Generative AI tools and assessment: Guidelines of the world's top-ranking universities

Benjamin Luke Moorhouse [a],[*], Marie Alina Yeo [b], Yuwei Wan [a]

[a] *Department of Education Studies, Hong Kong Baptist University, Hong Kong, China*
[b] *Senior Language Specialist, SEAMEO RELC, Singapore*

ARTICLE INFO

ABSTRACT

The public release of generative artificial intelligence (GAI) tools (e.g., ChatGPT) has had a disruptive effect on the assessment practices of higher education institutions (HEIs) worldwide. Concerns have largely been associated with academic integrity, cheating and plagiarism. HEIs have had to develop guidelines in response to GAI. As many of these guidelines were developed in haste and could affect a large number of instructors and students, there is a need to examine their content, coverage and suitability. This review examines the extent to which the world's 50 top-ranking HEIs have developed or modified their assessment guidelines to address GAI use and, where guidelines exist, the primary content and advice given to guide instructors in their GAI assessment design and practices. The findings show that just under half of the institutions have developed publicly available guidelines. The guidelines cover three main areas: academic integrity, advice on assessment design and communicating with students. Amongst the suggestions for teachers on assessment design, two appear particularly pertinent in helping develop effective assessment tasks and developing learners' AI literacy: first, running assessment tasks through GAI to check the extent to which the tool can accomplish the task and, second, having students use GAI as part of the assessment process. Overall, the review suggests that HEIs have come to accept the use of GAI and drafted assessment guidelines to advise instructors on its use. In the article, we argue that it may be beneficial to embrace GAI as a part of the assessment process since this is the reality of today's educational and job landscape. This will require instructors to develop a new competence - generative artificial intelligence assessment literacy - which is conceptualised in this article.

1. Introduction

The growth of Artificial Intelligence (AI), particularly generative AI (GAI) such as ChatGPT and Bard, has intensified calls for guidance of its use in higher education institutions (HEIs) worldwide [1,2]. This is in large part due to the ability of these tools to generate human-like texts that are difficult to detect even by experts [3], and the incredibly fast adoption of these technologies by students. Indeed, *The Times Newspaper* reported that nearly half of the students at Cambridge University in the United Kingdom admitted to using ChatGPT in their studies [4] while *Forbes* magazine claimed that one in five, or 20 % of college students in the United States, admitted the same [5]. These figures are likely to increase as GAI functionality becomes embedded in word processors and presentation software (e.g., Microsoft Co-Pilot). GAI represents a dramatic advancement from previous AI models. GAI leverages deep learning models to generate human-like content, including audio, code,

images, text, simulations, 3D objects and videos. These tools can create unexpected outputs in response to varied and complex prompts (e.g., languages, instructions, questions) [6].

A key concern of HEIs is that students may use GAI to cheat or plagiarise their written assignments and exams [7]. This could undermine academic integrity and ultimately the reputation of HEIs. In addition, scholars have argued that students could become over-reliant on GAI, leading to a decline in their writing and critical thinking skills [8] and negatively affect the quality of education and student learning outcomes [9].

However, other scholars have argued that GAI can bring benefits to education (e.g., [10,11]). For example, Hwang and Chen [12] provide six possible roles for GAI in education: teacher/tutor; student/tutee; learning peer/ partner; domain expert; administrator; and learning tool. Indeed, traditional (non-generative) AI tools have been found to have a number of positive utilities in higher education. Crompton and Burke

---

[13] in their systematic review of AI use in HEIs from 2016 to 2022 found various pedagogical benefits of using AI. These included providing more personalised feedback on assignments, helping instructors generate questions and tests (e.g., [14,15]), accurately predicting academic performance (e.g. [16]), and providing tailored and personal academic assistance (e.g. [17]). Similarly, Celik et al. [18] found that AI can help teachers define and support students' needs, provide immediate student feedback, and conduct automatic essay scoring.

Despite the proven and potential benefits of AI in education, the disruptive nature of the general release of GAI technologies to common assessment practices has led to widespread calls for HEIs to develop comprehensive guidelines pertaining to their use (e.g. [19,20]). For example, Rudolph et al. [19] suggests the HEIs, "develop policies and clear, easy-to-understand guidelines for the use of language models in learning and teaching - the guidelines should include information on the proper use of these tools and the consequence of cheating" (p.356).

Early reports on the reactions of HEIs to GAI suggested a range of responses to their use. Some universities banned the use of GAI tools in their academic programmes, with the use of AI seen as cheating [7,21], while others have guidelines allowing their use, as long as that use is declared and acknowledged [22]. Clearly, the development of GAI technologies has complicated the HEI assessment landscape and blurred the line between acceptable and unacceptable practices. The lack of clarity could be troubling for instructors and learners, who look to institutional guidelines to guide their practices [11,23]. state, "A culture of academic integrity in institutions is enhanced if policies and procedures are carefully written with all key stakeholders in mind" (p. 6).

Although HEIs were initially "caught off guard" by the development of GAI [20], an increasing number of them have begun to develop guidelines pertaining to their use in teaching, learning and assessment. We believe there can be much to learn from different HEIs' responses to GAI, allowing for a collective shaping of clear and consistent guidelines regarding GAI, academic integrity and assessments.

Within this rapidly advancing context, this review article aims, firstly, to identify the extent to which the world's top ranking HEIs have developed or modified guidelines specifically to address the complexities posed by the use of GAI tools in assessment tasks. Secondly, where guidelines exist, we seek to identify the primary content and advice contained within them that can help guide instructors regarding GAI and assessment. In this article, we define guidelines as non-mandatory recommendations, interpretations, administrative instructions, best practice guidance, or frameworks in which to operate. Policies, on the other hand, are mandatory, prescriptive and have far-reaching aims (e.g. institutional reputation, legal compliance, operational efficiency) [24], so they generally take a long time to develop and be approved. Based on this, it is likely that most of the examples from the websites we surveyed may, in effect, be guidelines, even if they are referred to as policies.

We explicitly focus on GAI guidelines, as they offer tangible examples of advice and suggestions that can be adopted by instructors in HEIs. At the same time, we hope that the findings of this paper offer practical suggestions for designing assessment in higher education in the GAI world. The review serves as a reference for HEIs in the process of designing or redesigning their guidelines around academic integrity and assessments.

1.1. Academic integrity and dishonesty in the GAI world

As previously stated, the primary concerns regarding AI and GAI and assessments in HEI have been the issue of academic integrity and academic dishonesty [2,7,19]. Various institutions and government agencies have policies regarding academic integrity and provide guidance on the implications of academic dishonesty. In order to operationalize these policies, clear definitions are needed. For example, The Australian government's Tertiary Education Quality and Standards Agency (TEQSA) [25] provides the following definition of academic

integrity as "the expectation that teachers, students, researchers and all members of the academic community act with: honesty, trust, fairness, respect and responsibility". Academic dishonesty, in contrast, refers to "a range of behaviours involving intentional violation of academic rules for personal gain, such as plagiarism, lying, and falsifications'' ([26], p. 100,054). TEQSA [25] further lists plagiarism, recycling or resubmitting work, fabricating information, collusion, exam cheating, contract cheating and impersonation as examples of academic dishonesty.

Within the literature, plagiarism appears to be the most frequent form of academic dishonesty [27]. Yet plagiarism itself is a multifaceted concept with diverse interpretations and definitions. One form of plagiarism is "textual plagiarism", which refers to "the use of words and/or ideas from another source, without proper attribution," and this may be intentional "prototypical plagiarism" or unintentional "patch-writing" ([28], p. 276). Another act of plagiarism is to list sources in the reference list that the student has not actually read [26]. Based on these conceptualizations of plagiarism, deliberately using GAI to construct assessment responses that include citations and references would still constitute textual plagiarism [11]. In addition to plagiarism, students' use of GAI to complete assignments might also be deemed as collusion, contract cheating, and impersonation. Although it is common for academic integrity policies to include clauses pertaining to these aspects, often they refer to collusion with "other students" or "commissioning to someone else" (e.g., [25]). These clauses were developed before the emergence of GAI and may require HEIs to reflect this change in any subsequent guidelines. GAI likely opens up the possibility of contract cheating to vastly more students [29].

To promote academic integrity and deter academic dishonesty, HEIs have generally taken a two-pronged approach: detection and prevention. Plagiarism detection software such as iThenticate and Turnitin are commonly used to "check" students' assessment submissions. How well these tools can detect original texts generated by AI remains questionable [30]. Some tools such as GPTZero, Turnitin, ZeroGPT and Winston. AI claim to be able to detect text generated by GAI. Despite these claims, many scholars have questioned the accuracy of these tools (e.g. [31]). Kohnke et al. [11] note that the use of these tools may "lead to a game of cat and mouse" (p.544) with GAI developers and detection tools constantly coming up with ways to out-pace each other, leaving users with no assurance of reliable detection.

1.2. Assessment design and academic integrity in the GAI world

Assessment design and processes have been another area of debate since the development of GAI [30]. These debates have focused on the integrity of the assessment process and the qualities of effective assessment may be compromised if students use GAIs to complete their assessment tasks. For example, Yeo [32] suggested that the validity, reliability, and fairness of the assessment process may be compromised if students are assessed based on work written partly or wholly by GAI tools. As an illustration, if instructors assign a written essay to test the writing sub-skills of topic development, organisation and coherence, lexical resource and grammatical range and accuracy, it is with the expectation that students will construct and produce a response on their own so that the instructors can assess them against the assessment criteria. However, if students use a GAI tool to complete the task, they will not be demonstrating these skills but may instead be using other skills, such as the ability to write initial and further prompts to guide the tool to produce and refine the generated output. In addition, some students may have access to specialist GAI tools, or more advanced tools (e. g. Chatgpt 3.5 vs GPT-4), and others do not. This can affect the reliability and fairness of an assessment as students do not have equal access to resources that can help them to succeed in the assessment [33].

An initial response to the above issues and risks by many HEIs was to ban the use of GAI [34] or control the environment in which students did their assignments and to return to proctored pen-and-paper tests [21]. Such measures were supported by evidence that the environment where

students do assignments, at home or in a controlled classroom environment, has an impact on academic integrity. For example, Yuzici et al. [35] noted, "The more the students work out-of-class or perform take-home exams and assessment, the greater the risk to their academic honesty" (p. 222). Janke's [26] study found that students were more likely to commit acts of academic dishonesty for assessments not done under proctored, timed conditions.

However, due to the known limitations of in-class assessments (e.g. limited scope, test-taking anxiety, lack of differentiation, decontextualization, irrelevance to today's workplace) [36], there has been a move away from in-class assessments towards take-home assessments in recent years [28]. Although this has opened up the kinds of assignment tasks instructors can assign, with various recommendations proposed to increase the authenticity of assignments and diversity of assignment tasks (e.g. portfolios, case studies) [37,38], there is still a dominance of written essays in many HEI programmes and courses. Common take-home assessments include responding to questions, conducting a literature review, writing an argumentative or analytical essay, producing presentation slides, or writing research reports [39]. These kinds of writing tasks can be accomplished by GAI tools [3]. Indeed, Cotton et al. [29] demonstrated the ability of ChatGPT to construct an academic text by publishing the results as an article in the journal, *Innovations in Education and Teaching International* with a discussion on the implications of the tool for academic integrity. Similarly, King [40] published a conversation he had with ChatGPT as an editorial in *Cellular and Molecular Bioengineering* to highlight the capabilities of the tool.

Scholars have suggested then that we need to reconsider the kinds of assessment tasks instructors assign to make them 'AI-resistant' by reducing the likelihood that GAI can complete the whole assignment task. For example, Rudolph et al. [19] suggest that students can be tasked with analysing videos or images, analysising in-class discussions, analysising long texts that do not fit in a prompt, writing about recent events that are not in the training data of GAI tools and writing about topics that are highly specific, niche or personal to the students. While there is an increasing amount of advice available to instructors on how to modify their assignment tasks in the GAI world (e.g. blogs, newsletters), many instructors will look to their institutions for guidance and direction regarding GAI [11]. Therefore, it is important to understand the kinds of suggestions HEIs are providing to their instructors.

2. Methodology

This review seeks to investigate the extent to which HEIs have developed guidelines in response to the emergence of GAI tools and identify the components of such policies. The following research questions guided the study:

1. To what extent have the world's top-ranking HEIs developed or modified assessment guidelines specifically to address the use of GAI tools in assessment tasks?
2. Where guidelines exist, what is the primary content of the guidelines and the advice contained within the guidelines that can help guide instructors regarding GAI and assessment?

To ensure we selected guidelines in a systematic and efficient way, we decided to use the Times Higher Education (THE) World University Rankings 2023 (https://www.timeshighereducation.com/world-univer sity-rankings) to identify universities for inclusion. We decided to include the top 50 ranked universities in the review. While this index is well regarded, we recognise using the THE does present limitations. First, the rankings tend to favour English-speaking countries in the Global North. Second, the guidelines may not be representative of all universities globally. However, despite these limitations the selection criteria allowed us to conduct the review rapidly and systematically. Below the process of conducting the review is presented.

2.1. Locating GAI guidelines

To locate the GAI guideline documents in each university's website, the researchers visited all the top 50 universities' official websites and conducted manual searches. The initial keywords used for the search included "AI policy", "Generative AI policy", "ChatGPT policy", "AI guidelines", "Generative AI guidelines'', "ChatGPT guidelines", "AI guide", "Generative AI guide", and "ChatGPT guide". Additionally, some related terms, such as "academic integrity", "academic honesty" and "plagiarism", were identified and incorporated into the search to ensure comprehensive coverage. After conducting the initial keyword search, we also took additional steps to ensure that no relevant information was missed. We visited specific sections or columns on each university's official website that could potentially contain GAI guidelines (e.g. the websites of teaching and learning centres). Of the 50 universities, 30 were found to have guidelines related to generative AI on their official websites (60 %). The websites were extracted for analysis. The search and extraction were conducted on 15th June 2023. Only publicly available documents or websites were included due to access restrictions.

2.2. Applying inclusion criteria

To ensure the websites were relevant to our review aims, we applied the following inclusion criteria:

- The audience must be instructors (exclude guidelines for researchers and academic publishing)
- The genre must be guidelines (exclude blogs, notes, memos and news)
- The guidelines must be issued at the university level (exclude faculty, school, and department level)

After applying the inclusion criteria, seven universities were excluded. The remaining 23 university guidelines are included in the review (See Appendix for the list of universities and corresponding websites included in the review. Please note that the websites are active, and content may have changed since they were extracted for this study). The study consists of policies or guidelines from 15 universities located in the USA; 4 in the UK; 2 in Canada; and 1 in Japan and Australia respectively.

2.3. Conducting the review

Due to the rapid nature of the development of the guidelines, it was important to conduct two levels of analysis: context level and content level.

2.3.1. Context-level

At the context level, data pertaining to the issuer of the guidelines (Table 1); issue or version date (Table 2); and tools (Table 3) are mentioned. It can be seen from Table 1 that the primary issuer of guidelines pertaining to the use of GAI in assessments is university centres for teaching and learning (Note that University of Toronto (UOFT), University of California, Berkeley (UCB) and. University of

Table 1

Issuer of guidelines pertaining to Assessment and AI.

issuer	quantity
Centre for teaching, learning or innovation	15
Office of Provost	3
office of ethics/academic integrity/community standard	2
Centre for technology	2
Senate	1
Not stated	3

Table 2

Issue month of guidelines pertaining to Assessment and AI.

Issue month	Quantity (HEI)
Without specific date	15
January 2023	1 (PU)
February 2023	2 (SU, UCL)
March 2023	2 (UCB, ICL)
April 2023	1 (UTokyo)
May 2023	1 (UOFT)
June 2023	1 (Edin.)

Table 3

GAI tools mentioned in the guidelines pertaining to Assessment and AI.

Tools mentioned	Quantity
Only ChatGPT	13
Not only ChatGPT (also include other tools, e.g. DALL-E, CoPilot, etc.)	10

British Columbia (UBC) had two issuers of guidelines). Interestingly, when we compared the scope and content of the documents issued by different issuers (e.g., senate, provost, centres for teaching and learning), we did not see any clear differences. The issuer may therefore reflect the university's decision-making and communication process and structure rather than a distinction between policies and guidelines. As Table 2 shows, most guidelines do not have a specific date of issue with one or two issued each month between January 2023 and June 2023. Table 3 shows that more than half of the guidelines (57 %) only mention ChatGPT and not other types of GAI tools. Many of the others mention other tools, but use ChatGPT as the example tool throughout the guidelines.

2.3.2. Content-level

At the content level, the process was more complex. First, two of the researchers read five websites independently and coded the content inductively. They recorded their codes and used them to create a framework for analysis. The two researchers met and discussed their initial coding and framework. They found consistency in their process and coding and discussed some minor discrepancies. They created a single framework comprising of nine fields: authorship; acknowledgement; plagiarism; advice on assessment design and tasks; detection of GAI use; responsible agent; proper use; improper use; and communication with students. Themes within fields and guiding questions were also
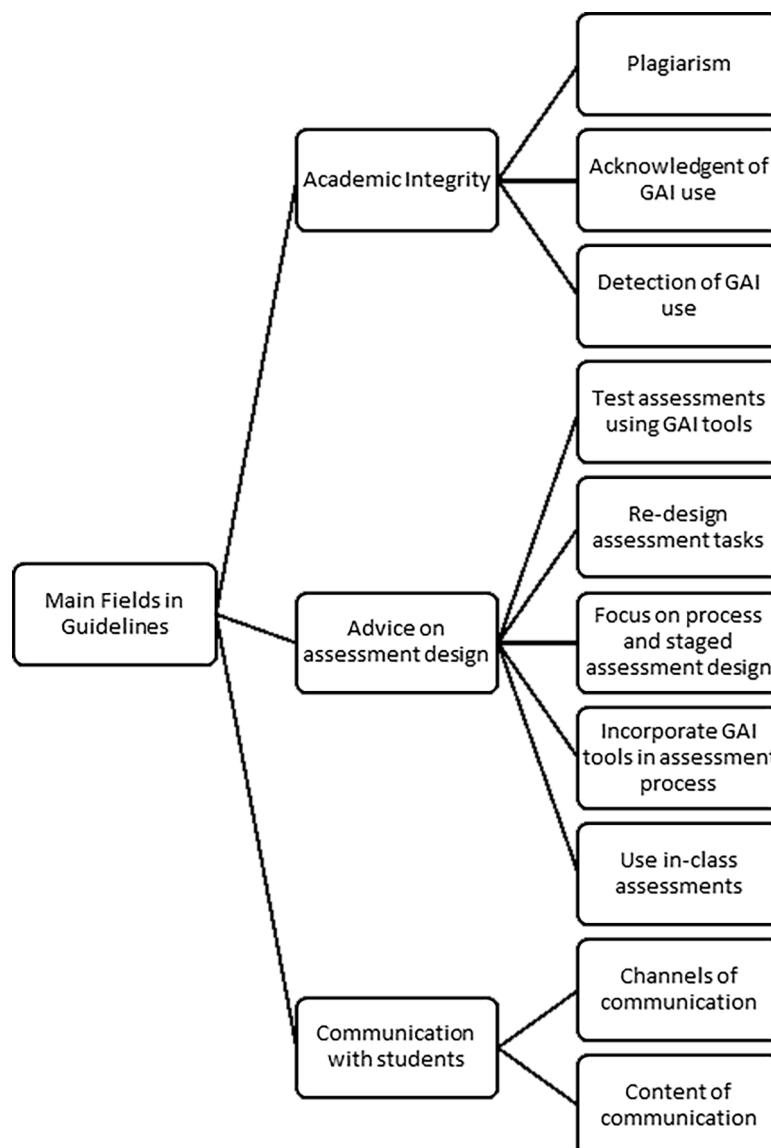
Fig. 1. Main Fields in GAI Guidelines.

included to aid the review. The two researchers then independently sorted, coded and extracted both quantitative (if the field was mentioned in the website) and qualitative information (what the guidelines included pertaining to the field) from the remaining websites. Additional fields or themes were included if content did not match the nine fields above. After the researchers completed the process, they met to discuss the outcomes. Each website was discussed one by one, with any discrepancies discussed until agreement was reached.

At this stage, the quantitative data was collocated and subjected to standard descriptive statistical analysis (quantities and percentages). The qualitative data was read and re-read again with data extracts illustrative of the themes within the fields identified for inclusion in this paper. In summary, our review of the GAI guidelines on the websites of 23 universities listed amongst the 50 top universities in the THE Index (2023) yielded three main areas of coverage, which can further be categorised as shown in Fig. 1. The findings regarding each area are presented in detail in the next section.

3. Findings

3.1. Academic integrity

Given the nature of the guidelines, they all included advice on academic integrity and GAI in assessments. However, not all the guidelines provided the same advice, nor did they include the same information. Three main fields were identified: plagiarism, acknowledgement of GAI and detection of GAI use. Below, data pertaining to each field are presented (See Table 4 for the themes related to academic integrity and HEI guidelines that exhibit the themes).

3.1.1. Plagiarism

Sixty percent of the university guidelines addressed plagiarism or academic misconduct in their GAI guidelines. The analysis of the documents identified three main themes: (1) Forms of plagiarism using GAI; (2) Ways to address plagiarism; and (3) Instructors' responses to plagiarism.

Eight university guidelines identified three forms of plagiarism using GAI. These were copying and pasting AI-generated response, running material through multiple AI generators to avoid detection, and inadequate documentation of the use of GAI in assessment tasks. UCB explicitly stated that "Lifting full sentences and paragraphs wholecloth, whether it's from an encyclopaedia, written article, or AI-generated text creation tool, is considered plagiarism." Similarly, Imperial College London (ICL) suggested that submitting assessments generated by AI as if it was students' own is plagiarism. University of California, Los Angeles (UCLA) guidelines suggested that to avoid being caught by plagiarism detection tools, some students may run their work through

Table 4
Themes related to academic integrity and HEIs HEI guidelines that exhibit the themes.

Themes	Sub-themes	Institutions
Plagiarism	Forms of plagiarism using GAI	UCB; ICL; UCLA UCL; Duke; NU; Edin.; UToyko
	Ways to address plagiarism	Caltech; UCLA; Duke; UIUC
	Instructors' responses to plagiarism	UPenn; NU; UCL
Acknowledgement of GAI	Acknowledge the use of GAI	SU; PU; UCB; UPenn; UOFT; UCL; Edin.; MU
	Cite AI-generated content	DU; CMU; Edin.; UBC; UIUC; UOFT
Detection of GAI use	Discourage the use of GAI detection tools	ICL; UOFT; UCB; YU; Cornell; DU Edin.; UTokyo; MU
	Try GAI detection tools	UPenn; NU; UBC; CMU; UIUC

several GAI paraphrasing tools (e.g., Quilbot). UCLA reminded its instructors to be mindful of such behaviour. In addition, the guidelines of University College London (UCL) and Duke University (DU) suggested that students may use AI-generated content as a source of information or inspiration but if they failed to properly cite or acknowledge the source, this behaviour would be considered as plagiarism.

To help students better understand the importance of academic integrity and the consequences of violating the university's plagiarism policy, UCLA and University of Illinois at Urbana-Champaign (UIUC) suggested instructors incorporate relevant information into their course syllabus. Some guidelines provided templates or examples of the policy language. For example, DU provided the following sample; "All work submitted in this course must be your own. Contributions from anyone or anything else—including AI sources—must be properly quoted and cited every time they are used".

When students are suspected of plagiarism, three universities mentioned how instructors should respond to such behaviour. Both University of Pennsylvania (UPenn) and Northwestern University (NU) suggested that instructors should refer to university's standard of academic integrity, but Upenn also emphasised that instructors have the discretion to decide whether a violation occurred or not. UCL recommended that if instructors have concerns about a student's work potentially containing AI-generated content without proper documentation, students should be invited to "an investigatory viva to probe the authorship" of their work.

3.1.2. Acknowledgement of GAI

More than half (57 %) of the universities provided guidelines on proper acknowledgement of GAI in students' assignments. The two main ways suggested are (1) acknowledging the use of GAI and (2) citing AI-generated content. Eight universities mentioned acknowledging GAI use in their policies or guidelines. Four of them only emphasised the importance of disclosing the use of GAI assistance without providing further guidance on how to properly document and attribute the AI-generated content. The remaining four universities offered more comprehensive guidelines for acknowledging GAI use, including mentioning the specific AI tool used and the date it was accessed (University of Edinburgh [Edin.], UOFT), documenting the process of using GAI tools (UCL, Monash University [MU]), and providing an appendix (UOFT). UOFT stated that the appendix for each assignment should include "what tools were used, how they were used, and how the results from the GAI were incorporated into the submitted work". Similarly, UCL pointed out that students must describe the prompts used, the output, and how the output generated by GAI was changed by students.

Six universities indicated how to cite or give credit to content generated by GAI following source citation academic conventions including in-text citation and including the GAI tool in the reference list. For the reference style, half of them recommended students use standard citation formats (e.g., MLA, APA, Chicago), as these organisations have already provided information on how to cite GAI. Given the interactive nature of GAI tools, Edin. suggested citing AI-generated content as "personal communication", typically using an in-text citation only. UIUC also recommended instructors consider different ways of citations for different uses, such as "ideation, inspiration, or brainstorming". There seems to be no standard way within HE for students to currently acknowledge GAI use in assessments.

3.1.3. Detection of GAI use

61 % of the universities mentioned the existence of GAI detection tools (e.g. GPTZero, Turnitin AI detection feature). The majority of the guidelines discouraged instructors from relying on the tools as a way to check if students have followed academic integrity policies. Reasons stated for discouraging their use included: inaccuracy in identifying AI generated content and privacy concerns. For example, ICL and DU reminded their colleagues that because AI detection tools are still in the infant stage, they are "unproven" and "by no means foolproof" in

detecting AI. In addition, MU suggested detection tools cannot "keep pace with the latest developments". Some universities suggested there could be privacy and security issues if instructors input students' work into the external sites (explicitly referencing AI detection tools). UCB suggested this may lead to ethics, privacy, and copyright violations. The guideline of DU cautioned that the use of such tools can "signal that students should not be trusted", therefore, breaking the trust between staff and students.

Five university guidelines suggested instructors can try GAI detection tools. However, these guidelines also point out the importance of being aware of the detection tools' limitations. All the guidelines reminded instructors that these tools should not be the "sole factor" to determine whether students have breached academic integrity (e.g. UBC).

3.2. Advice on assessment design

A majority, seventeen (74 %), of the reviewed universities provided instructors with advice pertaining to the design of assessment tasks in response to the development of GAI tools. There was a recognition in the documents that GAI tools required instructors to examine their current assessment tasks and practices. The advice provided was categorised into five themes: (1) Test assessments on GAI tools; (2) *Re*-design assessment tasks; (3) Focus on process and staged assessment design; (4) Incorporate AI tools in the assessment design and; (5) Use in-class assessments (See Table 5 for themes related to assessment design HEI guidelines that exhibit the themes).

3.2.1. Test assessments using GAI tools

Seven universities suggested instructors test their current and proposed assessment tasks using generative AI tools. When providing this advice, most guidelines specifically mentioned ChatGPT only. The testing of the assessments on GAI tools allows instructors to "assess their capabilities and limitations" (UBC) and "Familiarising yourself with the AI software available in your discipline" (ICL). It also raises instructors' awareness of the types of texts generated by AI so they can recognise it (UPenn). MU provided its instructors with step-by-step instructions on how to do this:

Table 5

Themes related to assessment design and HEI guidelines that exhibit the themes.

Theme	Sub-theme	Institution
Test Assessments Using Generative AI tools		ICL; UPenn; CMU; UTokyo; UBC; MU; UIUC
Re-design assessment tasks	Design assessments that require creativity and critical thinking	PU; ICL; UOFT; CMU; MU; UTA
	Incorporate contextual elements	UCB; UPenn; UW; CMU; UBC; MU; UIUC
	Design authentic assessments	YU; CU; Cornell; MU
	Provide alternative ways for students to represent their knowledge beyond text	YU; UPenn; UOFT; NU; CMU; UTokyo; UBC; PU; MU; UIUC
Focus on process and staged assessment design		Caltech; PU; CU; UPenn; UOFT; Cornell; NU; UW; CMU; MU; UIUC; UTA; UTokyo
Incorporate GAI tools in the assessment process		Caltech; UCB; YU; ICL; CU; UW; UBC; MU; UIUC; UTA
In-class assessments	Recommend the use of in-class assessments	UTokyo; UBC; UIUC; UTA
	Caution against the over-reliance of in-class assessments	CMU; MU

- Paste an entire assignment brief into ChatGPT and see what it produces.
- After reviewing the result, add additional instructions into ChatGPT to try to finetune the results.
- Sometimes, poor results are due to entering insufficient prompts into ChatGPT rather than because ChatGPT cannot produce a satisfactory response to an assignment. Test this by adding further instructions derived from the assessment rubric, marking guide, and taught material. Try a variety of combinations of instructions and see if this improves the results.

As the instructions above show, some documents suggested instructors try out variations of the task instructions and prompts and examine the results. By becoming familiar with the way the tools handle the assessment tasks, instructors can consider how they might re-design tasks that focus on current limitations of the tools. However, Carnegie Mellon University (CMU) guidelines emphasised an important point, that instructors must be cautious about only focusing on current limitations, as "ChatGPT will adapt as it is used. Consequently, designing assignments around any current limitations may be a temporary solution, but not a sustainable one". Additionally, UOFT and University of Texas at Austin (UTA) reminded instructors that there are privacy risks if examination papers or other assessment tasks are inputted into GAI tools. UOFT guidelines stated, "since exam questions are highly confidential documents, you should not, in principle, input them directly into GAI tools". UTA reminded their teaching staff to review the privacy and data collection policies of the tools before using them to test assessment tasks.

3.2.2. Re-design assessment tasks

All seventeen universities suggested different ways instructors can re-design assessment tasks in light of the development of GAI tools. The primary advice focused on (1) designing assessments that require creativity and critical thinking, (2) incorporating contextual elements, (3) designing authentic assessments, and (4) providing alternative ways for students to represent their knowledge beyond text.

The guidelines suggested that instructors design assessments that require creativity and critical thinking as they assume that GAI tools currently struggle with and they cannot "easily replicate" (ICL) these kinds of tasks. UOFT suggested that their colleagues design tasks that:

Ask real questions that stem from current debates in your discipline, and let students know that you expect engaged critical thinking that is appropriate for the level of your students and your discipline. Encourage speculation based on evidence and reasoning, not just compilation of existing information or expression of unsupported personal opinion.

UTA provided a long list of tasks that they suggest AI tools could struggle with. These include,

Give a hug: empathy, collaboration, communication, and leadership skills; Solve a mystery: generating questions and problem finding; and Tell a story: finding what's relevant in a sea of data or applying values, ethics, morals, or aesthetic principles to a situation.

They advised colleagues to consider these when re-designing tasks.

Seven of the universities advised colleagues to incorporate contextual elements into their assignment design. This could include asking "students to connect course content, class conversations, and lived experience" (University of Washington [UW]) and "design essay and exam prompts that require close discussion or analysis of the materials used for your class, including images, video, and other media." (UCB). UIUC suggested instructors "make assignments more personal, reflective, specific, local, based on scenarios/facts/topics covered in class, or that address more complex cognitive skills". The guidelines emphasised that GAI tools can struggle with these kinds of tasks, while students can be more motivated to complete tasks that they find more relevant to their lived experiences (UIUC, CMU). Similarly, authentic assessments, which are designed for students to apply course concepts in real world situations or problems (Columbia University [CU]) are promoted by four

of the universities. MU suggested various kinds of authentic assessments including, "case studies, exhibitions, reflective portfolios, and problem-based inquiries". The MU guidelines suggested that research supports the "role of authentic assessment (when done well) in improving engagement and preventing students from engaging in academic misconduct." They cited Sotiradou et al. [38] to support their advice.

Ten universities suggested that instructors consider providing alternatives and a variety of ways for students to represent their knowledge beyond text. These included drawing images, making slides, facilitating a discussion (Yale University [YU]), submitting slides or presentations, video or audio recordings, or infographics (UPenn); producing portfolios, logbooks or assessment notebooks (UOFT) and using live interviews (MU). Some universities encouraged instructors to give choice to students on how they present evidence of their learning (e.g., YU). While the guidelines suggested instructors consider these kinds of assessments, UBC asked colleagues to be aware that GAI tools are developing rapidly and may be able to create videos and other modes of texts in the near future.

3.2.3. Focus on process and staged assessment design

Nearly all the documents that provided advice on assessment design, 13 of 17, mentioned the need for instructors to put more emphasis in their assessment design on the process of completing an assessment task. This makes students think more about the process of completing a task. A number of ways were suggested for how they might do this:

- Instructors can ask students to "Submit notes they took on sources to prepare their papers or presentations" (UPenn)
- Instructors can add "elements such as proposals, drafts, annotation, or feedback into your assignments." (NU)
- Instructors can "Use more iterative processes of assessment such as student peer review which leads to revisions of the work." (MU)
- Instructors can ask students to provide "a list of specific steps they took, what they could have done differently, and why" (CMU)
- Instructors can scaffold assignment tasks by breaking "big assignments into smaller pieces" (UIUC)

Breaking larger assignments into smaller pieces was a particularly popular suggestion within the guidelines. There seemed to be a few key reasons for this suggestion. Notably, Cornell University (Cornell) suggested that a staged assignment can lower grade anxiety, and therefore reduce the chance students will consider cheating. Similarly, UTA suggests,

Carefully scaffold assignments with time and space for students to complete each step along the way, and consider whether the number of time-intensive tasks might require more bandwidth than students have to spend. Students are more likely to utilize a tool like ChatGPT when they are short on time.

CU emphasised the value of staged assignments in providing information to instructors on student performance, and the opportunity it gives for students to receive formative feedback and peer feedback. In addition to a focus on process and staged assignments, three universities suggested instructors consider more balanced and flexible submission policies. For example, Cornell suggested instructors can offer a set number of late submissions, allow more time for major assignments, and reduce the weight of major assignments. This, they suggested, can reduce anxiety, and reduce the likelihood students will use GAI tools improperly. CMU suggested that "Students may turn to ChatGPT if they are feeling stressed, overwhelmed, unsupported, or out of time" and therefore instructors should help provide a more balanced workload. This included timing assessments to take place outside of exam periods, providing longer deadlines, and building in time in-class for students to work on assignments.

3.2.4. Incorporate GAI tools in the assessment process

Ten of the universities provided suggestions to instructors on how

they might incorporate GAI tools into their assessment designs. Most of these suggest that instructors design an assessment task that requires students to generate responses on tools such as ChatGPT and then critique their responses. For example, the guidelines of YU recommended, "Engaging with ChatGPT as a tool that exists in the world and having students critically engage with what it is able to produce". ICL suggested students can be shown how to use ChatGPT as a formative assessment tool. CU and UW suggested that by designing assessments that require students to use generative AI tools, they can foster digital literacy skills. Importantly, MU reminded instructors that students should acknowledge their use of GAI tools. UBC reminded instructors that there may be ethical and privacy concerns if they require students to use tools in assessment tasks.

3.2.5. Use in-class assessments

Four universities recommended the use of in-class assessments as one approach to mitigate the use of GAI tools in assignments. For example, UIUC advised that, "Do writing assignments in class. It may be useful to have students compose a handwritten essay or write a shorter reflection paper on a current topic in class". Similarly, UBC suggested that,

One approach to mitigating the use of ChatGPT in assignment and assessment design is to incorporate more in-class or otherwise synchronous assignments, either written or oral, or change your current grade weighting to emphasise these.

However, they also acknowledged this strategy may not be feasible for all classes.

On the opposite side, there were words of caution from both MU and CMU that instructors should avoid an over reliance on in-class assessments, such as oral exams and presentations. MU mentioned that these assessments limit the kinds of knowledge students can demonstrate, while CMU reminded their colleagues that these kinds of assessments can disadvantage certain learners. Their guidelines stated that, "one or more of these approaches may appear to be a simple solution, these changes could raise more difficulties than they solve, particularly for reasons of equity and inclusion. For example, some of these approaches may inadvertently disadvantage non-native English speakers or students requiring accommodations for disabilities".

Interestingly, the majority of guidelines did not provide any guidance on in-class assessments.

3.3. Communication with students

Nearly all the guidelines (87 %) provided advice on how instructors can communicate with students about the use of GAI in assessments. They provided advice on: 1) the channels of communication, and 2) the content of the communication (See Table 6 for themes related to

Table 6
Themes related to communicating with students and HEI guidelines that exhibit the themes.

Themes	Sub-themes	Institutions
The channels of communication	Include a statement in syllabi or course outlines	SU; PU; YU; UPenn; UOFT; DU; NU; MU
	Engage students in open discussions around the use of GAI	SU; PU; UPenn; Cornell; DU; CMU; Edin.; UTA
	Collaborate with librarians	UCB; UIUC
2) Content of communication	Set clear expectations	CU; Cornell; UTokyo; UBC
	Partner with students to develop class policies	CU
	Discuss the ethics and limitations of GAI	UCLA; DU; UBC
	Mention the importance of originality	DU
	Highlight the importance of college learning, intellectual struggle, and process	UTokyo

communicating with students and HEI guidelines that exhibit the themes).

3.3.1. Channels of communication

Three primary suggestions are made for how teachers should communicate with students about the use of GAI in assessments. The first is to include a statement in syllabi or course outlines. The second is to engage students in open discussions around the use of AI and the third is to collaborate with librarians. Suggesting instructors inform students of their policies and expectations around the use of GAI in assessments seems common advice in the reviewed guidelines. Some guidelines stated that a statement should be included but do not provide specific examples (e.g. UOFT, MU), however, most guidelines provided examples of the statements instructors can use. For example, PU provided two statements depending on the GAI policy an instructor adopts:

1. Intellectual honesty is vital to an academic community and for my fair evaluation of your work. All work submitted in this course must be your own, completed in accordance with the University's academic regulations. (https://rrr.princeton.edu/2022/students-and-university/24-academic-regulations) You may not engage in unauthorized collaboration or make use of ChatGPT or other AI composition software.
2. Students must obtain permission from me before using AI composition software (like ChatGPT) for any assignments in this course. Using these tools without my permission puts your academic integrity at risk.

Along with these statements, the guidelines encouraged instructors to have open discussions with their students about the use of GAI. These discussions are suggested to be "open-minded" (CMU), and include discussions about "What will they lose intellectually if they use AI to complete their assignments?" (DU), "what ChatGPT or similar AI tools can and cannot do, emphasising those aspects of your assignments that cannot be outsourced" (Cornell) and what is "permissible use in your context" (Edin). UCB and UIIC suggest instructors collaborate with librarians when discussing with students the impact of using GAI in their assessments and the value of the process of completing assignments.

3.3.2. Content of communication

The university guidelines gave advice to their instructors on the content of their communications. These include: (1) setting clear expectations; (2) partnering with students to develop class policy; (3) the ethics and limitations of GAI; (4) Importance of originality; and (5) importance of college learning, intellectual struggle, and process. Due to the complexity of GAI tools and their wide range of capabilities, some guidelines suggested that instructors must make it clear if and how GAI tools can be used. For example, CU stated, "it is important to be explicit with students about the expectations around the usage of ChatGPT and other AI tools in your course. For example, ChatGPT is capable of reading a student's essay and providing meaningful feedback that can then be used by the student to make edits. As the instructor, it's important to be clear about these expectations: can students use the tool for feedback on their own writing? If so, how should they disclose their use? As with all course policies, especially those around academic integrity, it is essential for instructors to be explicit and transparent with their expectations, and to have frank conversations with their students".

CU also suggested that instructors partner with their students to discuss acceptable and unacceptable use of GAI. Cornell provided instructors with three appeals- logical appeal, emotional appeal and personal appeal - to help instructors consider ways to communicate the issues of using GAI in assessments. These are designed to raise students' awareness of the implications of using GAI inappropriately themselves and the community and the value of their degree. Many guidelines emphasised the need to discuss the limitations of AI with students to raise their AI literacy. For instance, UCLA suggested that instructors

should "Facilitate discussions with your students on the impacts of spreading disinformation or biased information, lack of regulation of companies that develop these technologies, and other dangers".

Finally, some of the guidelines suggested that instructors focus on the value of college learning, the intellectual struggle involved in learning and the importance of process in learning. For example, the University of Tokyo (UTokyo) stated, "'Emphasise the importance of the process of finding an answer rather than the final answer itself'. DU suggested, "Instructors can emphasise why original writing (or coding or creativity) matters and what it means to develop your own voice and ideas. Understanding how these skills will help them in careers and further study in your discipline can motivate students to avoid unwanted use of AI." Given the difficulty of detecting GAI content and the versatility in its abilities, it seems necessary that instructors help students realise the implications of their use of such tools in the short, medium and long term for themselves and the wider society.

4. Discussion

In this section, we discuss the findings in relation to our research questions. In response to our first research question, our review has found that 23 of the top 50 ranked universities in the THE World University Rankings 2023 have developed publicly available guidelines for their instructors pertaining to the use of GAI tools in assessment tasks. That less than half of the universities have accessible guidelines is concerning given that research suggests that such guidelines have the potential to raise awareness about academic integrity and reduce academic misconduct [23,41–43]. Furthermore, scholars have been calling for institutions to develop guidelines to support instructors as they adapt to the GAI world (e.g., [1,19]). Without clear guidelines, instructors may take a defensive approach to GAI and adopt more in-class assessments [7] or feel frustrated as they struggle to adapt their assessment practices without institutional guidance ([[44]]). It is therefore essential that HEIs develop GAI guidance and review them regularly as the technology develops.

Addressing our second research question, the guidelines show a concern for the effects GAI tools could have on assessment tasks and provide advice on how to address these concerns. There is an acknowledgement within the guidelines that GAI is here to stay and, therefore, HEIs need to adapt to accommodate it. Our review found that the guidelines included three main areas: the effects of GAI on academic integrity; advice on assessment design; and communication with students (See Tables 4,5, and 6 for the main themes and sub-themes). The primary aims of the guidelines seem to be to help instructors consider what might constitute improper use of GAI, develop procedures for students on how to acknowledge or cite GAI use, re-design their assessments so that it becomes more difficult for students to use GAI solely to complete them and discuss if and how GAI can be used in assignment tasks. The findings can increase instructors' awareness of the risks to their assessment practices and offer suggestions for how to adapt their practices for the GAI world. They also provide a good starting point for instructors and other HEIs who wish to develop or revise their own guidelines. HEIs can see if their guidelines cover the same areas and align with the top universities in the world.

Specifically regarding assessment design, the guidelines advice teachers to design assessments that require creativity and critical thinking, incorporate contextual elements by connecting course content to real-life experiences, implement authentic assessments that allow students to apply concepts in real-world situations, focus on the process and stages of assessment and provide alternative modes of representation beyond text, such as images, slides, discussions, videos, and audio recordings. These suggestions accord with prevailing research and suggestions [7,29,32]. However, the recommendation that teachers should "test" their assessment task by checking how well it can be accomplished solely by GAI and to have students use GAI as part of the assessment process deserve mention. We believe it offers a principled

way of integrating the use of GAI in assessment.

However, questions still remain. The emphasis in the guidelines still seem to focus on limiting or preventing GAI use in assessment tasks. Even when suggestions are made on how instructors can integrate them, the communications with students focus on the negative implications of their use. The goal seems to be to design assessments that can elicit evidence of learning through responses that cannot be fully auto-generated with minimal input from students. A noble goal in the short term. Furthermore, suggestions for proper ways GAI can be used in assessment tasks seem to be limited. This does not align with calls for authenticity in assessment tasks [38,45], which relates to using assessment tasks that resemble the kind of tasks students have to perform in their actual situations, for instance, their current or future jobs [7,20].

In future guidelines, a greater emphasis may need to be placed on how GAI tools can be made integral to assessment tasks and part of students' lived HEI experiences. Allowing or even requiring students to use GAI at various stages of the assessment process would, in fact, enhance the authenticity of assessments. Authenticity is concerned with the extent to which the assessment tasks replicate "real world" tasks [46]. In the same way as editing tools such as Grammarly are embedded in word processing systems, the use of GAI tools is likely to become similarly ubiquitous, hence allowing its use in assessment enhances the quality of authenticity.

Of course, many of the guidelines rightly point out the limitations of these tools and the real concerns their use poses, including bias, inaccuracy, privacy and copyright infringement [19,40,47]. Yet, humans have always used tools that help them complete tasks, and as new tools develop, they change the skills and knowledge we need to do certain tasks [48]. The jobs of the future will require graduates to be GAI literate and part of developing these skills will be using GAI tools in productive, responsible and ethical ways in assessments [49]. Understanding how we might assess GAI use, such as incorporating it into our rubrics, is also something we need to consider in future guidelines.

In addition, most of the guidelines focus on one GAI tool - ChatGPT. This is understandable as it was the first tool to receive widespread attention. However, there are now a plethora of GAI tools with different functionalities. GAI tools have been developed that can generate images, code, videos, and a wide range of texts [6]. By focusing on one tool predominantly, it may leave instructors with a limited view of GAI and its effect on assessments. Indeed, once GAI becomes embedded in common products, such as word processing software, the complexity of considering what are suitable uses of GAI in assessments will only increase. Perhaps, having rational and logical conversations with students about the implications of GAI use, as suggested by Cornell, is the way we need to go. In a sense, adopting trust-based assessment practices that emphasise the value of the task to student learning beyond grades might be more fruitful.

Clearly GAI has created a highly complex context. In order to navigate these new realities, it has been argued that instructors require AI literacies [14]. We argue that the specific nature of assessments means that instructors need to develop GAI assessment literacy. Based on the findings above, we provide a *preliminary conceptualization of GAI assessment literacy* (See Table 7). This conceptualization can help act as a framework in developing GAI assessment literacy.

5. Conclusion

This review has shone light on the guidelines adopted by the top universities around the world pertaining to the use of GAI in assessments. The findings highlight the need for instructors to consider the impact of GAI and develop appropriate assessment practices. As these guidelines have been developed rapidly due to an emerging technology, it is likely that many of the suggestions have not been tested in authentic HEI contexts. As instructors begin to put the suggestions in these guidelines into practice, we will get a better idea of their efficacy in authentic contexts and how assessments will need to be shaped in the

Table 7
A Preliminary Conceptualization of GAI Assessment Literacy.

Able to recognise the implications of GAI for academic and assessment integrity
Able to design assessment tasks that provide space for students to demonstrate learning while incorporating GAI tools in the assessment process
Able to communicate with students about the productive, responsible and ethical use of GAI in assessment tasks

GAI world. We suspect that any GAI guidelines will need to be reviewed and updated regularly as the technology and new assessment and pedagogical models are developed.

As stated earlier, the study has several limitations. First, the review only included publicly available guidelines extracted from the reviewed university websites. The remaining universities may have guidelines that are not accessible publicly. Second, the review focused on the top 50 universities in THE. While this allowed for the review to be done rapidly and systematically, it may not reflect the situation globally. Indeed, the Global North is disproportionately represented. Future studies can explore whether the findings here are consistent with guidelines in other contexts. Third, the review only explored guidelines targeted at instructors. Universities may have developed guidelines for students on GAI use or developed institutional policies governing their use. These guidelines and policies could be reviewed in a future study. Despite these limitations, the findings offer a rich resource for developing and reviewing HEI guidelines and can help inform policymakers, instructors and others on the important content of such guidelines.

5.1. Funding statement

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.caeo.2023.100151.

References

[1] Crawford J, Cowling M, Allen K. Leadership is needed for ethical ChatGPT: character, assessment, and learning using artificial intelligence (AI). J. Univ. Teach. Learn. Pract. 2023;20(3). https://doi.org/10.53761/1.20.3.02.

[2] Holmes W, Porayska-Pomsta K, Holstein K, Sutherland E, Baker T, Shum SB, Santos OC, Rodrigo MT, Cukurova M, Bittencourt II, Koedinger KR. Ethics of AI in education: towards a community-wide framework. Int. J. Artif. Intell. Educ. 2022; 32(3):504–26. https://doi.org/10.1007/s40593-021-00239-1.

[3] Casal JE, Kessler M. Can linguists distinguish between ChatGPT/AI and human writing?: a study of research ethics and academic publishing. Res. Methods Appl. Linguist. 2023;2(3). https://doi.org/10.1016/j.rmal.2023.100068.

[4] Sleator L. Almost half of Cambridge students admit they have used ChatGPT. April 21. The Times; 2023. https://www.thetimes.co.uk/article/cambridge-university-students-chatgpt-ai-degree-2023-rnsv7mw7z.

[5] Nietzel MT. More than half of college students believe using ChatGPT to complete assignments is cheating. March 20. Forbes; 2023. https://www.forbes.com/sites/michaeltnietzel/2023/03/20/more-than-half-of-college-students-believe-using-chatgpt-to-complete-assignments-is-cheating/.

[6] Lim WM, Gunasekara A, Pallant JL, Pallant JI, Pechenkina E. Generative AI and the future of education: ragnarök or reformation? A paradoxical perspective from management educators. Int. J. Manag. Educ. 2023;21(2):100790. https://doi.org/10.1016/j.ijme.2023.100790.

[7] Chan CKY. A comprehensive AI policy education framework for university teaching and learning. Int. J. Educ. Technol. High. Educ. 2023;20(1). https://doi.org/10.1186/s41239-023-00408-3. 38–25.

[8] Warschauer M, Tseng W, Yim S, Webster T, Jacob S, Du Q, Tate T. The affordances and contradictions of AI-generated text for second language writers. SSRN Electron. J. 2023. https://doi.org/10.2139/ssrn.4404380.

[9] Chan CKY, Lee KKW. The AI generation gap: are gen Z students more interested in adopting generative AI such as ChatGPT in teaching and learning than their gen X and millennial generation teachers? arXiv.org; 2023.

[10] Chiu TKF. The impact of generative AI (GenAI) on practices, policies and research direction in education: a case of ChatGPT and Midjourney. Interact. Learn. Environ. 2023. https://doi.org/10.1080/10494820.2023.2253861.

[11] Kohnke L, Moorhouse BL, Zou D. ChatGPT for Language Teaching and Learning. RELC Journal 2023;54(2):537–50. https://doi.org/10.1177/00336882231162868.

[12] Hwang GJ, Chen NS. Editorial position paper: exploring the potential of generative artificial intelligence in education: applications, challenges, and future research directions. Educ. Technol. Soc. 2023;26(2). https://doi.org/10.30191/ETS.202304_26(2).0014.

[13] Crompton H, Burke D. Artificial intelligence in higher education: the state of the field. Int. J. Educ. Technol. High. Educ. 2023;20(1):1–22. https://doi.org/10.1186/s41239-023-00392-8.

[14] Ng DTK, Leung JKL, Chu SKW, Qiao MS. Conceptualizing AI literacy: an exploratory review. Comput. Educ. Artif. Intell. 2021;2:100041. https://doi.org/10.1016/j.caeai.2021.100041.

[15] Yang ACM, Chen IYL, Flanagan B, Ogata H. Automatic generation of cloze items for repeated testing to improve reading comprehension. Educ. Technol. Soc. 2021;24(3):147–58. https://doi.org/10.30191/ETS.202107_24(3).0011.

[16] Chu H, Tu Y, Yang K. Roles and research trends of artificial intelligence in higher education: a systematic review of the top 50 most-cited articles. Australas. J. Educ. Technol. 2022;38(3):22–42. https://doi.org/10.14742/ajet.7526.

[17] Kim C, Bennekin KN. The effectiveness of volition support (VoS) in promoting students' effort regulation and performance in an online mathematics course. Instr. Sci. 2016;44:359–77. https://doi.org/10.1007/s11251-015-9366-5.

[18] Celik I, Dindar M, Muukkonen H, Järvelä S. The promises and challenges of artificial intelligence for teachers: a systematic review of research. TechTrends 2022;66(4):616–30. https://doi.org/10.1007/s11528-022-00715-y.

[19] Rudolph J, Tan S, Tan S. ChatGPT: bullshit spewer or the end of traditional assessments in higher education? J. Appl. Learn. Teach. 2023;6(1). https://doi.org/10.37074/jalt.2023.6.1.9.

[20] Sullivan M, Kelly A, McLaughlan P. ChatGPT in higher education: considerations for academic integrity and student learning. J. Appl. Learn. Teach. 2023;6(1). https://doi.org/10.37074/jalt.2023.6.1.17.

[21] Cassidy C. Australian universities to return to 'pen and paper' exams after students caught using AI to write essays. January 10. The Guardian; 2023. https://www.theguardian.com/australia-news/2023/jan/10/universities-to-return-to-pen-and-paper-exams-after-students-caught-using-ai-to-write-essays.

[22] Shepherd T. New AI tools that can write student essays require educators to rethink teaching and assessment. May 17. Blog. London School of Economics; 2022. http://eprints.lse.ac.uk/116271/.

[23] De Maio C, Dixon K. Promoting academic integrity in institutions of higher learning: what 30 years of research (1990-2020) in Australasia has taught us. J. Coll. Charact. 2022;23(1):6–20. https://doi.org/10.1080/2194587X.2021.2017972.

[24] UW-Madison Library. Is it a policy, procedure, or guideline? The authors;:2022. https. doi://development.policy.wisc.edu/2022/06/01/is-it-a-policy-procedure-or-guideline/.

[25] Tertiary Education Quality, Standards Agency (TEQSA). What is academic integrity? The author; 2022. https://www.teqsa.gov.au/students/understanding-academic-integrity/what-academic-integrity.

[26] Janke S, Rudert SC, Petersen Ä, Fritz TM, Daumiller M. Cheating in the wake of COVID-19: how dangerous is ad-hoc online testing for academic integrity? Comput. Educ. Open 2021;2:100055. https://doi.org/10.1016/j.caeo.2021.100055.

[27] Von Dran GM, Callahan ES, Taylor HV. Can students' academic integrity be improved? Attitudes and behaviors before and after implementation of an academic integrity policy. Teach. Bus. Ethics 2001;5:35–58.

[28] Pecorari D, Petrić B. Plagiarism in second-language writing. Lang. Teach. 2014;47(3):269–302. https://doi.org/10.1017/S0261444814000056.

[29] Cotton DRE, Cotton PA, Reuben Shipway J. Chatting and cheating: ensuring academic integrity in the era of ChatGPT. Innov. Educ. Teach. Int. 2023. https://doi.org/10.1080/14703297.2023.2190148.

[30] Perkins M. Academic Integrity considerations of AI Large Language Models in the post-pandemic era: ChatGPT and beyond. Univ. Teach. Learn. Pract. 2023;20(2):07.

[31] Dalalah D, Dalalah OM. The false positives and false negatives of generative AI detection tools in education and academic research: the case of ChatGPT. Int. J. Manag. Educ. 2023;21(2). https://doi.org/10.1016/j.ijme.2023.100822.

[32] Yeo MA. Academic integrity in the age of Artificial Intelligence (AI) authoring apps. TESOL Journal, 14(3) 2023;716. https://doi.org/10.1002/tesj.716.

[33] Phakiti A, Fernandez J, Steinhoff A, Yeo MA, Richards JC. Language assessment: a practical approach. Regional Language Centre; 2021.

[34] Yang M. New York City Schools ban AI ChatBOT that writes essays and answers prompts. January 6. The Guardian; 2023. https://www.theguardian.com/us-news/2023/jan/06/new-york-city-schools-ban-ai-chatbot-chatgpt.

[35] Yazici A, Yazici S, Erdem MS. Faculty and student perceptions on college cheating: evidence from Turkey. Educ. Stud. 2011;37(2):221–31. https://doi.org/10.1080/03055698.2010.506321.

[36] Pope N, Green SK, Johnson RL, Mitchell M. Examining teacher ethical dilemmas in classroom assessment. Teach. Teach. Educ. 2009;25(5):778–82.

[37] Pereira D, Flores MA, Niklasson L. Assessment revisited: a review of research in assessment and evaluation in higher education. Assess. Eval. High. Educ. 2016;41(7):1008–32. https://doi.org/10.1080/02602938.2015.1055233.

[38] Sotiriadou P, Logan D, Daly A, Guest R. The role of authentic assessment to preserve academic integrity and promote skill development and employability. Stud. High. Educ. 2020;45(11):2132–48. https://doi.org/10.1080/03075079.2019.1582015 (Dorchester-on-Thames).

[39] Ambrose SA, Bridges MW, DiPietro M, Lovett MC, Norman MK. How learning works: seven research-based principles for smart teaching. 1st Ed. Jossey-Bass; 2010.

[40] King MR. A conversation on artificial intelligence, chatbots, and plagiarism in higher education. Cell Mol. Bioeng. 2023;16(1):1–2. https://doi.org/10.1007/s12195-022-00754-8.

[41] Bretag T, Mahmud SW, James C, Green ME, McGowan UP. Core elements of exemplary academic integrity policy in Australian higher education. Int. J. Educ. Integr. 2011;7(2):3–12. https://doi.org/10.21913/IJEI.v7i2.759.

[42] Möller A. An analysis of university academic integrity policies in New Zealand. J. Furth. High. Educ. 2022;47(3):338–50. https://doi.org/10.1080/0309877X.2022.2130195.

[43] Tatum HE. Honor codes and academic integrity: three decades of research. J. Coll. Charact. 2022;23(1):32–47. https://doi.org/10.1080/2194587X.2021.2017977.

[44] Kohnke L, Moorhouse BL, Zou D. Exploring generative artificial intelligence preparedness among university language instructors. Comput. Educ.: Artif. 2023;5:100156.

[45] Darling-Hammond L, Snyder J. Authentic assessment of teaching in context. Teach. Teach. Educ. 2000;16(5–6):523–45. https://doi.org/10.1016/S0742-051X(00)00015-9.

[46] Brown HD, Abeywickrama P. Language assessment: principles and classroom practices. White Plains, NY: Pearson Education; 2010.

[47] Moorhouse BL, Wong KM, Li L. Teaching with Technology in the Post-Pandemic Digital Age: Technological Normalisation and AI-Induced Disruptions. RELC Journal 2023;54(2):311–20. https://doi.org/10.1177/00336882231176929.

[48] Jones RH, Hafner CA. Understanding digital literacies : a practical introduction 2012. https://doi.org/10.4324/9780203095317.

[49] Markauskaite L, Marrone R, Poquet O, Knight S, Martinez-Maldonado R, Howard S, Tondeur J, De Laat M, Buckingham Shum S, Gašević D, Siemens G. Rethinking the entwinement between artificial intelligence and human learning: what capabilities do learners need for a world with AI? Comput. Educ. Artif. Intell. 2022;3:100056. https://doi.org/10.1016/j.caeai.2022.100056.