

Rethinking Essays: Alternatives in Education using AI

KAMAL SHARMA (1541791), Frankfurt University of Applied Sciences, Germany

HAFIZ MUHAMMAD ALI (1556223), Frankfurt University of Applied Sciences, Germany

RANJIT KHUDÉ (1542561), Frankfurt University of Applied Sciences, Germany

RENCY RAJESHBHAI PADASALA (1612037), Frankfurt University of Applied Sciences, Germany

TANISHKA AGALE (1541379), Frankfurt University of Applied Sciences, Germany

Assessments are an important part of higher education because they show what students have learned and can do. In the past, written exams and essays were the main ways to assess students, but other methods such as oral presentations, step-by-step assignments, and performance-based projects are now being used more often. With the rise of artificial intelligence (AI) tools like ChatGPT, it has become easier for students to produce written work quickly, raising concerns about fairness and honesty in assessments.

This study explores how students would like to be assessed in higher education, focusing on three formats: oral, process-based (e.g., essays, journals), and performance-based (e.g., projects, portfolios) assessments. A qualitative approach was used, and data was collected through an open-ended online questionnaire. The responses were analyzed using thematic analysis to find common patterns in student preferences.

The results show that students value fairness, clarity, and the chance to show practical skills. Many students prefer performance-based projects for their real-world relevance, while others appreciate process-based methods for encouraging reflection and deeper learning. Oral assessments were valued by some for improving communication skills but were also seen as stressful by others.

These findings can help educators design assessment methods that are fair, engaging, and better suited for learning in the age of AI.

ACM Reference Format:

Kamal Sharma (1541791), Hafiz Muhammad Ali (1556223), Ranjit Khude (1542561), Rency Rajeshbhai Padasala (1612037), and Tanishka Agale (1541379). 2025. Rethinking Essays: Alternatives in Education using AI. *Proc. ACM Meas. Anal. Comput. Syst.* 37, 4, Article 111 (August 2025), 20 pages. <https://doi.org/XXXXXX.XXXXXXX>

1 Introduction

Topic

In higher education, student learning outcomes and skills are evaluated through various assessment methods [2]. These methods aim to measure students' overall abilities in a particular domain. Traditionally, written exams and essays have been the most common forms of assessment. In recent years, however, alternative methods have gained prominence, including oral presentations, step-by-step assignments (such as reflective journals or draft

Authors' Contact Information: Kamal Sharma (1541791), Frankfurt University of Applied Sciences, Frankfurt am Main, Germany, kamal.sharma@stud.fra-uas.de; Hafiz Muhammad Ali (1556223), Frankfurt University of Applied Sciences, Frankfurt am Main, Germany, hafiz.ali2@stud.fra-uas.de; Ranjit Khude (1542561), Frankfurt University of Applied Sciences, Frankfurt am Main, Germany, ranjit.khude@stud.fra-uas.de; Rency Rajeshbhai Padasala (1612037), Frankfurt University of Applied Sciences, Frankfurt am Main, Germany, rency.padasala@stud.fra-uas.de; Tanishka Agale (1541379), Frankfurt University of Applied Sciences, Frankfurt am Main, Germany, tanishka.agale@stud.fra-uas.de.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2025 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM 2476-1249/2025/8-ART111

<https://doi.org/XXXXXX.XXXXXXX>

essays), and performance-based work (such as projects or portfolios) [11]. Together, these approaches provide a more comprehensive picture of students' capabilities.

Problem

With the introduction of Artificial Intelligence (AI) tools such as ChatGPT, producing essays has become significantly easier for students. These tools can generate grammatically correct and well-structured essays with minimal effort [5]. This practice makes it increasingly difficult for instructors to evaluate whether submitted work genuinely reflects a student's knowledge and abilities. While many strategies have been proposed to detect or prevent AI-enabled cheating, less attention has been given to understanding which assessment methods students themselves perceive as fair, engaging, and an authentic test of their skills [12].

Motivation

The purpose of assessment is not only to detect dishonesty but also to foster meaningful learning and active engagement. Well-designed assessments should motivate students and enhance their educational experience. If instructors understand students' preferences, they can select formats that are fair, effective, and aligned with learners' needs [1]. Such alignment could lead to better learning outcomes and help reduce academic dishonesty in the age of AI.

Research Gap

Despite increasing interest in assessment reform, several key gaps remain in the literature:

- (1) **Limited focus on AI-integrated, student-centred models:** While AI tools are becoming more common, little research examines how assessments can be designed to responsibly integrate AI while focusing on student needs and preferences.
- (2) **Over-reliance on traditional essays:** Many institutions continue to depend on essays as the main form of assessment, despite concerns about academic dishonesty and their limitations in evaluating diverse skills.
- (3) **Lack of empirical evaluation of alternatives:** Although oral exams, portfolios, and project-based work have been proposed as alternatives, there is limited empirical research on their effectiveness and acceptance from students' perspectives.
- (4) **Underexplored implementation strategies:** Even when alternatives are discussed, practical guidance on implementing them effectively in classrooms is sparse.

This study addresses these gaps by collecting and analyzing student feedback on different assessment types, evaluating their preferences, and offering practical recommendations for fair, engaging, and AI-resilient assessment design.

Research Question

This study is guided by the following research question: *"How do students like to be assessed in higher education: preferences between oral, process-based (e.g., essays, journals), and performance-based (e.g., projects, portfolios) assessments?"*

2 Related Work

This section reviews prior research relevant to rethinking assessment practices in higher education, particularly in the context of artificial intelligence (AI) and academic integrity. Three closely related studies are discussed, each highlighting alternative approaches and their limitations.

2.1 Employing the Interactive Oral to Mitigate Threats to Academic Integrity from ChatGPT

Newell (2023) investigates the use of *interactive oral assessments* as a means of addressing academic integrity challenges posed by AI tools such as ChatGPT [7]. Instead of relying on traditional essays, students are assessed through live, structured oral discussions designed to evaluate their understanding and critical thinking. The findings suggest that this approach can significantly reduce plagiarism risks and foster authentic learning. However, the study is limited to a single subject area and lacks broader empirical validation across disciplines and larger student populations. This highlights a gap in understanding how interactive oral assessments can be applied and scaled across diverse academic contexts [8].

2.2 Rethinking Assessment for Generative AI: Beyond the Essay

Furze (2024) critiques the continued reliance on essay-based assessments in an era shaped by generative AI. The study proposes alternative assessment formats such as project-based learning, portfolios, and process-driven evaluations, which emphasize the thinking process and creativity rather than solely the written product [9]. While these proposals provide promising pathways for AI-resilient assessment, Furze notes the lack of empirical classroom evidence to confirm their effectiveness in practice. The gap lies in moving from theoretical suggestions to real-world implementation strategies and rigorous evaluation [3].

2.3 Automatic Assessment of Text-Based Responses in Post-Secondary Education: A Systematic Review

Gao et al. (2023) provide a systematic review of automated assessment systems that use AI and natural language processing to evaluate text-based student work, including essays and short answers [10]. These technologies offer efficiency and consistency in grading but raise concerns about fairness, transparency, and adaptability across educational contexts [6]. The review emphasizes that while automated assessment tools are advancing, little research explores how they might be integrated with student-centered approaches. This leaves space for developing hybrid models that combine automation with human judgment, ensuring both efficiency and meaningful evaluation [4].

2.4 Summary of Research Gaps

From these studies, several critical research gaps emerge:

- Limited exploration of AI-integrated, student-centered assessment models.
- Continued overreliance on essays as the dominant assessment practice despite known shortcomings.
- Insufficient empirical evaluation of alternative assessment formats in real classroom settings.
- Lack of practical implementation strategies for AI-resilient assessments.
- Minimal research on hybrid approaches that combine automation with human-centered evaluation.

This project builds on these findings by analyzing student preferences for oral, process-based, and performance-based assessments. By collecting empirical data and highlighting students' perspectives, it aims to provide practical recommendations for designing fair, engaging, and AI-resilient assessment practices in higher education.

3 Method

This study investigates students' preferences for assessment formats in higher education, focusing on oral, process-based (e.g., essays, journals), and performance-based (e.g., projects, portfolios) approaches, particularly in the context of AI-assisted learning. A qualitative survey was conducted to gather detailed insights into students' perceptions and preferences.

3.1 Study Design

We employed a qualitative research design using an online questionnaire created with Google Forms. The questionnaire included demographic questions, one multiple-choice item on preferred assessment format, and several open-ended questions prompting participants to explain their reasoning, experiences, and concerns. To ensure concise yet meaningful responses, a word limit of 30–50 words per question was specified.

3.2 Independent Variables / Factors

The study focused on students' perspectives regarding the following assessment formats:

- **Oral assessments** (e.g., presentations, interviews, vivas)
- **Process-based assessments** (e.g., essays, reflective journals)
- **Performance-based assessments** (e.g., projects, portfolios)

3.3 Dependent Variables / Measures

The dependent measures were based on participants' self-reported perspectives collected via open-ended responses. These included:

- Reasons for preferring a given assessment format
- Positive experiences with different assessment types
- Perceived challenges or frustrations (particularly with essays)
- Perceptions of fairness and authenticity
- Views on whether assessment formats encourage or discourage academic dishonesty
- Descriptions of an ideal assessment format

3.3.1 Objective Measures. No objective or numerical performance measures were collected. Descriptive counts and distributions (e.g., preference percentages, helpful vs. stressful counts) were used to support the qualitative findings.

3.3.2 Subjective Measures. Participants' preferences, perceptions, and rationales were gathered through open-ended survey responses. These were analyzed thematically and supported with computational text analysis.

3.4 Confounds

Potential confounds considered in the study included:

- **Academic pressure** influencing how students perceive fairness and difficulty
- **Social desirability bias** in reporting preferred methods
- **Cultural differences** shaping comfort with oral or written tasks
- **Incentives** (1 credit point for participation), which may have influenced participation but not the content of responses

3.5 Procedure

The survey link was distributed via university email and social media channels (e.g., WhatsApp groups). After reviewing the study description and consent form, participants provided demographic information and completed the questionnaire. All responses were collected online and exported from Google Forms into Excel for analysis.

3.6 Participants

A total of **67** students from the international master's program at Frankfurt University of Applied Sciences participated. The study was part of the *Human-Machine Interaction* module during the Summer Semester 2024.

Participation was voluntary, and personal identifiers (name, email, immatriculation ID) were collected solely for awarding credit points and certificates. These identifiers were stored separately from the response data, ensuring anonymity in analysis.

3.7 Data Analysis

Responses were analyzed using a hybrid approach that combined thematic coding with computational text analysis. This ensured that both qualitative depth and quantitative transparency were achieved.

- **Thematic analysis:** Open-ended responses were coded inductively to identify recurring patterns and grouped into higher-level categories such as preferences, challenges, and positive experiences.
- **Computational text analysis:** Responses were pre-processed by removing stopwords and irrelevant terms. TF-IDF vectorization and KMeans clustering were applied to identify themes, while word frequency analysis supported the thematic findings.
- **Visualizations:** Word clouds, bar charts, pie charts, and heatmaps were generated to highlight dominant terms, thematic clusters, and comparative patterns across assessment types.

To provide clarity, Table 1 presents an overview of how each survey question was analyzed, along with the type of visualization generated.

Survey Question	Analysis Method	Visualization / Output
Preferred assessment format (Q1)	Descriptive frequency analysis	Pie chart (distribution of oral, process-based, performance-based preferences)
Reason for preferred format (Q2)	Thematic coding + keyword frequency	Word cloud + bar chart of top terms
Positive experiences (Q3)	Thematic coding + clustering	Word clouds (by type) + comparative bar chart
Challenges with essays (Q4)	Thematic coding + frequency analysis	Bar chart of top challenge terms
Oral assessments: helpful vs. stressful (Q5)	Frequency comparison + keyword extraction	Stacked bar chart + word clouds (helpful vs. stressful)
Comparison of performance-based assessments (Q6)	Thematic coding	Word cloud of top terms
Academic dishonesty perceptions (Q7)	Cross-tabulation + clustering	Heatmap of encourage vs. discourage (by type)
Ideal assessment format (Q8)	Thematic coding + descriptive distribution	Word cloud + bar chart of preferred hybrid/individual models

Table 1. Overview of analysis methods and visual outputs per research question

This structured pipeline allowed us to align qualitative insights with visual evidence, ensuring that the conclusions drawn were supported by both thematic interpretation and transparent descriptive patterns.

4 Results

4.1 Introduction to Findings

This study investigated how students prefer to be assessed, what experiences they find positive or challenging, and how different assessment methods influence academic dishonesty. Through both qualitative clustering of open-ended responses and quantitative analysis of distributions, we obtained a rich picture of how students perceive oral, process-based, and performance-based assessments, as well as their ideal formats for the future.

The results below are structured according to questions asked in form, followed by a final synthesis that integrates the findings into an overall conclusion. Visual diagrams (pie charts, bar graphs, word clouds, and heatmaps) complement the thematic summaries, ensuring both breadth and depth of understanding.

For full code, data, and reproducibility materials, see the [Github Repository](#).

4.2 Student Preferences for Assessment Methods

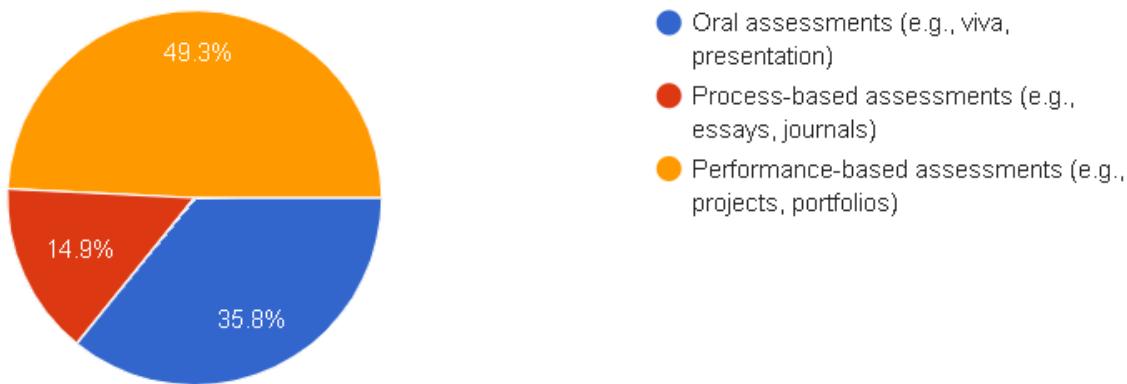


Fig. 1. Distribution of students' preferences across oral, process-based, and performance-based assessment formats.

Figure 1 illustrates students' preferences for different forms of academic assessment, highlighting a clear inclination toward more interactive and applied methods over traditional approaches. Nearly half of the respondents (49.3%) expressed a preference for *performance-based assessments* such as projects and portfolios. This suggests that students value opportunities to demonstrate their knowledge and skills through practical application, creativity, and problem-solving, rather than being confined to theoretical evaluations.

Meanwhile, 35.8% of students favored *oral assessments*, including vivas and presentations. This substantial proportion reflects the recognition of oral methods as a means to test not only knowledge but also communication skills, confidence, and the ability to articulate ideas in real time. Importantly, oral formats may also be perceived as less prone to plagiarism and artificial assistance, given their interactive and immediate nature.

By contrast, only 14.9% of participants preferred *process-based assessments*, such as essays or journals. This relatively low percentage underscores growing discontent with essay-based evaluations, which are often regarded as time-consuming, stressful, and increasingly vulnerable to academic dishonesty, particularly with the rise of AI-enabled writing tools.

Taken together, these findings reveal a clear shift in student attitudes toward more dynamic and authentic assessment practices that align closely with real-world skills and individual learning styles.

4.3 Reasoning: Student Preferences for Assessment Methods

4.3.1 Oral-Based Assessments. Students who expressed a preference for oral-based assessments often highlighted their value in fostering understanding, communication, and real-time interaction, as shown in Figure 2.

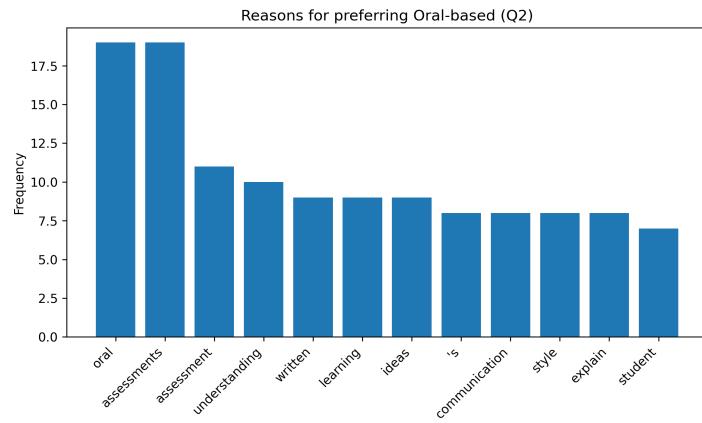


Fig. 2. Key terms associated with oral-based assessment preferences.

The most frequent terms, such as *oral* (19), *assessments* (19), *understanding* (10), and *communication* (8), reveal that learners associate oral formats with active engagement and demonstration of ideas. Oral assessments, such as presentations or vivas, are perceived as opportunities to articulate one's thoughts, test depth of knowledge, and build confidence in communication skills—abilities crucial for both academic and professional development. Additionally, words like *written* (9) and *learning* (9) suggest that students view oral assessments as complementary to written work, offering a different dimension of evaluation.

4.3.2 Process-Based Assessments. Despite being the least popular overall, process-based assessments still hold value for certain students who emphasize learning, reflection, and structured thinking, as illustrated in Figure 3.

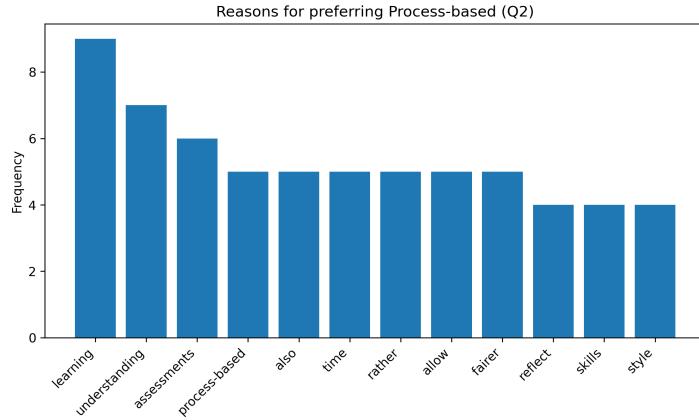


Fig. 3. Key terms associated with process-based assessment preferences.

The most frequent terms include *learning* (9), *understanding* (7), and *reflection* (4), highlighting that students see value in deep analysis and critical reasoning. Keywords like *fairer* (5) and *allow* (5) suggest that these methods are perceived as equitable, giving students time to organize their thoughts without the immediate stress of live performance. However, mentions of *time* (5) and *rather* (5) also point to frustrations, as the extended effort required is sometimes seen as burdensome.

4.3.3 Performance-Based Assessments. Performance-based assessments, such as projects and portfolios, emerged as the most preferred format, strongly associated with real-world application and skill development. Figure 4 presents the key terms driving these preferences.

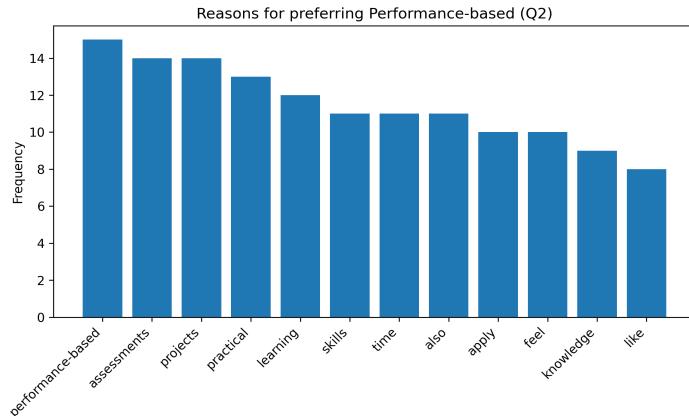


Fig. 4. Key terms associated with performance-based assessment preferences.

The most frequent terms, such as *performance-based* (15), *projects* (14), *practical* (13), and *skills* (11), show that students highly value applied knowledge and creativity. Words like *apply* (10) and *feel* (10) emphasize personal ownership and authentic learning experiences. Unlike traditional essays or oral exams, performance-based

work allows for teamwork, long-term projects, and multi-step problem solving, which align with professional expectations and make students feel their assessments are both fairer and more meaningful.

4.4 Positive Experiences Across Assessment Types

4.4.1 Overall Signal from Positive Experiences. Students' reflections on positive experiences with assessments reveal consistent patterns across different formats. As shown in Figure 5, the word cloud highlights the dominant concepts that participants associate with meaningful and satisfying assessments.

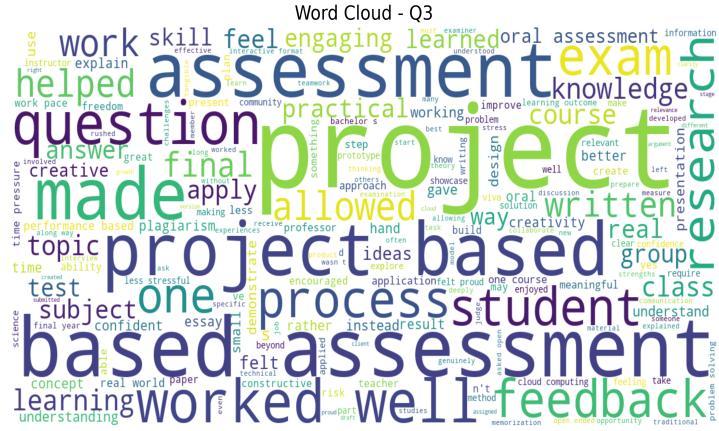


Fig. 5. Word cloud summarizing positive-experience responses.

Key terms such as *projects*, *work*, and *course* appear prominently, indicating that hands-on, course-embedded tasks are widely regarded as rewarding. Students emphasized the value of opportunities to apply knowledge, collaborate, and create tangible outputs, which fostered a sense of competence and progress. Unlike recall-driven or purely theoretical tasks, these formats were appreciated for enabling growth in communication, confidence, and practical problem-solving. Even when written or oral elements were involved, positive experiences concentrated around contexts that provided feedback, interaction, and real-world relevance. This suggests that satisfaction depends less on the assessment format alone and more on whether it affords authentic engagement and deeper learning.

4.4.2 Comparative Distribution of Positive-Experience Keywords. The word cloud is complemented by the quantitative analysis shown in Figure 6, which presents the most frequent terms across students' positive-experience responses.

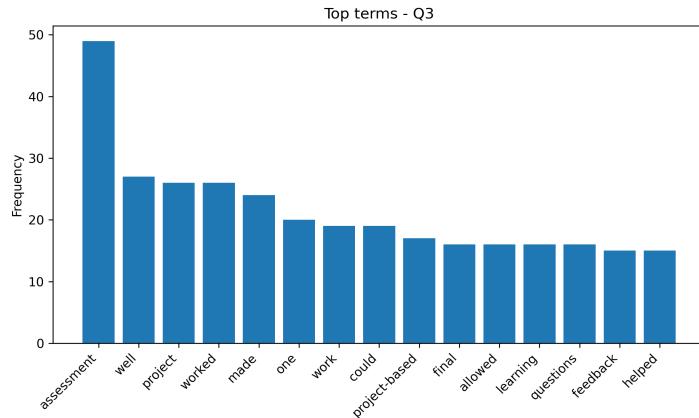


Fig. 6. Top-frequency keywords from positive-experience responses.

High-frequency terms such as *projects*, *practical work*, and *skills* reinforce the observation that students respond most positively to assessments emphasizing application and demonstrable competence. Oral formats were valued for the interaction and feedback they enabled, while process-based formats supported reflection and critical thinking. However, the overall distribution of terms indicates that positivity peaked in performance-based contexts where students could *do*, *build*, *present*, and *refine*. These findings converge with broader preference trends across the study: assessments are seen as most rewarding when they combine collaboration, iteration, and visible outcomes. Together, Figures 5 and 6 provide strong evidence that meaningful engagement—rather than ease or familiarity alone—drives positive student sentiment toward assessment.

4.5 Challenges with Process-Based Assessments

4.5.1 Dominant Challenges Identified in Responses. Students frequently described process-based assessments, such as essays and journals, as demanding and stressful. The word cloud in Figure 7 illustrates the most common terms students used when reflecting on these difficulties.

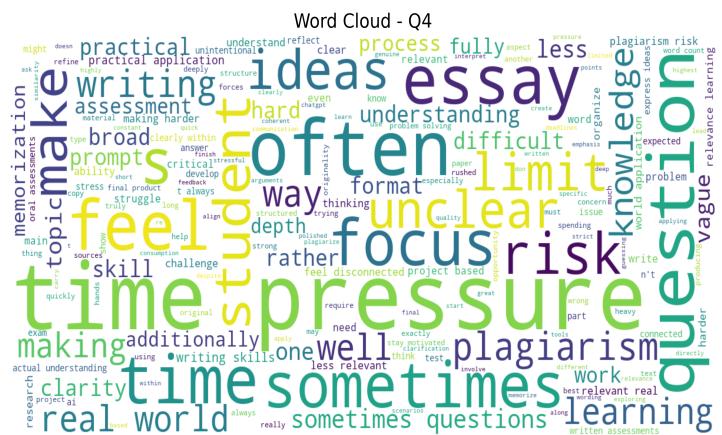


Fig. 7. Word cloud of challenges associated with process-based assessments.

Prominent terms such as *time*, *pressure*, and *focus* highlight the perception that these formats require prolonged effort and concentration. Unlike oral or performance-based tasks, which often allow interaction, creativity, or immediate feedback, process-based assessments place a heavy emphasis on sustained individual work. This structure can be demotivating when outcomes appear disconnected from practical application. Furthermore, repeated references to *plagiarism* and *unclear* reflect frustrations with academic dishonesty risks and vague instructions, making essays appear both less credible and less fair compared to alternative formats.

4.5.2 Frequency Distribution of Key Challenges. The quantitative distribution of frequently cited challenges, shown in Figure 8, reinforces these themes. Here, *time* (9 mentions), *pressure* (6 mentions), and *focus* (6 mentions) emerge as the most significant concerns.

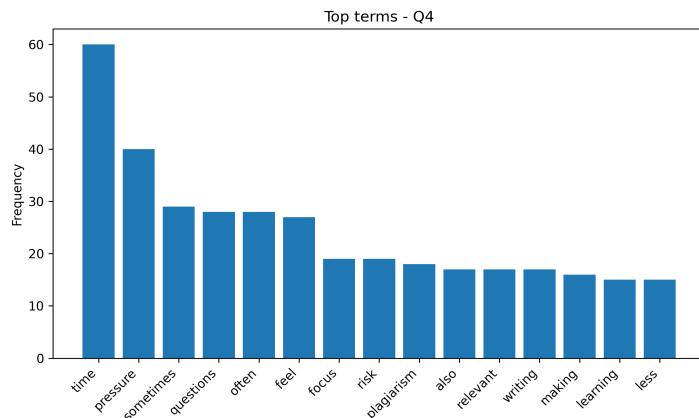


Fig. 8. Most frequently reported challenges in process-based assessments (script-generated bar chart).

Other recurring terms, including *sometimes*, *often*, and *final*, reflect students' anxieties about unpredictability and the high stakes typically associated with essays. Meanwhile, the persistence of *plagiarism* underscores

the vulnerability of these formats in an academic environment increasingly shaped by AI-generated writing. Collectively, these findings echo long-standing critiques of essay-based evaluation: while essays can cultivate critical analysis and structured reasoning, they often impose excessive burdens of time, lack transparency in expectations, and invite academic dishonesty. This combination helps explain their relatively low preference compared to oral or performance-based assessments.

4.6 Oral Assessments: Helpful or Stressful

4.6.1 Overall Distribution of Student Perceptions. Students' perceptions of oral assessments reveal a complex balance between their potential benefits and challenges. As shown in Figure 9, a majority of students (63) described oral assessments as helpful, while a smaller but notable group (48) found them stressful.

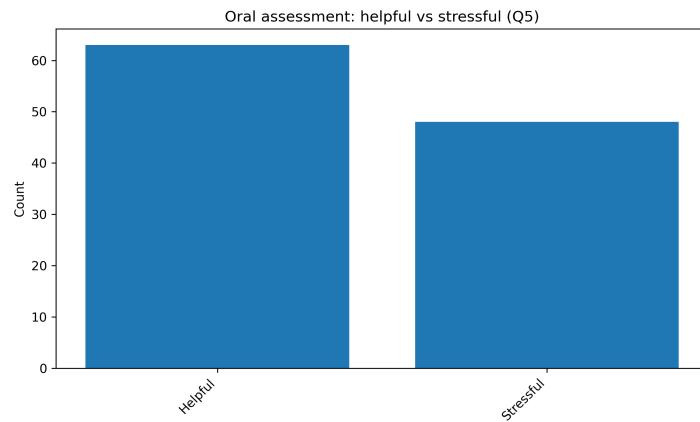


Fig. 9. Distribution of students who found oral assessments helpful vs. stressful.

This distribution illustrates the dual nature of oral formats: they are valued for building confidence, improving communication, and allowing real-time feedback, but they also generate performance anxiety due to immediacy and unpredictability. The results highlight that oral assessments must be carefully designed to maximize authentic learning opportunities while minimizing stressors that could hinder performance.

4.6.2 Keywords Associated with Helpful Experiences. The specific advantages of oral assessments are underscored by the keyword distribution shown in Figure 10. Terms such as *confidence* (31 mentions), *evaluators* (22), and *ideas* (21) were strongly linked to positive experiences.

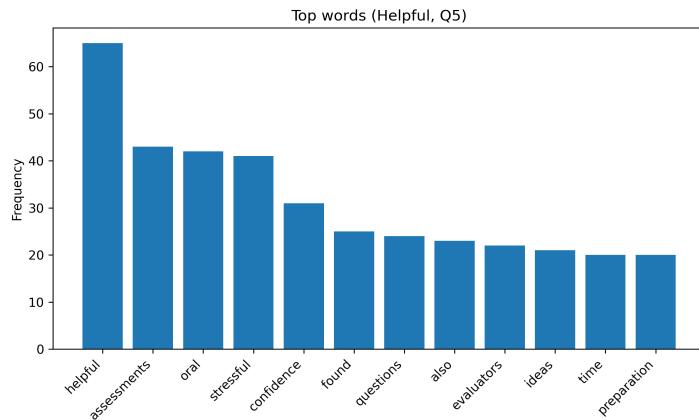


Fig. 10. Most common terms in responses describing oral assessments as helpful.

Students emphasized that oral formats provided opportunities to clarify ideas, receive direct feedback, and strengthen public speaking skills. These experiences were often described as enriching because they enabled students to demonstrate understanding beyond rote memorization and supported the development of interpersonal competencies crucial for professional and academic success.

4.6.3 Keywords Associated with Stressful Experiences. In contrast, Figure 11 highlights the terms most frequently associated with stressful experiences in oral assessments. Words such as *preparation* (20 mentions), *evaluators* (20), and *questions* (19) were common stressors.

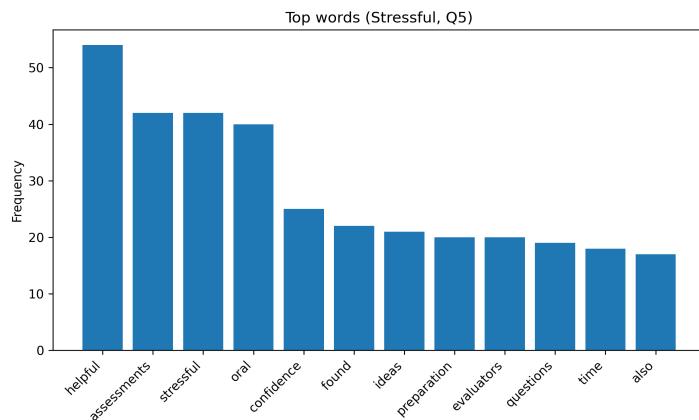


Fig. 11. Most common terms in responses describing oral assessments as stressful.

Students often reported anxiety about unexpected questions and the pressure of performing under observation. For some, oral assessments were perceived not only as evaluations of knowledge but also as tests of emotional resilience and composure. These findings underscore the importance of supportive and transparent assessment environments that balance the benefits of oral formats with the emotional demands they place on learners.

4.7 Performance-Based Assessments: Skills, Creativity, and Real-World Relevance

4.7.1 Key Concepts Highlighted in Student Responses. Students' reflections on performance-based assessments emphasized their value for authentic learning and practical application. As shown in Figure 12, the most frequent terms included *skills*, *practical*, *creativity*, and *projects*.



Fig. 12. Word cloud of student responses on performance-based assessments.

These responses illustrate that performance-based formats are closely linked with opportunities to apply knowledge in real-world contexts, engage in hands-on problem-solving, and demonstrate creativity. Unlike process-based tasks, which often feel abstract, or oral tasks, which emphasize immediacy of response, projects and portfolios were celebrated for their ability to promote deeper engagement, teamwork, and personal initiative. Many students highlighted the importance of collaboration and tangible outputs, which not only validated their knowledge but also gave them a sense of ownership and accomplishment.

4.7.2 Frequency Distribution of Core Themes. The frequency analysis in Figure 13 further reinforces these findings, with terms such as *skills*, *practical*, and *creativity* emerging as dominant.

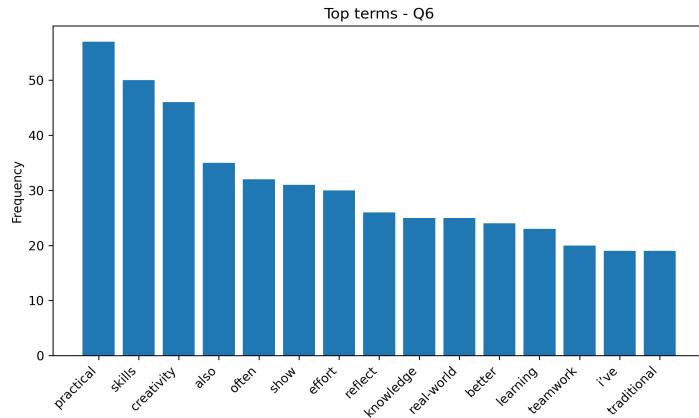


Fig. 13. Top-frequency keywords describing performance-based assessments.

Additional terms such as *team*, *involvement*, and *problems* highlight the collaborative and problem-solving aspects that students strongly associate with project work. This suggests that learners see performance-based tasks not only as fairer indicators of competence but also as direct preparation for professional environments. Collectively, these insights demonstrate why performance-based assessments received the highest overall preference: they integrate authenticity, collaboration, and skill development, making them both effective and engaging measures of student learning.

4.8 Academic Dishonesty Across Assessment Formats

4.8.1 Overall Signals in Student Responses. Students' perceptions of academic dishonesty varied significantly across assessment types. As illustrated in Figure 14, the most frequent terms included *based*, *time*, *thinking*, and *easy*, suggesting that process-based assessments, particularly essays, are viewed as especially prone to dishonest practices.

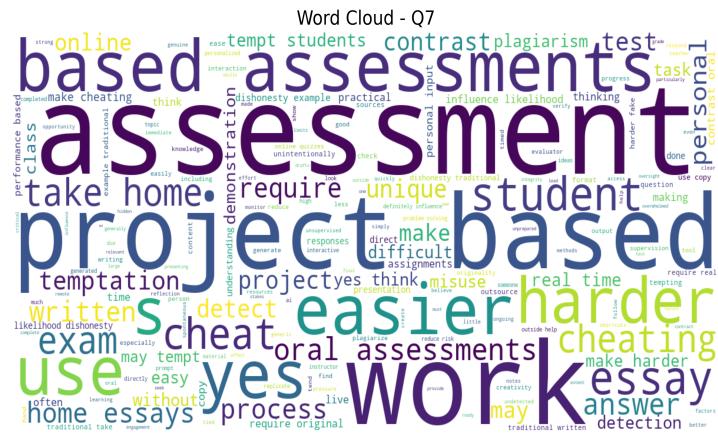


Fig. 14. Word cloud of terms describing academic dishonesty in assessments.

Many students highlighted that essays are “easy to cheat” because of their reliance on written output, which can be copied, outsourced, or generated using AI tools. In contrast, keywords such as *project* and *harder* point to the perception that performance-based assessments offer more resistance to dishonesty. These tasks typically require original contributions, collaboration, and tangible outputs, making them harder to fabricate. Oral formats were also identified as relatively secure due to their interactive and immediate nature, where real-time engagement makes dishonesty more difficult.

4.8.2 Comparative Vulnerability Across Assessment Types. To quantify these perceptions, Figure 15 presents a heatmap comparing how students evaluated the vulnerability of oral, process-based, and performance-based assessments to academic dishonesty.

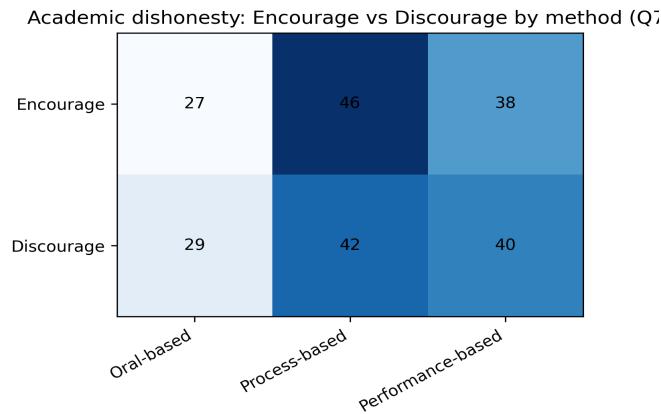


Fig. 15. Comparative vulnerability of assessment formats to academic dishonesty.

The results confirm that process-based assessments are perceived as most vulnerable, with 46 students suggesting they encourage dishonesty and 42 suggesting they do not. Performance-based tasks were viewed as relatively balanced, with slightly more students indicating they discourage dishonesty (40) than encourage it (38). Oral assessments were rated as the least vulnerable, showing a nearly even split (27 encourage vs. 29 discourage). Together, these findings underscore the challenge of safeguarding integrity in traditional essay formats while suggesting that applied, interactive, and demonstrative methods—particularly oral and performance-based assessments—are better positioned to maintain academic honesty in an AI-driven era.

4.9 Ideal Assessment Formats

4.9.1 Themes Emerging from Student Responses. Students’ reflections on ideal assessment formats reveal a strong preference for methods that emphasize authenticity, communication, and applied skills. As shown in Figure 16, the most frequent terms included *project*, *oral*, *presentation*, and *skills*.

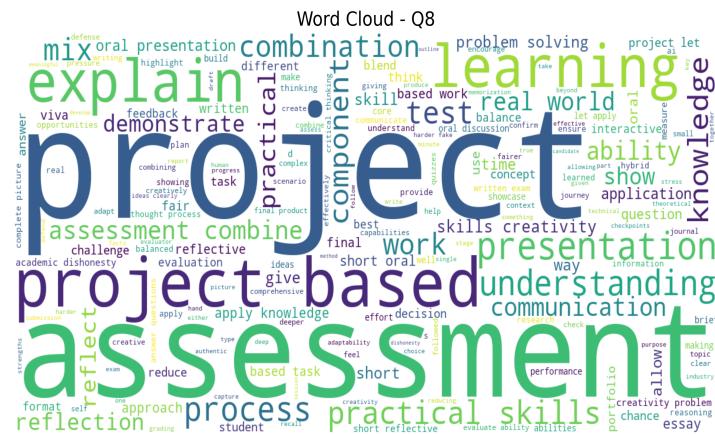


Fig. 16. Word cloud of terms used to describe ideal assessment formats.

These keywords suggest that students envision their ideal assessments as multifaceted and dynamic, blending performance-based and oral components. The appearance of terms such as *short* and *understanding* points to a desire for efficiency and clarity, where assessments test knowledge meaningfully without excessive time burdens. Similarly, mentions of *viva* and *practical* reinforce the emphasis on authenticity and real-world relevance, reflecting a collective call for assessments that mirror professional and applied contexts rather than abstract exercises.

4.9.2 Distribution of Preferred Ideal Formats. The distribution of students' ideal format choices is shown in Figure 17, combining responses into hybrid and individual assessment categories.

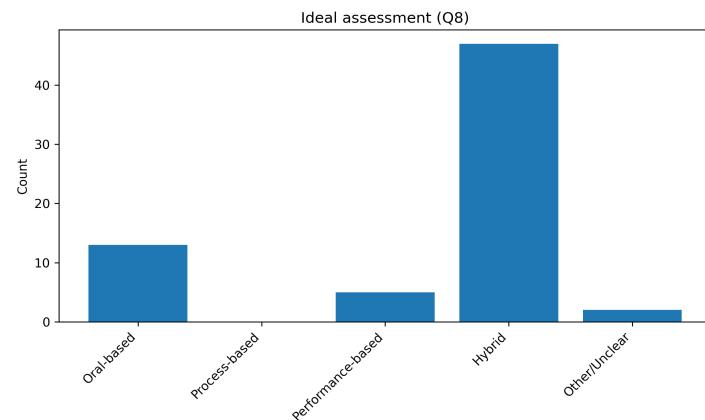


Fig. 17. Distribution of ideal assessment formats.

The results highlight that nearly half of the respondents (47) preferred hybrid models that integrate both performance-based and oral components. This dominant trend underscores students' desire for balanced evaluations that allow them to demonstrate practical competencies while also articulating and defending their

ideas. Interestingly, no student selected purely process-based formats, reflecting widespread discontent with essays and journals due to inefficiency and their vulnerability to AI-enabled dishonesty. Smaller groups favored exclusively oral (13) or performance-based (5) formats, but the overarching signal is clear: students want diverse, authentic, and student-centered assessments. Collectively, these findings reinforce the need to rethink essays as the dominant mode of evaluation and move toward hybrid models that align more closely with real-world demands and skills.

5 Discussion

5.1 Response to the Research Question

The guiding research question for this study was: “*How do students prefer to be assessed in higher education: oral, process-based, or performance-based assessments?*” The results reveal a clear preference for **performance-based assessments** (e.g., projects, portfolios, applied tasks), followed by **oral assessments** (e.g., presentations, vivas). Students consistently perceived these formats as more authentic, engaging, and effective in demonstrating knowledge, while also being less vulnerable to academic dishonesty.

In contrast, **process-based assessments**, such as essays and journals, were the least favored. Essays were frequently criticized as repetitive, time-consuming, and increasingly undermined by plagiarism and AI-generated text. Collectively, the findings point to a significant shift in student preferences toward **authentic, practical, and interactive assessment methods**, underscoring the need for higher education institutions to reconsider their continued reliance on essay-based evaluation models.

5.2 Implications

The results carry several important implications for higher education assessment practices. First, they suggest the need to **move beyond traditional essay-based evaluations** toward performance-based and oral assessments, which students perceive as more authentic, motivating, and reflective of real-world competencies. By emphasizing application, creativity, and communication, these methods also reduce risks associated with plagiarism and AI misuse, thereby strengthening academic integrity.

Second, performance-based and oral formats may improve student engagement by fostering collaboration, critical thinking, and applied knowledge. However, thoughtful design remains essential. Oral examinations should be structured to minimize anxiety while still ensuring rigor, and performance-based projects should include reflective components to fairly evaluate both individual and group contributions.

Finally, the findings provide concrete guidance for curriculum designers and policymakers. Assessment frameworks should be reoriented toward **21st-century learning goals**, prioritizing authenticity, adaptability, and practical skill development over rote essay writing. This shift would not only align with student preferences but also prepare learners for professional environments where communication, creativity, and applied problem-solving are valued.

5.3 Limitations

This study is subject to several limitations. The sample size was relatively small and limited to a single institutional context, which restricts the generalizability of findings. Furthermore, the study relied on self-reported qualitative responses, which may be shaped by social desirability bias or incomplete recall. Cultural and disciplinary differences—factors that strongly influence assessment preferences—were also not systematically captured.

From a technical standpoint, limitations arose from the use of standard natural language processing (NLP) tools for analyzing short open-text responses. These tools struggled to capture nuanced meanings, particularly when responses were brief, vague, or context-dependent. Consequently, while clustering and thematic extraction offered valuable insights, the analysis did not fully capture the depth and subtlety of student perspectives.

5.4 Future Work

Future research should address these limitations in several ways. Methodologically, larger and more diverse samples across institutions and disciplines are necessary to improve generalizability. Employing **mixed-methods approaches**—combining surveys, interviews, and performance data—would provide a more comprehensive understanding of student assessment preferences.

Technically, future studies could leverage **large language models (LLMs)** for text analysis. Unlike traditional NLP methods, LLMs offer advanced semantic understanding and contextual reasoning, enabling richer identification of themes and sentiments in student responses. A hybrid approach, integrating human coding with LLM-assisted analysis, may balance interpretability with scalability, offering more reliable insights.

Overall, extending this work will help refine student-centered, AI-resilient assessment models that preserve integrity while enhancing learning outcomes.

5.5 Conclusion

This study examined how students wish to be assessed in higher education by analyzing preferences, experiences, and challenges associated with oral, process-based, and performance-based assessment formats. Across survey responses and qualitative analysis, a clear trend emerged: students overwhelmingly favored assessment methods that emphasize authenticity, application, and interaction over traditional essay-based approaches.

Performance-based assessments, such as projects and portfolios, were consistently rated most positively. Students highlighted their practicality, opportunities for creativity, and real-world relevance. These formats were seen as the most effective at fostering collaboration, applying knowledge, and cultivating transferable skills. Oral assessments, including vivas and presentations, were also highly valued, particularly for their ability to test communication, build confidence, and discourage academic dishonesty. However, they also carried challenges, with many students reporting stress and performance anxiety. This dual perception indicates that oral methods are best utilized as complementary elements in hybrid designs, reinforcing authenticity while requiring careful structuring to minimize undue stress.

By contrast, process-based assessments, particularly essays, received the least support. While a minority of students recognized their value for reflection and structured thinking, the majority associated them with time pressure, stress, and vulnerability to plagiarism or AI-enabled dishonesty. This declining credibility suggests that essays should no longer occupy the central role in higher education assessment.

Taken together, the findings point toward a student-driven rethinking of assessment design. The dominant preference was for hybrid models that combine performance-based projects with oral components, striking a balance between applied skills and communicative competence. Such formats not only align with student expectations but also respond to wider educational challenges of ensuring fairness, integrity, and real-world preparation.

In conclusion, students envision assessment not as a test of rote reproduction, but as a multidimensional process of demonstrating knowledge through practice, communication, and creativity. The future of higher education assessment lies in moving beyond the essay, embracing authentic, skill-oriented, and hybrid methods that prepare learners for both academic and professional success. s

References

- [1] Anonymous. 2025. AI-Based Digital Cheating At University, and the Case for New Ethical Pedagogies. *Journal of Academic Ethics* (2025), 1–20.
- [2] Adrian Furnham, Andrew N. Christopher, Jeanette Garwood, and Neil Martin. 2008. How would you like to be evaluated? The correlates of students' preferences for assessment methods. *Educational Psychology* 28, 6 (2008), 1–21.
- [3] L. Furze. 2024. Rethinking assessment for generative AI: Beyond the essay. *Rethinking Assessment* (2024).

- [4] R. Gao, H. E. Merzdorf, S. Anwar, M. C. Hipwell, and A. Srinivasa. 2023. Automatic assessment of text-based responses in post-secondary education: A systematic review. *arXiv preprint arXiv:2308.16151* (2023).
- [5] M. García et al. 2024. Rethinking educational assessment in the age of generative AI: Actionable strategies to mitigate academic dishonesty. *Preprint* (2024).
- [6] S. Herbold, A. Hautli-Janisz, U. Heuer, Z. Kikteva, and A. Trautsch. 2023. AI, write an essay for me: A large-scale comparison of human-written versus ChatGPT-generated essays. *arXiv preprint arXiv:2304.14276* (2023).
- [7] A. Kirwan. 2024. ChatGPT and university teaching, learning and assessment: some initial reflections on teaching academic integrity in the age of Large Language Models. *Irish Educational Studies* 43, 4 (2024), 1389–1406.
- [8] S. J. Newell. 2023. Employing the interactive oral to mitigate threats to academic integrity from ChatGPT. *Scholarship of Teaching and Learning in Psychology* (2023).
- [9] K. Obed, W. A. L. Anangisye, and P. Sanga. 2025. Academic integrity considerations of using ChatGPT in assessment activities among university student teachers. *Quality Assurance in Education* 33, 2 (2025), 305–320.
- [10] T. Susnjak and T. R. McIntosh. 2024. ChatGPT: The End of Online Exam Integrity? *Education Sciences* 14, 6 (2024), 656.
- [11] G van de Watering, D Gijbels, F Dochy, and P Van den Bossche. 2008. Students' assessment preferences, perceptions of assessment and their relationships to study results. *Higher Education* 55 (2008), 1–17.
- [12] Tal Waltzer, Celeste Pilegard, and Gail D. Heyman. 2024. Can you spot the bot? Identifying AI-generated writing in college essays. *International Journal for Educational Integrity* 20, 11 (2024), 1–15.