a mandate for punitive measures. Institutions can invite students to identify ambiguous examples, test whether instructions are understood similarly and revise guidance when an interpretation is unclear. Moore and Lookadoo's communication perspective is relevant to that process.[31]

[29] Hasib and Islam, 'How University Students' (n 5).
[30] Chan (n 3); Fengchun Miao and Wayne Holmes, Guidance for Generative AI in Education and Research (UNESCO 2023); Perkins and others, 'AI Assessment Scale' (n 12).
[31] Moore and Lookadoo (n 6).

The weak predictive results argue for offering educational support broadly rather than targeting students through these fitted scores. If a university wants to identify where communication fails, it should examine the clarity of tasks and the availability of advice as well as students' answers. Any allegation about a submitted assignment requires evidence tied to the applicable instruction, an opportunity for explanation and a proportionate decision process. A survey response is not a substitute for that evidence.

6.3 Implications for theory and evaluation

The theoretical contribution is a boundary around the awareness construct. Perceived knowledge of rule status can be studied as an informational condition without being relabelled attitude, control or integrity. Future planned-behaviour research should measure those concepts separately and specify the relevant assignment and time period. Academic integrity values can guide the normative interpretation, while behavioural evidence should establish what students actually do.

Prospective institutional evaluation should collect the policy text, dates of communication, evidence of exposure, scenario-based comprehension and subsequent task-specific practice. A comparison of communication formats could then assess whether clearer examples improve understanding and compliance. Random allocation, where feasible, or a credible longitudinal comparison would address intervention questions more directly than cross-sectional endorsement. Predictive evaluation should retain a prevalence benchmark and validation across cohorts or institutions, reporting uncertainty and potential consequences before proposing any operational use. The present findings justify that stronger design; they do not supply its missing evidence.

7. Conclusion and Future Research

Perceived institutional AI policy awareness and responsible-use endorsement are related questions, but the supplied records do not support treating them as interchangeable. The complete database contains N = 385 records, of which n = 245 are current law students. Awareness and permitted-use endorsement each occur in approximately three fifths of both the complete database and the student subsample. The complete-database awareness association is also uncertain (adjusted OR 1.17; 95% CI 0.89-1.53). The adjusted awareness association is positive and uncertain, and its standardised probability comparison includes zero. The secondary practices show no consistent positive awareness association. High endorsement of prospective training and enforcement benefits concerns anticipated institutional arrangements rather than verified implementation.

The held-out assessment further limits practical interpretation. Awareness contributes little incremental probability information beyond the measured covariates, and neither fitted covariate model outperforms a prevalence-only benchmark. These results do not support screening students for compliance through the available awareness response. They also do not determine whether an intervention that

improves task-specific understanding would change behaviour. That causal question remains open.

Several limitations constrain the article. The numeric data's provenance is unresolved because the codebook uses synthetic labels for identifiers and contextual comments without documenting how substantive responses were generated. Empirical conclusions require confirmation of authentic collection records. Recruitment, dates, response rate and sampling probabilities are undocumented. The complete database reaches 385 records, but only 245 identify as current law students; the overall planning target does not establish precision or representativeness for that subsample. National representativeness cannot be claimed. Institutional identifiers are absent, preventing multilevel analysis and adjustment for common institutional environments.

Measurement limitations are equally material. Single items capture narrow self-reported orientations; agreement does not verify policy knowledge or actual compliance. The principal exposure does not establish whether a policy exists, and the outcome combines course and institutional permission. Dichotomisation improves interpretability but groups neutrality with disagreement. Strong-agreement sensitivity does not remove that measurement problem. Low block consistency precludes a validated latent model, while the available covariates leave unmeasured differences in assignments, resources and student circumstances.

Future studies should begin by documenting their ethics pathway and data provenance, then recruit an explicitly defined student population across identifiable institutions under appropriate confidentiality protections. Measures should separate policy availability, communication, comprehension, legitimacy and task-specific intention. Subsequent assessments can supply evidence of actual permitted practice, with consent and proportionate data handling. A longitudinal or experimental evaluation can examine changes following clearer guidance; external validation can assess whether useful predictive information transfers across cohorts. The present article supplies the analytical distinctions needed for that programme, with all empirical interpretation conditional on the source records being verified.

Declarations

Ethics Approval

Formal ethical approval was not obtained. At the time of the study, the authors report that their institutional setting did not provide an accessible formal review mechanism applicable to this non-interventional adult questionnaire research. This circumstance is not represented as a formal exemption or independent ethics clearance. Participant-protection provisions included informed consent, voluntary participation, the option to discontinue, an anonymous questionnaire and aggregate reporting, as described in Section 3.6. Responsibility for ethical conduct remained with the researchers.

Informed Consent

The questionnaire includes a consent screen stating that participation is voluntary, that no name is required, that responses will be analysed in aggregate, and that participants may stop at any time.

Data Availability

The analysed Group 7 database is held by the research team. Access is subject to participant-consent conditions, confidentiality and applicable institutional requirements. Aggregate findings and analytical definitions are reported in this article and its appendices.

Conflict of Interest

No competing interests are reported in the study documentation.

Funding

No external funding is identified in the available study documentation.

Author Contributions

Sifa Sultana: Conceptualisation; preparation of study materials; data organisation and curation; writing-original draft. The project was conducted under the supervision of Prof. Dr Kazi Abdul Mannan.

Kazi Abdul Mannan: Supervision; conceptual and methodological guidance; interpretation of findings; writing-review and editing.

AI/Generative AI Disclosure

Generative AI assisted manuscript drafting, language editing and analytical support. The authors retain responsibility for the accuracy of the data analysis, references, legal interpretation and conclusions.

Bibliography

Ajzen I, ‘The Theory of Planned Behavior’ (1991) 50 Organizational Behavior and Human Decision Processes 179

Altman DG and Bland JM, ‘Absence of Evidence Is Not Evidence of Absence’ (1995) 311 BMJ 485

An Y, Yu JH and James S, ‘Investigating the Higher Education Institutions’ Guidelines and Policies Regarding the Use of Generative AI in Teaching, Learning, Research, and Administration’ (2025) 22 International Journal of Educational Technology in Higher Education 10

Beck L and Ajzen I, ‘Predicting Dishonest Actions Using the Theory of Planned Behavior’ (1991) 25 Journal of Research in Personality 285

Chan CKY, ‘A Comprehensive AI Policy Education Framework for University Teaching and Learning’ (2023) 20 International Journal of Educational Technology in Higher Education 38

Cochran WG, *Sampling Techniques* (3rd edn, John Wiley & Sons 1977)

Cotton DRE, Cotton PA and Shipway JR, ‘Chatting and Cheating: Ensuring Academic Integrity in the Era of ChatGPT’ (2024) 61 Innovations in Education and Teaching International 228

Cronbach LJ, ‘Coefficient Alpha and the Internal Structure of Tests’ (1951) 16 Psychometrika 297

Dahl M and others, ‘Large Legal Fictions: Profiling Legal Hallucinations in Large Language Models’ (2024) 16 Journal of

Legal Analysis 64

Efron B and Tibshirani RJ, *An Introduction to the Bootstrap* (Chapman & Hall 1993)

Foltynek T and others, 'ENAI Recommendations on the Ethical Use of Artificial Intelligence in Education' (2023) 19 International Journal for Educational Integrity 12

Furze L and others, 'The AI Assessment Scale (AIAS) in Action: A Pilot Implementation of GenAI-Supported Assessment' (2024) 40 Australasian Journal of Educational Technology 38

Gneiting T and Raftery AE, 'Strictly Proper Scoring Rules, Prediction, and Estimation' (2007) 102 Journal of the American Statistical Association 359

Gruenhagen JH and others, 'The Rapid Rise of Generative AI and Its Implications for Academic Integrity: Students' Perceptions and Use of Chatbots for Assistance with Assessments' (2024) 7 Computers and Education: Artificial Intelligence 100273

Hasib M and Islam MS, 'How University Students in Bangladesh Engage with ChatGPT: A Qualitative Study' (2025) 20 PLOS ONE e0333089

Holm S, 'A Simple Sequentially Rejective Multiple Test Procedure' (1979) 6 Scandinavian Journal of Statistics 65

International Center for Academic Integrity, *The Fundamental Values of Academic Integrity* (3rd edn, ICAI 2021)

Jiang Y, Xie L and Cao X, 'Exploring the Effectiveness of Institutional Policies and Regulations for Generative AI Usage in Higher Education' (2025) 79 Higher Education Quarterly e70054

Johnston H and others, 'Student Perspectives on the Use of Generative Artificial Intelligence Technologies in Higher Education' (2024) 20 International Journal for Educational Integrity 2

Kasneci E and others, 'ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education' (2023) 103 Learning and Individual Differences 102274

Liang W and others, 'GPT Detectors Are Biased against Non-Native English Writers' (2023) 4 Patterns 100779

Long D and Magerko B, 'What Is AI Literacy? Competencies and Design Considerations' in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (ACM 2020)

McCabe DL and Treviño LK, 'Academic Dishonesty: Honor Codes and Other Contextual Influences' (1993) 64 Journal of Higher Education 522

McCabe DL, Treviño LK and Butterfield KD, 'Cheating in Academic Institutions: A Decade of Research' (2001) 11 Ethics & Behavior 219

McDonald N and others, 'Generative Artificial Intelligence in Higher Education: Evidence from an Analysis of Institutional Policies and Guidelines' (2025) 3 Computers in Human Behavior: Artificial Humans 100121

Miao F and Holmes W, *Guidance for Generative AI in Education and Research* (UNESCO 2023)

Moore S and Lookadoo K, 'Communicating Clear Guidance: Advice for Generative AI

Policy Development in Higher Education' (2024) 87 Business and Professional Communication Quarterly 610

Perkins M, 'Academic Integrity Considerations of AI Large Language Models in the Post-Pandemic Era: ChatGPT and Beyond' (2023) 20(2) Journal of University Teaching and Learning Practice, DOI: 10.53761/1.20.02.07

Perkins M and others, 'The Artificial Intelligence Assessment Scale (AIAS): A Framework for Ethical Integration of Generative AI in Educational Assessment' (2024) 21(6) Journal of University Teaching and Learning Practice, DOI: 10.53761/q3azde36

Shmueli G, 'To Explain or to Predict?' (2010) 25 Statistical Science 289

Sijtsma K, 'On the Use, the Misuse, and the Very Limited Usefulness of Cronbach's Alpha' (2009) 74 Psychometrika 107

Steyerberg EW and others, 'Assessing the Performance of Prediction Models: A Framework for Traditional and Novel Measures' (2010) 21 Epidemiology 128

von Elm E and others, 'The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for Reporting Observational Studies' (2007) 4 PLOS Medicine e296

Wang H and others, 'Generative AI in Higher Education: Seeing ChatGPT through Universities' Policies, Resources, and Guidelines' (2024) 7 Computers and Education: Artificial Intelligence 100326

Wasserstein RL and Lazar NA, 'The ASA Statement on p-Values: Context, Process, and Purpose' (2016) 70 American Statistician 129

Weber-Wulff D and others, 'Testing of Detection Tools for AI-Generated Text' (2023) 19 International Journal for Educational Integrity 26

Whitley BE Jr, 'Factors Associated with Cheating among College Students: A Review' (1998) 39 Research in Higher Education 235

Appendix A Focal Questionnaire Items

Table A1 preserves the original statements so that the scope of each measure can be assessed independently of the abbreviated labels used in the results.

Table A1. Exact wording of focal substantive items

Code	Questionnaire statement
IPR1	I know whether my institution has rules or guidance on generative AI use.
IPR2	Clear institutional AI rules help me distinguish acceptable from unacceptable use.
IPR3	Training on responsible AI would improve my compliance with academic integrity requirements.
IPR4	I would follow institutional restrictions on entering confidential or personal data into AI tools.
IPR5	Consistent enforcement of AI rules would encourage responsible student behaviour.
GAI1	I use generative AI only in ways permitted by my course or institution.
GAI2	I distinguish between AI assistance and work that I must complete independently.
GAI5	I avoid submitting substantially AI-generated work as entirely my own.
AIL1	I understand that generative AI can produce plausible but inaccurate information.
AIL3	I understand that AI outputs may reflect bias or incomplete data.

Common response scale: one strongly disagree, two disagree, three neutral, four agree and five strongly agree. Training and

frequency categories are reported in Table 2.

The wording explains why awareness, anticipated benefits and reported practices remain separate. The statements do not constitute an objective test of policy comprehension or an audit of submitted legal work.

Appendix B Reproducibility Specification

The source is the Data sheet of Group-4-Law-SDG-Database(1).xlsx, containing N = 385 records. For the primary analysis, select D3_Current_Status equal to Law student, retaining n = 245. Retain all N = 385 records for the complete-sample sensitivity model; exclude only the nine student never-users for the n = 236 current-user sensitivity model. Define endorsement as GAI1, GAI2 or GAI5 greater than or equal to four. The design matrix contains an intercept, IPR1 minus three, AIL1 minus three, AIL3 minus three, any D8 training versus No, and D7 frequency coded zero through four minus two. Fit each outcome separately by unpenalised logistic maximum likelihood. No weights, imputation, feature selection or interactions are used.

Let p denote the fitted probability and X the design matrix. The covariance is the inverse observed information multiplied on both sides of the sum of respondent score outer products, with HC1 multiplier n/(n–k). Confidence limits exponentiate coefficient plus or minus 1.96 robust standard errors. Holm adjustment applies only to the two secondary awareness tests. The quadratic check adds (IPR1–3) squared and compares twice the log-likelihood improvement with a one-degree-of-freedom chi-squared reference.

For standardisation, set the awareness column to zero and then one, corresponding to original scores three and four, and average predicted probabilities over the student covariates. In each of 2,000 respondent resamples, refit and repeat standardisation using that resample’s covariate distribution. Use seed 20261020 and the 2.5th and 97.5th percentiles for pointwise intervals. The probability difference is calculated within each replicate before obtaining its interval.

For prediction, use twenty shuffled stratified five-fold partitions with seeds 20261020 through 20261039. Fit the same unpenalised logistic models on each training fold. The four-covariate model excludes IPR1; the augmented model includes it. The intercept benchmark predicts the training-fold endorsement proportion. Compute pooled held-out Brier score as the average squared difference between probability and outcome, log loss as the mean negative Bernoulli log likelihood, and receiver operating characteristic area from those same probabilities. Average each metric across repetitions. Differences use matched repetitions; their range is partition sensitivity, not an inferential interval. Python calculations use pandas, NumPy, SciPy and scikit-learn; document formatting is independent of estimation.