RedFame

Journal of Education and Training Studies
Vol. 15, No. 1; January 2027
ISSN 2324-805X E-ISSN 2324-8068
Published by Redfame Publishing
URL: http://jets.redfame.com

When AI Writes the Text: AI-Assistance Depth and Recorded Disclosure in University Application Essays

Karan Gupta

Correspondence: Karan Gupta, Karan Gupta Consulting, Mumbai, India.

Received: August 10, 2026 Accepted: September 15, 2026 Online Published: September 16, 2026

doi:10.11114/jets.v15i1.9173 URL: https://doi.org/10.11114/jets.v15i1.9173

Abstract

Generative artificial intelligence has made assistance with university application essays easier to obtain, but disclosure rules often treat all use as a single category. This study examined whether the absence of AI disclosure varied with the functional depth of assistance. Of 684 applicants recruited through 10 counselling organisations, 617 met the analytic criteria. AI involvement was classified as none, editing only, idea generation or substantive generation using process logs, draft histories and student and counsellor reports, and disclosure was coded from the final application materials. The primary analysis covered the 497 applicants with recorded AI involvement, clustered across 58 counsellors. No disclosure was recorded for 39.9% of editing-only cases, 44.4% of idea-generation cases and 60.4% of substantive-generation cases. After adjustment, idea generation did not differ detectably from editing only, whereas substantive generation carried more than twice the odds of no recorded disclosure relative to editing only (OR = 2.32, 95% CI [1.49, 3.62]) and almost twice the odds relative to idea generation (OR = 1.94, 95% CI [1.31, 2.87]). Intermediary control was not associated with AI disclosure, and no association was found with authorial displacement, ownership, explainability or self-attributed authorship. The pattern is consistent with a functional threshold rather than a gradient. Universities should define material assistance by function and record which disclosure route each applicant was offered.

Keywords: generative artificial intelligence; disclosure; academic ethics; university admissions; authorship; personal statements

1. Introduction

University application essays occupy an unusual ethical position. They are submitted as evidence of motivation, judgement, character, fit and personal history, yet they are rarely produced in isolation. Applicants discuss topics with parents, take advice from teachers and counsellors, use commercial editing services, and increasingly rely on generative artificial intelligence (GenAI) for brainstorming, revision or direct text production, yet the document is still submitted under one name. Research on selective admissions has long shown that application writing is shaped by social advantage, school context and professional coaching (Gebre-Medhin et al., 2022; T. J. Huang, 2024; Jones, 2013); GenAI adds a scalable and less visible contributor (Kasneci et al., 2023).

Academic-integrity responses initially concentrated on detection, which is weak at the individual level. Detection systems vary across models, prompts, languages and post-editing conditions, and can misclassify human writing, particularly by people who use English as an additional language (Liang et al., 2023; Perkins et al., 2024; Weber-Wulff et al., 2023). Human readers are also unreliable (Waltzer et al., 2024): selection committees could not reliably separate genuine personal statements from generated ones (Lum et al., 2024), while machine assessment of personal qualities is itself feasible (Lira et al., 2023). These limits have encouraged a shift from provenance toward transparency (Dawson et al., 2024; Eaton, 2023; Foltýnek et al., 2023).

Disclosure appears to offer a less punitive alternative: rather than asking institutions to infer hidden tool use from linguistic traces, it asks the applicant to state whether AI contributed and, ideally, how. In practice disclosure operates under uncertainty, because students may not know what counts as reportable assistance, where it should appear, or whether reporting will disadvantage them. Kirsanov et al. (2025) found AI disclosure rare among 31 university students and read silence as shaped by ambiguous rules and fear of penalty rather than dishonesty; Yusuf et al. (2025) found non-disclosure associated with psychological factors rather than intention to deceive.

The ethical problem is therefore whether disclosure rules remain effective precisely when assistance becomes most consequential. A regime that elicits acknowledgement of low-level editing but loses visibility when AI writes substantive passages fails where information matters most. Functional differentiation is accordingly central to current guidance on educational AI use (Foltýnek et al., 2023; UNESCO, 2023).

A second unresolved issue is whether disclosure can stand in for authorship integrity. It is tempting to assume that applicants

who disclose nothing are those whose essays have moved furthest from their own priorities, who feel least ownership, or who cannot explain the final text, but that assumption has not been tested. Disclosure is a communicative act governed by policy, risk and interpretation; displacement, ownership and explainability describe the applicant's relation to the text itself.

This study examines AI disclosure in a multi-source dataset of university applicants, using process evidence rather than detection or self-report. It asks whether disclosure changes gradually across assistance levels or at a functional boundary, whether counsellor control predicts it independently of AI depth, and whether it is associated with the integrity outcomes measured in a companion analysis.

2. Conceptual Framework and Research Questions

2.1 Disclosure as Transparency Rather Than Confession

Academic-integrity systems often inherit a misconduct model in which disclosure is treated as an admission. That framing suits GenAI poorly, because many forms of use are permitted, encouraged or necessary for accessibility. Ethical disclosure is better understood as information enabling a legitimate evaluator to interpret an artefact: the question is not whether a tool was present but what function it performed. This shifts responsibility toward the institution that defines the reporting rule (International Center for Academic Integrity, 2021; Whitley & Keith-Spiegel, 2001; Birks & Clare, 2023).

Transparency requires a usable disclosure architecture: applicants need to know what assistance is material, where to report it and how it will be interpreted, and institutions need categories specific enough to support judgement without demanding a history of every spelling correction. Research on scholarly AI disclosure reaches a similar conclusion — generic statements establish presence but reveal little about function (Chen & Jia, 2026; Lund et al., 2023; Ruttenberg, 2026; Bjelobaba et al., 2025).

Proportionality follows. If disclosure exists to support an inference, the detail required should be calibrated to the inference at risk. An admissions office drawing conclusions about motivation needs to know whether that account is the applicant's own; it does not need to know whether a grammar checker was run over the final draft.

2.2 A Functional Threshold in AI Assistance

AI assistance can be arranged along a functional continuum. Editing modifies expression without necessarily altering the underlying claim; idea generation proposes themes, structures or examples while the applicant composes the substantive content; substantive generation supplies sentences or paragraphs incorporated into the final essay. The categories are not perfectly discrete but represent ethically meaningful differences in control (Cotton et al., 2024; Foltýnek et al., 2023). Ucan and Demirel Ucan (2026) formalise this into a six-level taxonomy graded by how far the tool contributes to substance rather than surface.

Two rival patterns are possible, and they differ in where they locate the ethically relevant boundary. A gradient account treats AI contribution as a matter of degree: each increase in assistance makes disclosure somewhat less likely, so non-disclosure should rise steadily from editing through idea generation to substantive generation, with no step larger than the others. A threshold account treats the boundary as functional: what changes the applicant's relation to the text is not how much help was received but whether the tool supplied language the applicant then submitted as their own. It therefore predicts little difference between editing and idea generation, both of which leave composition with the applicant, followed by a break at substantive generation. The two accounts are distinguished by a single comparison. A gradient predicts that idea generation sits between the other two; a threshold predicts that substantive generation exceeds idea generation specifically, while editing and idea generation do not reliably differ. Showing only that the deepest category differs from the shallowest is consistent with both and settles neither. The threshold account matters because it identifies the boundary policy must define: not any contact with AI, but material contribution to the content of the application.

2.3 AI Disclosure and Human Intermediary Control

Application essays are also shaped by human intermediaries. Private counsellors, school counsellors and university-commissioned agents may give feedback, collaborate on revision, or determine structure and wording. Counselling becomes ethically concerning when the intermediary substitutes judgement or text while the application still functions as an unaided voice, so function and control are more informative than occupational title or payment status (Bretag et al., 2019; I. Y. Huang et al., 2022; Lancaster, 2020; Nikula & Kivistö, 2020).

Many emerging policies ask explicitly about GenAI while leaving comparable human assistance unrecorded. The dataset contains a measure of intermediary control but no parallel intermediary-disclosure outcome, so the study can test whether control predicts the AI-disclosure decision but cannot estimate disclosure of human assistance. That gap is itself a governance asymmetry.

2.4 Admissions as a Distinctive Disclosure Context

Most empirical work on AI disclosure concerns coursework and assessment, where an established teaching relationship supplies context that admissions lacks. Three features distinguish the application essay. First, the evaluator cannot ask: a tutor who doubts an essay can set an oral examination, whereas an admissions officer reads one document from a stranger,

so disclosure is often the only structured information about production available.

Second, the stakes are concentrated: an admission decision is singular and not recoverable within the cycle, so an applicant weighs a place against candour rather than a marginal grade. Third, the genre is already collaborative — schools run personal-statement workshops, teachers comment on drafts, and paid consultants are established in the international market (T. J. Huang, 2024; Jones, 2013) — so generative AI enters a genre whose disclosure norms were already weak.

2.5 Research Questions

RQ1. Does the probability that no AI disclosure is recorded vary across editing-only, idea-generation, and substantive-generation assistance, and if so, does it change gradually or at a discontinuity?

RQ2. Does intermediary control predict AI non-disclosure independently of the depth of AI involvement, and is there evidence that the two forms of assistance interact?

RQ3. Which applicant characteristics are associated with no recorded AI disclosure?

RQ4. Is AI disclosure associated with authorial displacement, perceived ownership, explainability, or self-attributed authorship after accounting for assistance depth?

3. Method

3.1 Design, Setting, and Participants

The study used a prospective, multi-source observational design embedded in one international university-application cycle. Applicants were recruited through 10 counselling organisations, generating 684 screened cases of which 617 entered the analytic sample. Sixty-seven were excluded: baseline narrative not completed (29), final narrative unavailable (18), orientation profile incomplete (12), process log unusable (8). Cases were nested within 59 counsellors; the 497 in the primary model were handled by 58. Data were collected between 18 June and 30 August 2025. By the end of collection, 438 of the 617 applications had been submitted and 179 finalised but not yet transmitted; case-level submission status was not retained and therefore could not be modelled.

Mean age was 19.8 years (SD = 3.0), with 395 undergraduate and 222 postgraduate applicants. Intended destinations were the United States (n = 181), United Kingdom (n = 163), continental Europe (n = 97), Canada (n = 79), Australia (n = 55) and Singapore or Hong Kong (n = 42). One hundred and forty-seven were first-generation students and 176 reported English as a first language, reflecting the proportion educated in English-medium international curricula. Full characteristics are in Table 1.

Students completed a supervised baseline narrative before substantive assistance. Drafts, tracked-change documents and version histories were retained through the normal organisational workflow, and participants completed structured process logs identifying who suggested, drafted, revised or approved substantive sections. AI interaction records were included only when voluntarily supplied. After completion, students rated ownership and sat a structured interview testing whether they could explain the origin and selection of central claims.

3.2 Ethics and Consent

The study was approved by an independent ethics review board prior to recruitment. Full identifying details of the board, its registration, its assurance and the protocol number are provided; the consent arrangements for applicants, parents and counsellors are set out in the Declarations.

3.3 Measures

AI involvement

AI involvement was classified into four mutually exclusive levels, drawing on student and counsellor reports, process logs, draft histories, tracked changes and voluntarily supplied AI records. Editing only covered grammar, concision or sentence-level clarity; idea generation covered proposed themes, questions, examples or structures where the student wrote the substantive text; substantive generation covered AI-generated sentences or paragraphs materially incorporated into the final submission. Where accounts diverged, documentary evidence was weighted more heavily and disagreements were resolved by discussion; no formal agreement statistic was computed. Cases coded none had a mean estimated AI text share of 1.07% (maximum 5.8%), against 10.5% for editing only, 24.7% for idea generation and 53.7% for substantive generation.

AI disclosure

AI disclosure was coded from the final version of the application materials held in the counselling record rather than from retrospective self-report: either the version already submitted or the one finalised for intended submission. The receiving institutions either provided a dedicated AI-disclosure field or permitted an additional-information statement, so a route to disclose existed in every case. The two are ethically distinct and were not recorded per case: a dedicated field requests

information; an additional-information box only allows it to be volunteered. The analysis therefore cannot separate an applicant declining an explicit request from one not volunteering information never solicited, and no claim about non-compliance is made. Four states were coded: no disclosure recorded, where none appeared despite documented AI involvement; general, acknowledging AI without specifying function or extent; detailed, identifying the tool and what it did; and not applicable, where no AI involvement was documented.

For the principal analysis, applicants with recorded AI involvement were classified according to whether no AI disclosure was recorded or whether a general or detailed disclosure was present. Cases without recorded AI involvement were retained for descriptive checks but excluded from the principal logistic model. The specificity analysis compared detailed with general disclosure among applicants for whom some disclosure was recorded. Disclosure was coded once per case, so no inter-rater reliability statistic is available for this variable.

Estimated AI text share

Two researchers compared the final narrative against draft histories, tracked changes and available AI records, and coded each sentence as student-generated, AI-generated, jointly produced or indeterminate. Estimated AI text share is the proportion of the final word count in sentences classified as AI-generated or substantially derived from AI-generated text. It is defined only where the applicant voluntarily supplied AI records and is available for 448 of the 497 cases (90.1%); availability did not differ by disclosure state ($p = .710$) or AI depth ($p = .192$).

Intermediary control

Intermediary control was coded independently of AI involvement. Feedback only involved questions or comments without substantive rewriting; collaborative revision involved joint selection of examples or reworking of substantial sections; intermediary-directed drafting indicated that the intermediary substantially determined structure, emphasis, examples or wording. Documentary evidence was weighted above retrospective self-report where accounts differed.

Companion integrity constructs

Authorial displacement was the normalised Euclidean distance between independently coded baseline and final narrative profiles across Security, Status and Freedom orientations. Ownership was rated on a seven-point scale and explainability scored 0 to 10 through a structured interview by assessors blinded to the displacement score. Applicants also named the principal author.

Covariates

The adjusted model included first-generation status, application level, gender, household-income band and English-first-language status, with AI depth and intermediary control entered simultaneously. These were selected before interpreting the results.

3.4 Analytic Strategy

The primary analysis was restricted to the 497 applicants with editing-only, idea-generation or substantive-generation AI involvement. A binary logistic regression estimated the odds that no disclosure was recorded, using editing only and no intermediary involvement as reference categories. Standard errors were clustered by counsellor, yielding 58 clusters and F-based joint tests. Counsellor-level variance was negligible, so clustered estimates differ little from unclustered ones. Because the threshold account requires a discontinuity rather than a difference from the lowest category, substantive generation was also contrasted directly with idea generation using a linear combination of the fitted coefficients.

The AI-depth-by-intermediary-control interaction was assessed with a cluster-robust joint F test on the product terms. Among applicants who made any disclosure, an exploratory model compared detailed with general disclosure, and crude and within-depth comparisons of estimated AI text share distinguished compositional confounding from within-category concealment. The primary model was refitted excluding editing-only cases, on the ground that sentence-level correction may not reasonably attract a disclosure expectation. Association analyses regressed authorial displacement, ownership and explainability on non-disclosure, adjusting for AI depth, intermediary control and the prespecified covariates; self-attributed authorship was assessed using the phi coefficient. Two one-sided tests for equivalence were conducted at three margins. Benjamini-Hochberg false discovery rate correction was applied across the inferential tests reported in Tables 3 and 4, with adjusted q values shown there.

The same cohort is examined in a companion manuscript focused on authorial displacement. To avoid duplicate reporting, disclosure is the sole primary outcome here; displacement, ownership, explainability and authorship attribution are used only to test whether disclosure represents a distinct construct. All analyses are associational.

4. Results

4.1 Sample and Disclosure Structure

AI involvement was recorded in 497 of the 617 analytic cases: editing only in 173, idea generation in 180 and substantive generation in 144. The remaining 120 had no recorded AI involvement.

Table 1. Characteristics of the analytic sample (N = 617)

Characteristic	n	Value
Age, years	617	19.8 (3.0)
Undergraduate	395	64.0%
Postgraduate	222	36.0%
Woman	299	48.5%
Man	298	48.3%
Non-binary / prefer not to say	20	3.2%
First-generation student	147	23.8%
English first language	176	28.5%
United States destination	181	29.3%
United Kingdom destination	163	26.4%
Other destination	273	44.2%

Note. Percentages use N = 617. Age is reported as mean (SD).

The disclosure states are separated almost perfectly by recorded AI involvement (Table 2). All 236 cases with no disclosure recorded occurred among applicants with recorded AI involvement and none among applicants without it; the 106 coded not applicable occurred exclusively among the latter. Fourteen applicants without classified AI involvement nevertheless entered a general or detailed statement, with low estimated AI text share (M = 1.7%, maximum 5.1%); these were excluded from the principal model.

Table 2. AI disclosure state by depth of AI involvement

AI involvement	n	No disclosure recorded	General	Detailed	Not applicable
None	120	0 (0.0%)	13 (10.8%)	1 (0.8%)	106 (88.3%)
Editing only	173	69 (39.9%)	83 (48.0%)	21 (12.1%)	0
Idea generation	180	80 (44.4%)	80 (44.4%)	20 (11.1%)	0
Substantive generation	144	87 (60.4%)	39 (27.1%)	18 (12.5%)	0
Total	617	236	215	60	106

Note. Percentages are within AI-involvement category. The principal model includes the 497 applicants with recorded AI involvement.

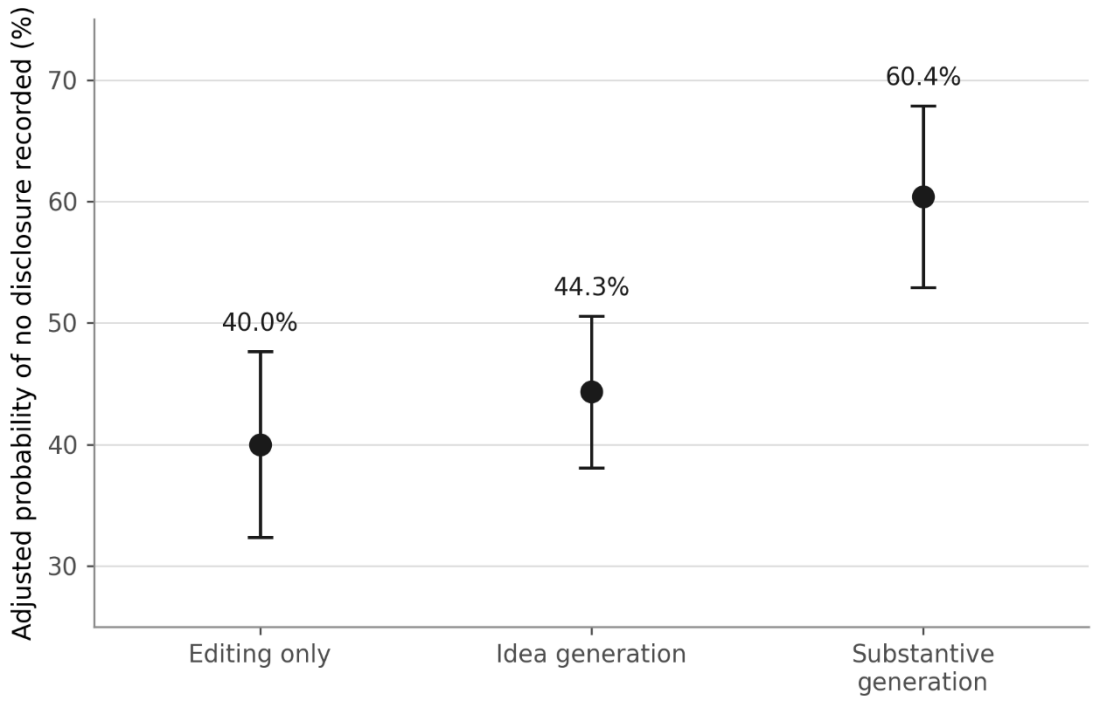

Figure 1. Adjusted probability of no AI disclosure being recorded, by functional depth of AI assistance, with 95% confidence intervals. Estimates are average marginal predicted probabilities from the model in Table 3

4.2 The Disclosure Threshold

The descriptive pattern rose from 39.9% with editing only to 44.4% with idea generation and 60.4% with substantive generation. Table 3 reports the adjusted model. The apparent difference between editing and idea generation was not reliable, OR = 1.20, 95% CI [0.75, 1.93], p = .455. Substantive generation was associated with more than twice the odds that no disclosure was recorded relative to editing only, OR = 2.32, 95% CI [1.49, 3.62], p < .001, q = .002, and the joint contribution of the AI-depth terms was significant, F(2, 57) = 9.11, p < .001.

The direct contrast required by the threshold interpretation also held. Substantive generation differed from idea generation, the level immediately below it, OR = 1.94, 95% CI [1.31, 2.87], F(1, 57) = 10.88, p = .002, q = .003. Adjusted predicted probabilities of no recorded disclosure were 40.0% (95% CI [32.4, 47.6]) for editing only, 44.3% ([38.1, 50.6]) for idea generation and 60.4% ([52.9, 67.9]) for substantive generation, shown in Figure 1.

Equivalence between editing and idea generation could not be demonstrated. Two one-sided tests failed at a strict margin of OR 0.80 to 1.25 (p = .430) and at a lenient margin of OR 0.67 to 1.50 (p = .177), passing only at OR 0.50 to 2.00 (p = .017), too wide to inform. The data fail to reject equality between the two lower categories but do not establish it, so a discontinuity has not been demonstrated in the statistical sense.

Because disclosure may not reasonably be expected for sentence-level correction, the model was refitted excluding editing-only cases, leaving 324 applicants across 55 counsellors. The estimate was essentially unchanged: substantive generation carried almost twice the odds of no recorded disclosure relative to idea generation, OR = 1.96, 95% CI [1.28, 3.01], p = .002. The English-first-language association weakened to p = .087 (OR = 1.49, 95% CI [0.94, 2.37]), consistent with the loss of a third of the cases.

Table 3. Adjusted odds of no AI disclosure being recorded (N = 497)

Predictor	OR	95% CI	p	q
Idea generation vs editing only	1.20	0.75–1.93	.455	.853
Substantive generation vs editing only	2.32	1.49–3.62	< .001	.002
Substantive vs idea generation	1.94	1.31–2.87	.002	.003
Feedback only vs no intermediary	0.82	0.40–1.68	.591	.887
Collaborative revision vs no intermediary	0.98	0.51–1.88	.941	.997
Intermediary-directed vs no intermediary	0.93	0.41–2.11	.865	.997
Undergraduate vs postgraduate	1.22	0.79–1.89	.364	.853
First-generation student	1.07	0.70–1.64	.759	.997
English first language	1.67	1.17–2.39	.005	.038
Woman vs man	0.93	0.69–1.27	.655	.941
Non-binary / prefer not to say vs man	0.84	0.28–2.56	.760	.997
Income ₹15–30 lakh vs under ₹15 lakh	0.96	0.58–1.61	.889	.997
Income ₹30–60 lakh vs under ₹15 lakh	0.99	0.52–1.90	.985	.997
Income ₹60 lakh–₹1.2 crore vs under ₹15 lakh	1.02	0.58–1.80	.950	.997
Income over ₹1.2 crore vs under ₹15 lakh	1.12	0.54–2.31	.759	.997

Note. Binary outcome: no AI disclosure recorded. N = 497, 58 counsellor clusters, cluster-robust standard errors. References are editing only and no intermediary involvement. q values are Benjamini-Hochberg adjusted and are reported in Tables 3 and 4

4.3 Intermediary Control and Interaction

Intermediary control was not detectably associated with AI non-disclosure. None of the three contrasts approached statistical significance, and the joint intermediary-control test was also non-significant, F(3, 57) = 0.28, p = .841.

The interaction requires care. A cluster-robust joint test of the six product terms was significant, F(6, 57) = 3.31, p = .007, but this was driven by a single cell: of the ten applicants who used substantive generation with no intermediary involvement, nine had no disclosure recorded. Excluding that cell removed the effect entirely, F(2, 57) = 0.02, p = .984, and excluding the no-intermediary category left it non-significant, F(4, 57) = 1.82, p = .138. The interaction is an artefact of one sparse cell.

This null result is informative but narrow: it concerns the disclosure of AI, not of intermediary assistance. The dataset contains no parallel outcome showing whether applicants reported feedback, collaborative revision or intermediary-directed drafting to the university.

4.4 Applicant Characteristics

English-first-language applicants had higher adjusted odds that no AI disclosure was recorded, OR = 1.67, 95% CI [1.17, 2.39], p = .005, q = .038. First-generation status, application level, gender and household income were not associated; the

gender terms were jointly non-significant, $F(2, 57) = 0.12$, $p = .885$, as were the income terms, $F(4, 57) = 0.07$, $p = .990$. All coefficients appear in Table 3.

4.5 Robustness to Destination and Organisation

Institutions differ in whether and how they ask about AI, and organisations differ in advising practice and destinations served, so the model was refitted with intended destination and with organisation fixed effects. The estimate for substantive generation was unchanged or slightly larger throughout: OR = 2.36 [1.50, 3.71] with destination, OR = 2.43 [1.54, 3.84] with organisation, and OR = 2.46 [1.54, 3.92] with both, all $p < .001$. The association is therefore not an artefact of deeper AI use concentrating in particular destinations or organisations.

4.6 Disclosure Specificity and AI Text Share

Among the 261 AI-involved applicants who made any disclosure, detailed rather than general disclosure was recorded for 20.2% of editing-only cases, 20.0% of idea-generation cases and 31.6% of substantive-generation cases. In the adjusted model, substantive generation carried roughly twice the odds of detailed disclosure relative to editing only, but imprecisely, OR = 2.07, 95% CI [0.91, 4.72], $p = .084$; idea generation did not differ, OR = 1.08, 95% CI [0.52, 2.27], $p = .832$, and the joint test was non-significant, $F(2, 52) = 1.75$, $p = .184$. The analysis is conditional on having disclosed and rests on 59 detailed disclosures; treat it as exploratory.

Applicants with no recorded disclosure had a higher crude estimated AI text share than applicants who disclosed (32.6% vs 24.8%), but that difference was compositional rather than evidence of additional concealment within assistance levels. It disappeared after stratification: editing only, $p = .134$; idea generation, $p = .850$; substantive generation, $p = .343$. The crude comparison reflects the greater concentration of substantive-generation cases in the no-disclosure group.

4.7 Disclosure and the Companion Integrity Constructs

Table 4 reports these analyses. After adjustment for AI depth, intermediary control, and applicant characteristics, no evidence of association was found between non-disclosure and the Authorial Displacement Index, b = -0.002, 95% CI [-0.855, 0.851], $p = .997$, perceived ownership, b = 0.067, 95% CI [-0.057, 0.191], $p = .289$, or explainability, b = -0.059, 95% CI [-0.238, 0.120], $p = .517$. None approached significance before or after false discovery rate adjustment.

Precision differs across the three. Expressed in standard deviations of each outcome, the displacement interval spans -0.16 to +0.16, excluding all but small effects; the ownership interval spans -0.08 to +0.25 and the explainability interval -0.23 to +0.11, both compatible with effects of up to roughly a quarter of a standard deviation. The displacement null is therefore reasonably precise; the other two are consistent with absence of association but do not establish it.

Missing data on these two measures were limited and unpatterned: 22 of 497 cases (4.4%) lacked an ownership rating and 26 (5.2%) an explainability score. Missingness did not differ by disclosure state ($p = 1.000$ and $p = .733$) or by AI depth ($p = .619$ and $p = .702$), and complete-case analysis was used.

Self-attributed authorship showed the same pattern. The unadjusted association between no disclosure and naming oneself as principal author was negligible, $\varphi = -.037$, $p = .409$, and an adjusted logistic model returned the same result, OR = 1.08, 95% CI [0.72, 1.61], $p = .725$. Applicants with no recorded disclosure were therefore no more displaced, no less able to claim or explain their essays, and no less likely to identify themselves as the author.

Table 4. No recorded disclosure and the companion integrity constructs

Outcome	Effect of no disclosure	95% CI	Standardised CI	p	q
Authorial Displacement Index	-0.002	-0.855 to 0.851	-0.16 to +0.16	.997	.997
Perceived ownership (1–7)	0.067	-0.057 to 0.191	-0.08 to +0.25	.289	.853
Explainability (0–10)	-0.059	-0.238 to 0.120	-0.23 to +0.11	.517	.862
Names self principal author (unadjusted)	$\varphi = -.037$	—	—	.409	.853
Names self principal author (adjusted)	OR = 1.08	0.72–1.61	—	.725	.997

Note. Linear models adjusted for AI depth, intermediary control, first-generation status, application level, gender, household income and English-first-language status, with counsellor-clustered standard errors. Standardised CI expresses the interval in standard deviations of the outcome. Authorship is reported both unadjusted (phi) and from an adjusted logistic model with the same covariates.

5. Discussion

5.1 The Shape of the Pattern

The central finding is specific. AI disclosure did not decline steadily with every increase in assistance. Editing and idea generation were not distinguishable, and the only reliable difference appeared when AI crossed into substantive text generation, holding both against editing-only use and against idea generation directly. Two qualifications belong with that claim: equivalence testing failed at conventional margins, so editing and ideation are not shown to be the same; and the categories were defined by function before the analysis, so they fix in advance where any break could appear.

This matters because policy requires a boundary. A rule based on the mere presence of AI treats spell-checking, brainstorming and generated paragraphs as equivalent, which is both overinclusive and hard to enforce. The result is consistent with a functional threshold: disclosure becomes especially important when AI produces language the institution may reasonably attribute to the applicant. It identifies, among the boundaries examined, the one at which disclosure behaviour changed.

The pattern also cautions against moralising a descriptive result. The outcome is that no disclosure appeared in the final materials, not that an applicant concealed use. Disclosure behaviour depends on institutional wording, perceived stakes and available channels; Kirsanov et al. (2025) similarly read silence under ambiguous rules as risk management rather than dishonesty.

5.2 Disclosure Is Not Authorship Integrity

These analyses change how disclosure should be understood. Once AI depth was controlled, no association was found between non-disclosure and displacement, ownership, explainability or self-attributed authorship. For displacement the interval is narrow enough to exclude all but small effects; for ownership and explainability it is not, so the correct statement is that no evidence of association was found rather than that the constructs are independent. Even so, disclosure is not a usable shortcut for determining whose essay it is.

This separation is conceptually plausible. A person can endorse and understand language produced by a tool while failing to report the production process, and another may disclose minor use in detail even though the essay remains close to their unaided account. Transparency concerns what information reaches the evaluator; displacement, ownership and explainability concern the applicant's relation to the text. Treating these as interchangeable would allow a university to infer too much from a single missing statement.

The converse error is equally available. An institution that treats detailed disclosure as evidence of integrity may be rewarding candour rather than authorship, misclassifying in both directions: penalising the transparent applicant whose essay is largely their own, and clearing the silent applicant whose essay is not.

The result also marks the boundary between this analysis and the companion displacement paper, which uses the same cohort but a different question and primary outcome.

5.3 The Asymmetry Between AI and Human Assistance

Intermediary-control category was not detectably associated with recorded AI disclosure after adjustment, and the apparent interaction rested on a single ten-case cell. Intensity of intermediary control is not the same thing as whether a counsellor advised, discouraged or facilitated disclosure, which was not measured. The finding establishes only that the degree of control counsellors exercised over the text did not track disclosure.

The absence of that variable nevertheless exposes an institutional asymmetry. Application systems may ask whether an applicant used a chatbot while remaining silent about intensive human rewriting, even though consultants and commissioned agents have long shaped application texts (T. J. Huang, 2024; I. Y. Huang et al., 2022; Bretag et al., 2019; Lancaster, 2020). A proportionate policy should define assistance by function across both machine and human contributors.

This is not an argument for treating all assistance as misconduct but for consistency: institutions should apply the same functional test to human and machine contributors.

5.4 English-first-language Applicants

The English-language result was unexpected and survived false discovery rate adjustment. Applicants reporting English as a first language were more likely to have no disclosure recorded, which runs against a simple interpretation in which non-disclosure arises from linguistic difficulty among multilingual applicants. It may instead reflect confidence in using AI without perceiving it as compensatory, or differing school and destination norms: such applicants were predominantly those educated in English-medium international curricula.

The result should remain secondary: the variable is not a direct measure of policy comprehension, and replication with direct measurement is needed before assigning a mechanism.

5.5 Why the Threshold May Fall Where It Does

The data cannot establish why the difference falls where it does, and the accounts below are propositions for future research rather than findings. The most parsimonious is that the categories differ in what they leave behind. Editing and idea generation produce no artefact a reader could attribute to a machine. Substantive generation does: an applicant who has incorporated generated wording holds information the reader does not, so disclosing becomes a decision to surrender an informational advantage.

A second concerns how applicants construe the rule: if reportability is understood as tracking authorship rather than tool use, editing and ideation may not register as reportable at all (Yusuf et al., 2025). A third is institutional: if prompts ask about writing the essay rather than using a tool, applicants who used AI for structure may reasonably answer that they wrote it themselves. None of these supports treating a missing statement as evidence of intention.

5.6 Implications for Admissions Governance

One implication follows from the evidence: applicants whose essays incorporated AI-generated text were more likely to have no disclosure recorded. The recommendations below are design proposals consistent with that finding.

1. Define material assistance by function, distinguishing correction, brainstorming, restructuring, substantive generation and fabrication. The present evidence places particular weight on incorporated AI-generated text.

2. Provide a standard disclosure channel, and record which one was offered. A dedicated field reduces the ambiguity of deciding whether disclosure belongs in the essay, an additional-information section, or an email. Institutional responses remain divergent (UNESCO, 2023), and applicants applying across several countries meet inconsistent expectations within a single cycle. UCAS distinguishes acceptable use for brainstorming and structuring from generating the statement itself, and requires an explicit attestation (UCAS, 2025).

3. Ask for proportional detail and separate disclosure from adjudication. A useful statement identifies what the tool did and whether generated wording was incorporated, without a complete interaction log; a disclosed use should inform whether the artefact supports the intended inference, not automatically trigger an adverse decision.

4. Apply functional symmetry to human assistance, stating when third-party drafting or substantial rewriting must be reported, including by paid counsellors and commissioned agents, and requesting no more private data than the judgement requires.

5. Use verification where the inference is high stakes. A short follow-up question or supervised response tests whether an applicant can extend central claims, shifting the question from what a text looks like to what its author can account for. Institutions should not demand unrestricted access to private accounts.

5.7 Recording the Disclosure Route

The clearest methodological lesson is that the disclosure route must be recorded. Two applicants coded identically here may have faced entirely different situations: one was asked a direct question about AI and answered nothing; the other was never asked and simply did not volunteer the information. Those are different acts warranting different institutional responses.

A usable classification would separate at least five conditions, ordered by the strength of the expectation placed on the applicant: no identifiable route; a general additional-information section; AI disclosure invited but optional; a dedicated mandatory AI question; and an attestation that the work is the applicant's own. Only the last three request information rather than allow it to be volunteered. Recording this would show whether the gap at substantive generation reflects applicant judgement, institutional design, or their interaction.

5.8 Limitations

1. The observational design does not support causal claims, and unmeasured differences may influence both assistance and disclosure.

2. The disclosure route was not recorded per case. A dedicated AI question and a general additional-information box are ethically different instruments, and the study cannot distinguish an applicant declining an explicit request from one not volunteering unsolicited information. This is the largest limitation on the interpretation of the outcome.

3. The functional categories were defined before analysis and therefore fix in advance where a discontinuity could be observed. Equivalence between editing and idea generation was also not demonstrated, so the absence of a detectable difference is not evidence that the two elicit the same behaviour.

4. Of the 617 applications, 179 had been finalised for intended submission but not yet transmitted. Those applicants had not reached the point at which a submission-stage disclosure field would be completed, so non-disclosure may be overstated among them.

5. Disclosure was coded once per case without a second independent coder, so no reliability statistic is available for the primary outcome, for exposure classification, or for the sentence-level coding underlying estimated AI text share.

6. The category none denotes AI involvement below the classification threshold rather than demonstrable absence of tool contact, so the separation between disclosure states and involvement categories is partly a product of that definition. Estimated AI text share is defined only where AI records were voluntarily supplied and is used descriptively.

7. AI involvement was reconstructed from multiple sources but depended partly on self-report, and misclassification toward under-recording is plausible. The sample overrepresents advised applicants, so prevalence estimates should not be treated as population rates.

8. The disclosure outcome concerned AI only; the study cannot estimate whether applicants disclosed human-intermediary assistance. The specificity analysis was conditional on having disclosed and is exploratory, and English-first-language status is a coarse proxy.

9. The data represent one application cycle across several destinations with different policies, and disclosure behaviour may change as forms and rules evolve. Disclosure categories also simplify a drafting process in which applicants may move between editing, ideation and generation across drafts.

6. Conclusion

AI disclosure in university applications did not decline gradually across all forms of assistance. Editing and idea generation were not distinguishable; the only reliable difference appeared where AI generated substantive text incorporated into the final essay, separating substantive generation from the level immediately below it as well as from editing alone. That is the boundary at which reporting rules most need to be unmistakable, although the study locates it among the categories it defined rather than along a continuum.

No evidence was found that recorded disclosure was associated with authorial displacement, ownership, explainability or self-attributed authorship. A missing disclosure statement therefore cannot establish that an applicant does not understand or identify with the submitted essay. Transparency and authorship should be treated as separate ethical dimensions.

The practical response is neither universal prohibition nor detector-led suspicion. Universities should specify material assistance by function, create a standard and proportionate disclosure route, record which route each applicant was offered, apply equivalent principles to human and machine drafting, and verify high-stakes claims through evidence the applicant can explain.

Declarations

Ethics approval and consent to participate: The study was approved by an independent ethics review board prior to recruitment. Participating organisations distributed a standard invitation; counsellors did not obtain consent, were not told who declined, and were instructed not to discuss participation during advising. Applicants consented directly to the research team, and participation or refusal had no effect on the counselling received. Written informed consent was obtained from applicants aged 18 and over; those under 18 completed a written assent form with separate parental or guardian consent.

Consent for publication. Not applicable. No identifiable individual information is reported.

Availability of data and materials. The de-identified case-level dataset, codebook and analysis scripts will be retained for at least ten years after publication and are available from the author on reasonable request, subject to the agreed consent terms and to a use that protects participant confidentiality. Identifiable consent records, original application materials and the linkage key are stored separately under restricted access and will be securely destroyed five years after publication. Direct identifiers were removed before analysis.

Competing interests. The author works professionally in university-application counselling and has a commercial interest in application guidance, but did not recruit participants, did not counsel any participating applicant, and had no access to applicant identities during analysis. Participating organisations were independent of the author’s practice.

Funding. No external funding was received.

Author contributions. The author conceived the study, specified the analytic framework, supervised data collection, conducted and verified the analysis, interpreted the results, and wrote the manuscript.

Use of generative AI. An AI assistant (Claude, Anthropic) was used at one stage of this work: verification of the reference list. Each entry was checked against publisher records for author names, year, title, journal, volume, issue, pages and DOI, and every correction was made by the author. The tool was not used to generate, classify or analyse the data, to code AI involvement, disclosure or intermediary control, or to draft the argument. The research questions, design, coding framework, statistical analysis, findings and conclusions are the author's own. No artificial intelligence system is an author, and the author takes responsibility for the content and integrity of the work.

Acknowledgments

The author thanks the ten counselling organisations that agreed to take part and made their records available for independent coding, and the applicants, parents and counsellors who gave their time to the study. The manuscript was prepared by the author without personal assistance.

Author contributions

Karan Gupta is the sole author. He conceived the study, specified the conceptual framework and research questions, designed the coding scheme, supervised data collection, conducted and verified the statistical analysis, interpreted the results, and drafted and revised the manuscript. The author read and approved the final manuscript.

Funding

This work received no external funding. No grant or financial support was received from any agency in the public, commercial or not-for-profit sectors.

Competing interests

The author is the founder of Karan Gupta Consulting and serves as Managing Director for India and South Asia at IE University; both roles are relevant to the sector examined here and are disclosed as potential competing interests. The author did not recruit participants, did not counsel any participating applicant, and had no access to applicant identities during analysis. Participating organisations were independent of the author's practice, and no application to IE University formed part of the dataset.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Redfame Publishing.

The journal’s policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

References

Birks, D., & Clare, J. (2023). Linking artificial intelligence facilitated academic misconduct to existing prevention frameworks. *International Journal for Educational Integrity, 19*, 20. https://doi.org/10.1007/s40979-023-00142-3

Bjelobaba, S., Waddington, L., Perkins, M., Foltýnek, T., Bhattacharyya, S., & Weber-Wulff, D. (2025). Maintaining research integrity in the age of GenAI: An analysis of ethical challenges and recommendations to researchers. *International Journal for Educational Integrity, 21*, 18. https://doi.org/10.1007/s40979-025-00191-w

Bretag, T., Harper, R., Burton, M., Ellis, C., Newton, P., van Haeringen, K., Saddiqui, S., & Rozenberg, P. (2019). Contract cheating and assessment design: Exploring the relationship. *Assessment & Evaluation in Higher Education, 44*(5), 676–691. https://doi.org/10.1080/02602938.2018.1527892

Chen, C., & Jia, X. (2026). When researchers use AI: Public trust, ethical judgments, and the perceived value of academic research. *AI and Ethics, 6*, Article 223. https://doi.org/10.1007/s43681-026-01039-w

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

Dawson, P., Bearman, M., Dollinger, M., & Boud, D. (2024). Validity matters more than cheating. *Assessment & Evaluation in Higher Education, 49*(7), 1005–1016. https://doi.org/10.1080/02602938.2024.2386662

Eaton, S. E. (2023). Postplagiarism: Transdisciplinary ethics and integrity in the age of artificial intelligence and neurotechnology. *International Journal for Educational Integrity, 19*, 23. https://doi.org/10.1007/s40979-023-00144-1

Foltýnek, T., Bjelobaba, S., Glendinning, I., Khan, Z. R., Santos, R., Pavletić, P., & Kravjar, J. (2023). ENAI recommendations on the ethical use of artificial intelligence in education. *International Journal for Educational Integrity, 19*, 12. https://doi.org/10.1007/s40979-023-00133-4

Gebre-Medhin, B., Giebel, S., Alvero, A. J., Antonio, A. L., Domingue, B. W., & Stevens, M. L. (2022). Application essays and the ritual production of merit in US selective admissions. *Poetics, 94*, 101706. https://doi.org/10.1016/j.poetic.2022.101706

Huang, I. Y., Williamson, D., Lynch-Wood, G., Raimo, V., Rayner, C., Addington, L., & West, E. (2022). Governance of agents in the recruitment of international students: A typology of contractual management approaches in higher education. *Studies in Higher Education, 47*(6), 1150–1170. https://doi.org/10.1080/03075079.2020.1861595

Huang, T. J. (2024). Translating authentic selves into authentic applications: Private college consulting and selective college admissions. *Sociology of Education, 97*(2), 174–192. https://doi.org/10.1177/00380407231202975

International Center for Academic Integrity. (2021). *The fundamental values of academic integrity* (3rd ed.). International Center for Academic Integrity.

Jones, S. (2013). "Ensure that you stand out from the crowd": A corpus-based analysis of personal statements according to applicants' school type. *Comparative Education Review, 57*(3), 397–423. https://doi.org/10.1086/670666

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., & … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kirsanov, O., Kushwah, L., & Selvaretnam, G. (2025). Beyond detection: How students use—and hide—AI in online assessments and what authentic tasks can do about it. *Journal of Academic Ethics, 24*, Article 14. https://doi.org/10.1007/s10805-025-09691-3

Lancaster, T. (2020). Academic discipline integration by contract cheating services and essay mills. *Journal of Academic Ethics, 18*, 115–127. https://doi.org/10.1007/s10805-019-09357-x

Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023). GPT detectors are biased against non-native English writers. *Patterns, 4*(7), 100779. https://doi.org/10.1016/j.patter.2023.100779

Lira, B., Gardner, M., Quirk, A., Stone, C., Rao, A., Ungar, L., Hutt, S., Hickman, L., D'Mello, S. K., & Duckworth, A. L. (2023). Using artificial intelligence to assess personal qualities in college admissions. *Science Advances, 9*(41), eadg9405. https://doi.org/10.1126/sciadv.adg9405

Lum, Z. C., Guntupalli, L., Saiz, A. M., Leshikar, H., Le, H. V., Meehan, J. P., & Huish, E. G. (2024). Can artificial intelligence fool residency selection committees? Analysis of personal statements by real applicants and generative AI, a randomized, single-blind multicenter study. *JBJS Open Access, 9*(4), e24.00028. https://doi.org/10.2106/JBJS.OA.24.00028

Lund, B. D., Wang, T., Mannuru, N. R., Nie, B., Shimray, S., & Wang, Z. (2023). ChatGPT and a new academic reality: Artificial intelligence-written research papers and the ethics of large language models in scholarly publishing. *Journal of the Association for Information Science and Technology, 74*(5), 570–581. https://doi.org/10.1002/asi.24750

Nikula, P.-T., & Kivistö, J. (2020). Monitoring of education agents engaged in international student recruitment: Perspectives from agency theory. *Journal of Studies in International Education, 24*(2), 212–231. https://doi.org/10.1177/1028315318825338

Perkins, M., Roe, J., Postma, D., McGaughran, J., & Hickerson, D. (2024). Detection of GPT-4 generated text in higher education: Combining academic judgement and software to identify generative AI tool misuse. *Journal of Academic Ethics, 22*(1), 89–113. https://doi.org/10.1007/s10805-023-09492-6

Ruttenberg, D. (2026). The AIR framework for research transparency: A critical analysis of stage-specific AI disclosure in the context of accessibility and research integrity. *AI & Society*. Advance online publication.

https://doi.org/10.1007/s00146-026-03082-x

Ucan, S., & Demirel Ucan, A. (2026). Generative AI in student writing: A six-level taxonomy for responsible use in higher education assessment. *Journal of Academic Ethics, 24*, Article 63. https://doi.org/10.1007/s10805-026-09739-y

UCAS. (2025). *A guide to using AI and ChatGPT with your personal statement*. UCAS. https://www.ucas.com/applying/applying-university/writing-your-personal-statement/guide-using-ai-and-chatgpt-your-personal-statement

UNESCO. (2023). *Guidance for generative AI in education and research*. UNESCO.

Waltzer, T., Pilegard, C., & Heyman, G. D. (2024). Can you spot the bot? Identifying AI-generated writing in college essays. *International Journal for Educational Integrity, 20*, 11. https://doi.org/10.1007/s40979-024-00158-3

Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltýnek, T., Guerrero-Dib, J., Popoola, O., Šigut, P., & Waddington, L. (2023). Testing of detection tools for AI-generated text. *International Journal for Educational Integrity, 19*, 26. https://doi.org/10.1007/s40979-023-00146-z

Whitley, B. E., & Keith-Spiegel, P. (2001). Academic integrity as an institutional issue. *Ethics & Behavior, 11*(3), 325–342. https://doi.org/10.1207/S15327019EB1103_9

Yusuf, A., Mouas, S., & Al-khresheh, M. H. (2025). Understanding the psychological factors associated with (non) disclosure behaviour after GenAI usage. *Journal of Academic Ethics, 24*, Article 34. https://doi.org/10.1007/s10805-025-09712-1