Universiteit Leiden
Governance and Global Affairs

FAIR EDUCATIONAL ASSESSMENT IN THE AGE OF GENERATIVE AI

Advisory report following the Leiden Deliberative Assembly on GenAI Use in Assessment

Edited by Josette Daemen, Andrei Poama, Eden Lutz and Praneet Khandal

Co-authored by deliberative assembly participants Ambra Brizi, Anna Dawid, Arnout van Ree, Brandon Zicha, Cristina Del Real, Daný van Dam, David Ehrhardt, Eldad Vainstoc, Elmer Veldkamp, Guido Band, Inge 't Hart, Ingrid Samset, Jean-François Blanc, Jos Rohling, Lisa Wu Jianshun Pieters, Olga Bogolyubova, Rüya G. Koçer, Tanja Masson-Zwaan, Alexander Schuch, Amber Zandstra, Amy McKimm, Athena Thang, Athos Voutsakis, Bram Alkema, Ceddric Koekenbier, David Urban, Deveney Muller, Egor Buryi, Gaia Sterre Mirotti, Irinie Youssef, Jessie Zhu, Nao Watanabe, Omeima Khedhri, Peppi Tenkanen, Peter Bosch, Sheila Varadan, Simon Klaver, Sofía Paz Espinoza Gutierrez, Stefan Bojilov, Tamara van Koppen, William Chan, and Willow Miller

Four participants of the deliberative assembly, Maximilian Albers, Andrea Reyes Elizondo, Femke Mostert, and Daniël Tulkens, do not identify as co-authors because they feel that the report does not reflect their values, but they deserve recognition here for making an important contribution to the deliberative assembly.

For questions about the FAIR-ASSESS project and this report, please contact Josette Daemen at j.a.m.daemen@fgga.leidenuniv.nl or Andrei Poama at a.poama@fgga.leidenuniv.nl.

Three groups of people deserve special mention for their assistance or advice. First, the advisory team that provided feedback on the materials for the deliberative assembly: Praneet Khandal, Julian van der Kraats, Ella Akin, Katie Evans and Norbert Paulo. Second, the expert team that stood ready to answer questions from participants throughout the deliberative assembly: Katie Evans, Julia Dawitz, Veit Kubik, Maryam Alqassab, Francesco Walker, Zohre Mohammadi Zenouzagh, Julian van der Kraats, Joy Lee, and Mario de Jonge. Special thanks go to Katie Evans, Maryam Alqassab and Mario de Jonge for delivering expert presentations during the first deliberative assembly session. Finally, gratitude goes out to all members of the FAIR-ASSESS project: Josette Daemen, Andrei Poama, Eduard Fosch Villaronga, Praneet Khandal, Tom Theuns, Bram Klievink, Mariëlle Linting, Jelena Belic, Michael Lew, Marga Sikkema-de Jong, Eamon Aloyo, Min Cho, Dorota Mokrosinska, Mario de Jonge, Julian van der Kraats, Roeland van der Rijst, Katie Evans, Joy Lee, Matthijs Westera, Ella Akin, and Eden Lutz.

The *Leiden Deliberative Assembly on GenAI Use in Assessment* is an initiative of the project *Fair Educational Assessment in the Age of AI* (FAIR-ASSESS) and was co-created by dembrane. This project is supported and funded by the Faculty of Governance and Global Affairs (FGGA).

Leiden University
The Hague, April 2026

Executive summary

This report documents the outcome of the *Leiden Deliberative Assembly on GenAI Use in Assessment*. The assembly took place in January-February 2026, and was organised by the project *Fair Educational Assessment in the Age of AI* (FAIR-ASSESS).

The FAIR-ASSESS project addresses the increasing use of Generative Artificial Intelligence (GenAI) in various stages of the assessment process in higher education. Not only do many students use GenAI when completing assessments, there are also teachers who make use of the technology for designing assessments, and some experiment with using GenAI for feedback and even grading purposes, although especially the latter is controversial under the EU AI Act and many higher-education institutions' own examination rules. These actual and potential uses of GenAI confront higher education in general, and Leiden University in particular, with a pressing challenge: *how do we make assessments more, rather than less fair in the age of GenAI?*

To answer this question, we convened a relatively small but diverse group of students and teachers from all seven Leiden faculties. Gathered during four half-day sessions spread out over five weeks, 47 participants moved through various stages of deliberation. They were informed about the basics of GenAI technology and scientific research on the impact of GenAI use on higher-education assessment, reasoned about potentially problematic or beneficial forms of GenAI use at various assessment stages, reflected at length about how to regulate the use of GenAI in assessment practices at Leiden, and cast their votes on rules meant to guide or constrain GenAI use in assessment practices at our university.

The result of this process is a set of 32 rules, complemented by broader recommendations concerning GenAI access and use in assessment practices. The rules correspond to a number of key principles that student and teacher participants believed should guide assessment at Leiden University: validity, reliability, transparency, usability, trust, authenticity, responsibility, accountability, equality and integrity. Safeguarding these principles was seen as crucial to ensuring and potentially enhancing the fairness of educational assessments in the age of GenAI.

Voting showed that none of the rules proposed during the assembly obtained full support from all participants. The list of rules presented in this report should therefore not be read as an expression of unanimous agreement, but rather as a collection of guidelines that gained moderate to substantial support from participants. The rules offer resources for students, teachers and university administrators who are trying to navigate assessment practices against the current technological background, and can contribute to shaping regulations of GenAI use in assessment across our university.

Table of contents

Introduction 4

Section 1: Deliberative assembly participants 7

Section 2: Deliberative assembly process 8

Section 3: Rules for GenAI use in assesment 10

Section 4: Recommendations on GenAI and assessment 15

Section 5: Limitations of the deliberative assembly 18

Appendix I: Unedited rules and voting results 19

Appendix II: Similar reports 28

Introduction

Generative Artificial Intelligence (GenAI) raises significant challenges for higher-education assessment. While certainly not all students and teachers make use of this technology, instances of GenAI use can be found in various stages of the assessment process (see Figure 1).

When designing assessments, some teachers use GenAI to formulate exam questions or practicums resembling challenges of the kind that students will encounter in their future careers. When completing assessments, many students use GenAI to browse ideas for essays, edit their writing, or generate texts, code or images. When evaluating students' performance on assessments, some teachers experiment with using GenAI as an assistant for feedback or even grading, although especially the latter is controversial under the EU AI Act and many higher-education institutions' own examination rules. Notwithstanding varying levels of adoption or integration, GenAI is getting so widespread that institutions face an urgent need to guide and constrain its use.

A core concern when regulating GenAI in higher-education assessment practices is *fairness*. An assessment can be considered fair if it is guided by principles that cannot reasonably be rejected by those who participate in assessment practices - most notably, students and teachers. Well-known principles for higher-education assessments include validity (does an assessment measure what it is intended to measure?), reliability (is the measurement consistent?), and transparency (do students know what is expected of them?). In addition to these three, the website listing basic principles for adequate assessment at Leiden University mentions the principle of usability (is the assessment method also practicable?).

For abstract principles such as these to be achieved in practice, specific rules must be decided on that indicate what teachers and students may and may not or should and should not do at different stages of the assessment process. Figure 2 offers a rough illustration of how the value of fairness can inform different assessment principles, and how these principles can in turn offer the basis for concrete rules that constrain or otherwise guide assessment practices.

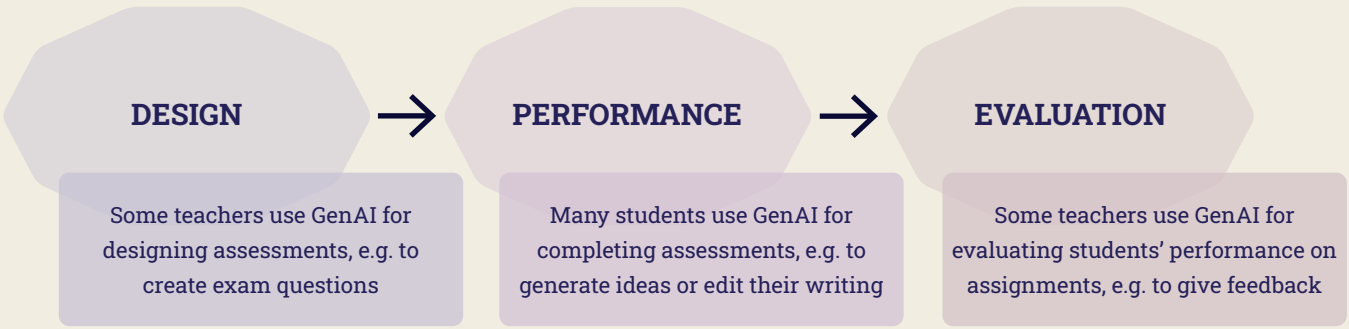


Figure 1. GenAI use across the three assessment stages

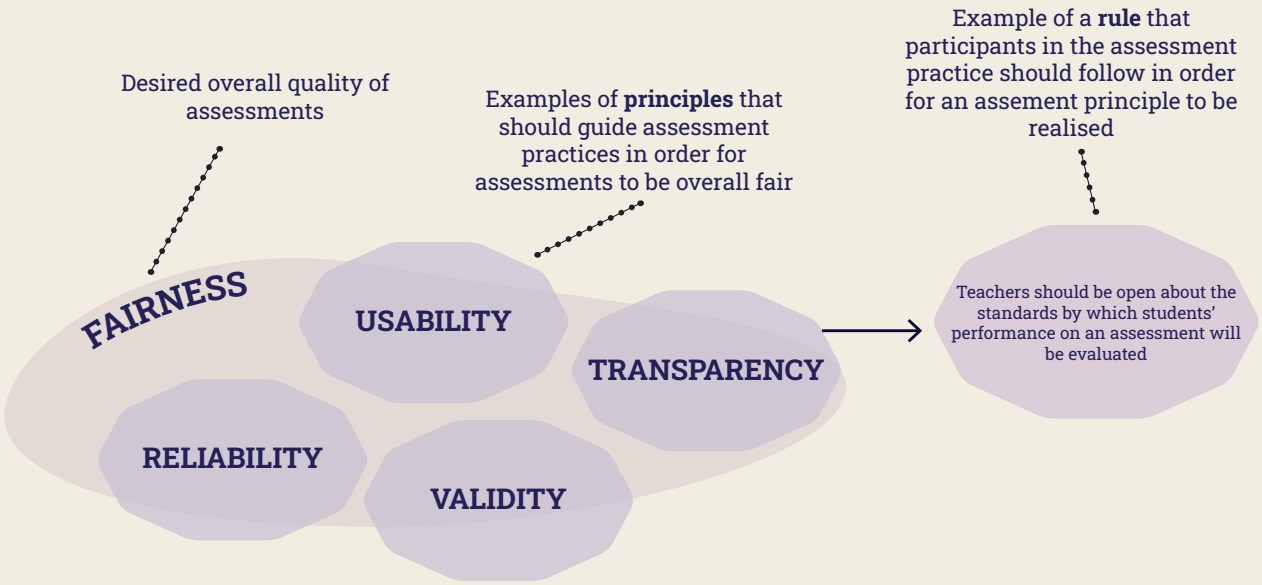


Figure 2. Fairness in the context of educational assessment

GenAI technologies can bring opportunities for higher education, but they also come with serious challenges for realising key assessment principles. How can students' knowledge and skills be assessed in a way that is valid and reliable, when they have the option of outsourcing many tasks to GenAI? What does transparency demand of teachers when it comes to instructing students about GenAI use in assessments? And what are the requirements when it comes to teachers' own use of GenAI? In short, which ways of using GenAI (if any) are compatible with fairness in educational assessment?

These questions drive the action research project *Fair Educational Assessment in the Age of AI* (FAIR-ASSESS). This multi-disciplinary project, which started in November 2024 with the support of the Faculty of Governance and Global Affairs (FGGA), unites 20 scholars from five Leiden University faculties, collaborating with a number of researchers from other higher-education institutions in the Netherlands and beyond. FAIR-ASSESS has resulted in two academic papers: *'Uses of GenAI in higher education assessment practices: a scoping review and typology'* and *'Fair educational assessment in the age of GenAI: a contractualist framework'* (working versions of which are openly accessible on the website of the FAIR-ASSESS project). In addition to producing this scientific output, FAIR-ASSESS aspires to make a hands-on contribution to the revision of current assessment regulations in the context of proliferating GenAI use at Leiden University and beyond.

This report presents the results of the *Leiden Deliberative Assembly on GenAI Use in Assessment* that took place in January-February 2026. The deliberative assembly brought together students and teachers from all seven Leiden faculties to discuss the question: *how can we make higher-education assessments more fair in the age of GenAI?* Higher-education institutions around the world have been struggling with this question ever since the launch of ChatGPT in November 2022. Various institutions have produced insightful reports on the future of assessment against the background of developments in AI, often with the help of experts in educational science, educational technology and higher-education policy (an indicative list of such reports is included in Appendix II of this report).

What distinguishes this report from others is that it treats the development of GenAI guidelines for educational assessment as a challenge that has to be approached in a way that is evidence- and expertise-sensitive, but remains as much as possible led from the bottom up. The question of how assessments can remain or become fair

in a time of GenAI should not be exclusively placed in the hands of experts, education administrators and policy-makers. It is also, and perhaps primarily, a question for teachers and students to discuss and determine. A first reason for this is epistemic: bringing together the perspectives of those who are directly involved in assessment practices ensures that rules are informed by experiential knowledge, and can thus improve the quality of assessment regulation. A second reason is moral: following a commitment to democratic values, we can argue that people who are subject to a rule should also have a say in the formation of this rule, and the same applies to students and teachers in the regulation of assessment practices. Third, there is a practical reason for taking a bottom-up approach to rule-making about GenAI and assessment: actively involving students and teachers in this process increases the sense of ownership of the academic community over the rules that are adopted, and can lead to more commitment and support among those whom the rules are supposed to guide.

To facilitate a bottom-up approach to the regulation of GenAI in assessments at Leiden University, FAIR-ASSESS organised a *deliberative assembly*. This is a gathering of a diverse group of people who discuss a complex problem and develop ideas and solutions for tackling it. For the design and facilitation of the deliberative assembly, FAIR-ASSESS collaborated with dembrane, a start-up company specialised in organising stakeholder engagement processes and citizen forums. The deliberative assembly included 47 members of the Leiden academic community. Participants were convened in four five-hour sessions over the course of five weeks, moving through various stages of deliberation. The objective was to jointly arrive at a set of rules that can help guide the use of GenAI in assessment practices at our university. The next sections of this report detail both the deliberative assembly process and its outcome.

Section 1: Deliberative assembly participants

After a university-wide campaign inviting all students and teachers of Leiden University to apply for participation in the deliberative assembly, 135 persons submitted an application. In the sign-up form, applicants were asked to register: (1) their faculty, (2) their department, (3) for students, the level of their study programme (bachelor/master), (4), for teachers, their job title (instructor, PhD student, postdoc, assistant professor, associate professor, full professor), (5) their nationality, (6) their gender, and (7) their attitude towards GenAI ('What is your general feeling about GenAI use in higher education? Indicate on a scale from -5 (very negative) to +5 (very positive)').

A stratified random sample was drawn of 49 participants, ensuring as much diversity as possible in terms of the seven factors listed above. Differences in size between the faculties were considered when filling up the spots left open as a result of the low number of applications from the smallest faculty (Archaeology): these spots were given to participants from larger faculties with high numbers of applications (Social and Behavioural Sciences, FGGA, Humanities). Two persons who were selected did not attend the assembly in the end, so the final group counted 47 participants.

Table 1 shows the composition of the group per faculty and participant category (teachers/students). Figure 3 shows the distribution of participants across the categories of answers to the question concerning their general feeling about GenAI use in higher education.

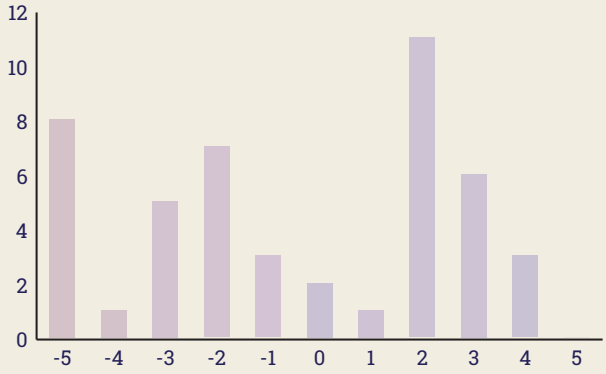


Figure 3. Distribution of participants across the categories of answers to the question 'What is your general feeling about GenAI use in higher education? Indicate on a scale from -5 (very negative) to +5 (very positive)'

The names of participants who wished to be included as co-authors of this report are listed on the cover page, as well as the names of four participants who did not identify as co-authors because they felt that the report did not reflect their values; one other participant preferred to remain anonymous.

At the end of the deliberative assembly, all participants received a participation certificate. Student participants also received a one-time remuneration of €150.

Faculty	Teachers	Students	Total
Archaeology (FdA)	0	2	2
Governance and Global Affairs (FGGA)	5	4	9
Humanities (FGW)	4	4	8
Law (FdR)	2	4	6
Medicine (LUMC)	2	3	5
Science (FWN)	3	4	7
Social and Behavioural Sciences (FSW)	5	5	10
Total	**21**	**26**	**47**

Table 1. Deliberative assembly participants

Section 2: Deliberative assembly process

Before the start of the deliberative assembly, participants received a handbook with background information and instructions (available on the FAIR-ASSESS website). To ensure an evidence- and expertise-informed deliberative process, the handbook featured basic information on GenAI and its impact on higher-education assessment processes and outcomes; a normative analysis on fairness questions raised in higher-education assessment practices; and instructions about deliberative procedures and discussion rules. The handbook included a glossary of key terms used in the debate about the topic (e.g. 'Generative Artificial Intelligence', 'formative assessment', 'summative assessment', 'validity', 'reliability') and summaries of the two papers written by the FAIR-ASSESS team (also available on the FAIR-ASSESS webpage). Finally, the handbook listed the names and contact details of nine experts that participants could consult over the course of the deliberative assembly on any substantive questions, the FAIR-ASSESS project organisers, the moderator from dembrane (the company supporting the organisation of the deliberative assembly), and an an independent confidential counsellor.

The deliberative assembly spanned four sessions of five hours each. The sessions took place on 15 January, 22 January, 29 January and 12 February 2026 in conference rooms in The Hague (sessions 1, 2 and 3) and Leiden (session 4). Over the course of these meetings, the 47 deliberative assembly participants moved through various stages of deliberation. The process could be visualised in the shape of a fish, starting with the specific question of how we can make higher-education assessments more fair in the age of GenAI, then broadening up participants' perspectives to explore the challenge in its full complexity, and in later stages finding focus when working towards a proper response, to finally end up with a set of rules for GenAI use in Leiden University assessment practices (see Figure 4).

Throughout the deliberative assembly, an AI tool developed by dembrane was used to support the process. The tool recorded participants' conversations, transcribed these in real time, and (when prompted) provided fast feedback at several points during the deliberative assembly. This enabled the moderator to highlight key themes that came up in group discussions. The tool was also prompted to make a summary of each meeting, which the organisers uploaded to a Teams environment uniting all participants of the deliberative assembly on the day after the session. Participants were invited to notify the organisers when they disagreed with any information included in these summaries, after which the organisers corrected the text. At no point during the assembly was the AI tool used to formulate, edit or otherwise modify rules or recommendations. Nor was the dembrane tool (or any other AI instrument) used for the writing of this report. The tool is fully GDPR and ISO270001 compliant. The data that were collected were only used for the purposes specified above, and will be deleted once the FAIR-ASSESS project is completed.

Figure 4. The deliberative assembly process

Session 1 was dedicated to learning about the basics of GenAI and the various aspects of fairness in educational assessment. Participants shared their first reasons for worrying or being excited about GenAI in assessment. They listened to talks from experts on what GenAI is and how it works. They were presented with the principles that officially guide assessment practices at Leiden University (listed on the University website). Based on the two papers written by the FAIR-ASSESS team, they were informed about how and with what outcomes students and teachers are currently using GenAI in assessments in higher-education institutions worldwide, and they were introduced to a specific mode of reasoning about fairness in the context of educational assessment. In smaller groups, participants discussed the assessment principles of validity, reliability, transparency and usability, and could propose other principles that they considered important for fair assessment, broadly speaking.

Session 2 focused on learning about the problems and opportunities created by GenAI in the assessment context. Participants were divided into seven faculty groups. Each group was presented with a list of practice-based hypothetical cases (available on the FAIR-ASSESS website) revolving around the use of various sorts of GenAI tools (commercial/in-house), by various participants in the assessment practice (students/teachers), across various stages of the assessment process (design/performance/evaluation), for various assessment types (exams/presentations/practicums/take-home assignments/theses). Based on these cases and additional cases proposed by the faculty groups themselves, participants reflected on the consequences of various GenAI uses. These included benefits, such as a reduction of workload for instructors and increased feedback for students, but also burdens, such as a setback in cognitive skills and ambiguity about responsibility for assessment outcomes. Participants discussed concerns connected to these consequences in the light of the assessment principles discussed previously, as well as the overarching value of fairness.

Session 3 marked the transition from understanding the challenge in its full complexity to thinking about how to respond to it. Specifically, participants were asked to think about rules to guide the use of GenAI in educational assessment in a way that meets the concerns or issues identified in Session 2. To do so, participants were again divided into faculty groups. Based on their discussions during the previous session, each group decided which concerns and issues they thought needed to be most urgently addressed. Examples included addressing lack of information about responsible GenAI use, where to draw the line between allowed and forbidden forms of GenAI employment in assignments, and how to account for students with ethical objections against using the technology. Trying to resolve such issues, each group constructed a first set of rules that they thought should apply either within their faculty or university-wide. The different faculty groups then presented their rules to the entire assembly, taking questions, comments and suggestions from the audience. Based on this feedback, each group revised and complemented their own set of rules.

Session 4 built toward a final set of rules and the outline of additional recommendations. Participants were once again split into faculty groups. After another round of refining the rules that they formulated during the previous session, each faculty group submitted its rules to the moderator. The moderator created an online form listing all the rules, 52 in total. Through this form, participants indicated for each rule to what extent they agreed with it, using a scale of -5 ('do not agree at all') to +5 ('fully agree'). The moderator presented the results of this voting round to the participants, highlighting which rules were most controversial (obtaining a very positive rating from some participants and a very negative rating from others) and which rules received low scores. Participants split into groups to discuss the seven most controversial rules and propose revisions or decide to drop any specific rule. Following the submission of the revised rules (five were revised, two dropped), participants went through a last voting round on all rules. Finally, participants were invited to write down any general or specific recommendations regarding GenAI and/or assessment at the university that did not fit the strict format of a *rule*. To put it in a slogan: rules tend to be about *don'ts*, while recommendations tend to be about *do's*.

Section 3: Rules for GenAI use in assesment

This section presents a list of rules proposed by participants to guide the use of GenAI in assessment practices at Leiden University. These rules are edited versions of the rules submitted by the participants at the end of the deliberative assembly. To avoid duplicates, in editing this report, several rules were merged, leading to a final list of 32 rules (instead of the 50 rules that were proposed and voted on by participants in the final round of session 4). Unedited versions of all initial rules are included in Appendix I of this report. The list below indicates for each edited rule the corresponding unedited rule(s).

Appendix I also includes the results of the final voting round, during which participants indicated the extent to which they agreed with each rule, using a scale of -5 ('do not agree at all') to +5 ('fully agree'). Importantly, although none of the rules in their final formulation obtained a negative aggregated score from all participants' votes combined, no rule obtained a full agreement from all participants either. The list below should therefore not be read as a set of rules that participants unanimously agreed on. Rather, it should be read as a collection of rules that could count on moderate to substantial support from the deliberative assembly as a whole. Not every participant finds their ideas and opinions reflected in every single rule.

Editorial revisions included in the list of rules presented below were based on the documentation kept by the faculty groups who discussed the rules during the deliberative assembly, as well as norms for clear formulation of rules that were shared with participants during the deliberative assembly. As editors, we aimed to retain participants' initial formulations as much as possible. Some of the rules listed below may appear similar on a first reading, but we ensured that the content of each rule is distinct from that of other rules. Following the editorial process, all participants received a request via email and Teams to review the edited list and raise objections or make final comments to the edits made. Eight participants submitted a written response, resulting in a total of 96 suggestions for revision. All suggestions that were reasonably in line with what was proposed by the deliberative assembly were adopted in the final version of the report.

For each *rule*, the list below indicates one or more corresponding assessment *principle(s)*. The principles featured here include the four assessment principles officially implemented at Leiden University (according to the University website): validity, reliability, transparency and usability. In addition, there are principles proposed by participants themselves: trust, responsibility, accountability, equality and authenticity (in the context of assessment, the latter means that assessments should focus on tasks that are worthwhile for students to be able to perform in the real world).

Connecting specific rules to general assessment principles reflects a normative approach distinctly favored by the FAIR-ASSESS project and the reasoning process that informed the deliberative assembly. During the assembly, participants were invited to start by reflecting on abstract assessment principles, and encouraged to use these reflections as a basis for subsequent deliberations about specific issues prompted in concrete cases of GenAI use in assessment, with the ultimate aim of proposing rules for dealing with them. When submitting their rules at the end of the deliberative assembly, participants were asked to indicate to which assessment principles their rules connected. Their input was taken as a basis for matching specific rules to the abstract principles below.

The assessment principles included in the following list can be read in two ways. First, they offer general normative categories that can help teachers, students and educational administrators to make sense of the general spirit of a rule and explain how a rule can contribute to valuable assessment practices. Second, they can assist other university participants to derive similar rules for GenAI use in assessment in the future, at Leiden University or elsewhere in higher education.

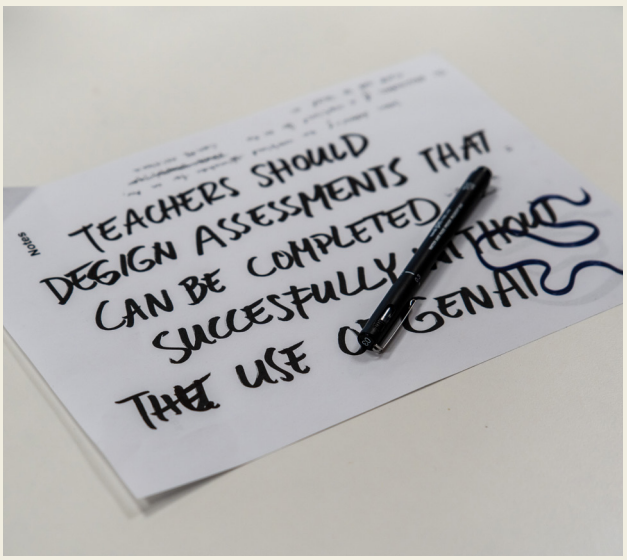
These are the 32 rules put forward by the deliberative assembly participants.

See Appendix I

	Rule	Corresponding principle(s)	Original rule
1.	Teachers should design their assessments in such a way that these assessments accurately measure the extent to which students have achieved the learning objectives of the course or degree programme, without the measurement being distorted by (possible) GenAI use by students.	validity	22
2.	If a study programme chooses to incorporate assignments permitting the use of GenAI, it should implement a curricular plan to ensure that students continue to attain the fundamental skills and learning objectives of the programme without relying on GenAI.	validity	32
3.	Educational directors should ensure that a study programme's curriculum includes obligatory checkpoints to assess whether students have attained the transferable skills essential to the discipline and can apply these skills without relying on GenAI.	validity, authenticy	40
4.	Teachers should ensure that at least 50% of the grade for each course is based on assessment components that students complete in conditions where they have no access to GenAI, and students have to pass these components to pass the course. The rule does not apply if developing students' GenAI literacy is the course objective.	validity, reliability	48
5.	A course grade should be based on at least one assessment method whereby students have no access to GenAI.	validity, reliability	47
6.	Students should be allowed to use GenAI if and only if its use complies with the course rules and learning objectives.	transparency, trust, integrity	49
7.	If a course or teacher requires that students use a GenAI tool, access to this tool should be provided by the university.	equality	11
8.	Students should not be obligated to use GenAI systems in their assessments.	equality	38, 43
9.	Students should always be able to explain how they arrived at their answer on a particular assessment, and to provide a reasoned argument for their answer.	transparency	12
10.	For all assessments in which they used GenAI, students should explicitly mention in a statement which specific GenAI tools, models and/or platforms they used, how they used them, and whether GenAI output was verified.	transparency, trust	17
11.	For all assessments in which they used GenAI, teachers should explicitly mention in a statement which specific GenAI tools, models and/or platforms they used, how they used them, and whether GenAI output was verified.	transparency, trust	14, 23, 28, 42

	Rule	Corresponding principle(s)	Original rule
12.	In research-intensive assignments such as theses, internship projects and field research, students must disclose any GenAI use and provide a consideration of the ethical implications of this use.	transparency, trust, authenticity	35
13.	For all assessments, faculties should provide a standardised AI use disclosure form that includes at least the type of AI use and the reason for using AI.	reliability, transparency	30
14.	In group settings, teachers should facilitate open and respectful discussions when there are different opinions about GenAI use among students.	transparency, trust	46
15.	In group assignments, individual group members should in principle bear joint responsibility for GenAI use, but teachers should ultimately assess individual group members separately if group members have different views about GenAI use.	validity, accountability	36, 45
16.	Students and teachers should never share sensitive information (e.g., medical information and personally identifying information) when using external GenAI tools.	trust	4
17.	Students and teachers should never share data that is part of an assessment with external GenAI providers without the consent of the (fellow) student or teacher who authored the work.	trust, transparency	15, 18
18.	When calculating the workload in EC for a given course, course coordinators should estimate the workload based on the assumption that students will not make use of GenAI tools to complete assessments, unless GenAI use is explicitly required in these assessments or the course learning objectives are related to the use of GenAI.	usability, validity	24, 33
19.	For all assessments, teachers must, at the start of a course or in the syllabus, provide clear and specific instructions about allowed uses of GenAI (if any), based on intended learning outcomes.	transparency, trust	7, 8, 10
20.	Teachers should clearly indicate in the syllabus or other course document how they define and respond to unauthorised uses of GenAI systems in students' assignments.	transparency, trust	21, 27
21.	If the use of GenAI systems is mandatory for a course, teachers should include this information in the learning objectives listed in the course guide and/or prospectus.	transparency, trust	16, 41
22.	At the start of every semester or programme, the board of examiners should inform students about the relationships between learning objectives, assessment methods and the (possible) use of GenAI systems.	transparency, trust, usability	44

	Rule	Corresponding principle(s)	Original rule
23.	To allow students who do not want their assignments evaluated by GenAI to not enroll in the course, teachers may use GenAI for evaluation purposes only if they disclose such uses in the course study guide/prospectus.	transparency	50
24.	Teachers should always carry the final responsibility and accountability for the assessment method adopted and the feedback and/or grade following from it.	transparency, trust, accountability	1, 2, 5
26.	When GenAI is used to generate questions and/or answer models for assessments, the output of the GenAI model should be reviewed and approved by at least two human assessors.	validity	37
27.	Given their unreliability, AI use detectors should not be used as an assessment tool or aid.	trust, validity, reliability	39
28.	The university should provide a list of approved GenAI tools that students can use when they are allowed to use GenAI.	transparency, usability	25, 26
29.	Any university-approved GenAI tool must adhere to GDPR standards, to ethical standards around legally sourced training data and transparency, and to sustainability standards.	transparency, integrity	3, 6, 20
30.	Each faculty should institute a counsellor whose function it is to promote AI literacy and mediate discussions around GenAI use and assessments.	responsibility	34
31.	Study programmes should implement measures to increase AI literacy and more responsible use of AI among students.	responsibility	19
32.	If GenAI use in assessment is allowed within a programme, learning about the risks and best practices of GenAI use should be included in the programme as a mandatory course.	responsibility	29



Section 4: Recommendations on GenAI and assessment

In addition to the process of forming rules, the deliberative assembly included a short time slot in session 4 where participants were invited to propose any recommendations that could safeguard or improve the quality and value of our assessments in the context of proliferating GenAI use.

Participants wrote the recommendations individually. This means that the recommendations do not reflect any group position, nor are they necessarily consistent with one another. An important difference to note between rules and recommendations is that rules are generally meant to focus on constraints for how GenAI tools should be accessed, used or regulated at Leiden University, while recommendations emphasise initiatives or actions that should be more positively undertaken to deal with or capitalise on increasing GenAI accessibility. To put it in a slogan: rules tend to be about don'ts, while recommendations tend to be about do's.

Individual participants came up with 27 distinct recommendations, which address two types of actors: the university administration, and the student and teacher community at Leiden. We grouped the university-level recommendations into four thematic sub-sets: assessment & GenAI governance, (costs of) cooperation with third-party GenAI providers, student protections, and teacher development. For the sake of clarity, recommendations were rephrased and categorised by the editors of this report. The recommendations are included in the list below.

University administration

Assessment & GenAI governance

To meet the challenges raised by AI in general, try to move from a university culture that is centered on productivity, metrics, documentability and/or unenforceable rights to a university culture that emphasises character, academic ethos, accountability and judgment.

In setting up any rules concerning GenAI, act with prudence (do no harm!) and with restraint (avoid introducing too many rules).

GenAI-related regulations should not reduce or otherwise negatively impact the diversity and creativity of different assessment forms or social interactions between students and teachers.

The university should set up ongoing and/or periodic workshops where GenAI implications for teaching and assessment practices can be discussed.

The university should introduce a standard framework and vocabulary (e.g., definitions of AI) to allow for consistent disclosure in assessment guidelines.

The university should establish an AI officer advisory position where students and teachers can address questions concerning (Gen)AI use.

The university should initiate and/or conduct large-scale studies to track the adoption of GenAI applications at Leiden University.

The university should introduce human-AI collaboration as a fourteenth transferrable skill at Leiden University.

The university should convene an expert group who can arrive at a set of recommendations concerning how to reform (bachelor and master) theses.

The university should develop a long-term plan concerning the requirements of fair assessment practices, such as student/teacher ratio in assessment.

The university should take into account that AI is developing rapidly, which makes it hard to develop lists of 'approved' GenAI tools and to make students and staff adhere to rules surrounding GenAI use. It should also take into account that which models and platforms may be helpful is likely to vary across disciplines and fields. Against this background, creating a long list of rules with limited feasibility may cause frictions within the organisation.

Cooperation and costs

The university should not without evidence accept product performance or other educationally relevant claims that are presented by AI service providers.

To avoid a premature and/or inefficient financial lock-in, the university should be responsible in contracting with GenAI providers.

GenAI should not be introduced or implemented as a cost-saving measure or based on financial constraint considerations.

The university should prioritise investing in AI-proof teaching facilities.

The university should try to minimise any dependence on commercial AI providers, such as providers of GenAI, cloud services or campus licenses for office applications.

The university should provide students with pre-paid GenAI services through their student accounts.

Student protections

The university should as much as possible protect student data from commercial use.

The university should pay attention to the inequalities that AI can create among members of our teaching and student community.

To ensure that academic skills are preserved, the university should promote an AI-free learning environment.

Every university programme should contain a mandatory course on risks and best practices of GenAI use in education.

Teacher development

An AI literacy module should become a mandatory element in the BKO training and be taken by teachers who already hold a BKO or equivalent qualification.

The university should allow teachers to develop and work with pedagogical AI assistants via the Leiden University Cognitive Agent (LUCA) platform.

Teachers should be incentivised to integrate AI use to enhance the impact of their teaching.

Teachers and/or students

Teachers and students should shift as much as possible towards process-based, instead of product-based assessment practices.

Teachers should seriously pay attention to the non-quantifiable aspects of AI use by students, such as the establishment of rapport and/or trust between teachers and students.

Assessments should incentivise students to use and experiment with GenAI in a way that promotes their learning agency and autonomy, rather than make them dependent on it.

Section 5: Limitations of the deliberative assembly

The deliberative assembly was designed and organised with much thought and care. Many participants indicated that they considered it a valuable experience, felt that their voice was heard in the process, and were happy with the rules and recommendations that the four deliberative sessions generated. Even so, there is room for improvement on some points. To learn from the deliberative assembly for possible future editions at Leiden University or similar initiatives at other higher-education institutions, we briefly reflect on a few limitations of the deliberative assembly.

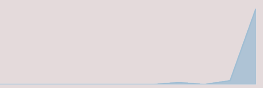
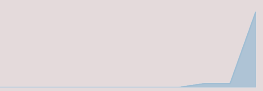
At the end of the assembly, some participants indicated that they had hoped to achieve a more concrete outcome than the rules and recommendations listed in this report. Some noted that they would have preferred to go through a more extensive training by experts to learn what GenAI is and how it affects teaching, learning, and assessment. A few participants remarked that they would have liked deeper discussions about the ethics of AI technology. A couple of participants also found it problematic that an AI tool (the tool developed by dembrane, the company supporting the organisation of the assembly) was used for giving recaps and making summaries of conversations during the assembly.

Upon reading the first version of this report, participants mentioned two additional points of improvement. First, they indicated that it would have been good to have had the final session of the assembly with a concept version of this report in hand. This could have focused the conversation on the specifics of this document, engaging with the remaining points of contention more deeply, and coming to a more coherent list of rules, recommendations, and open questions. Second, participants indicated that they could have used more time for going through the first version of the report. They now had five working days to send in their comments. Seven working days would have been a preferable revision interval.

As organisers, we value these feedback points from participants and recommend that they be taken into account in future initiatives. Given the ongoing developments in AI and the continuous need for higher-education institutions to keep their assessment policies up to date, we see ample room for follow-up events along the lines of the deliberative assembly in the years to come. In the meantime, the rules and recommendations presented in this report provide important mileposts for students, teachers and university administrators who are trying to navigate assessment practices against the current technological background, and form a basis for shaping official regulations of GenAI in assessment across the university.

Appendix I: Unedited rules and voting results

Below is a list of the unedited formulations of the rules submitted at the end of the deliberative assembly, including the results of the last voting round. During this voting round, participants indicated for each rule to what extent they agreed with it, using a scale of -5 ('do not agree at all') to +5 ('fully agree'). The table below lists all unedited rules from the rule obtaining most agreement to that obtaining least agreement. The first column indicates the place of the rule within this ranking. The third column shows the aggregated score that a rule obtained from all participants' votes combined. The fourth column visualises the voting distribution of the scores, with the horizontal axis displaying the scores from -5 to +5, and the vertical axis displaying the number of participants assigning a certain score. The fifth column indicates the faculty group who submitted the rule during the deliberative assembly. The last column gives for each unedited rule the number of the edited rule to which it corresponds within the final list of 32 rules presented in section 3 of this report.

Rank	Unedited rule	Score	Score distribution	Faculty	Edited rule
1	Examiners should always be responsible and remain accountable for the content of their evaluations	210		FGGA	24.
2	The teacher must carry the final responsibility and accountability over feedback or grading on an assessment.	209		FGW	24.
3	Any University-endorsed Generative AI tool must adhere to the legal standards of the GDPR, et cetera.	201		FGW	29.
4	Sensitive information (e.g. medical info/ personally identifying information/etc) should not be shared with external AI tools.	197		FWN	16.

Rank	Unedited rule	Score	Score distribution	Faculty	Edited rule
5	The teacher must carry the final responsibility and accountability over the methods and content of an assessment.	197		FGW	24.
6	Any University-approved Generative AI tool must adhere to ethical standards around legally sourced training data, transparency etc.	189		FGW	29.
7	At the start of a course, teachers must provide students with specific, clear AI usage rules for assignments in that course, based on intended learning outcomes.	182		LUMC	19.
8	Teachers must specify in all assessment instructions if, and how, the use of Gen AI is permitted.	182		FdR	19.
9	In assessments where Gen AI use is permitted, students must disclose how Gen AI was used.	182		FdR	29.
10	The teacher must clearly state if and how the use of GenAI is allowed in the assessment.	181		FWN	19.

Rank	Unedited rule	Score	Score distribution	Faculty	Edited rule
11	If GenAI-tool usage is prescribed/demanded by Leiden University, access to said tool should be provided by Leiden University.	181		FSW	7.
12	The student must always be able to explain how they came to their answer and argue the answer.	181		FGW	9.
13	The teacher must always be able to explain how they came to the format of assessment and feedback on assessments	176		FGW	11.
14	When examiners use GenAI for grading purposes, they should always disclose how they used it	171		FGGA	17.
15	Any data part of an assessment must not be shared with external parties, including providers of GenAI, without consent of the respective author.	171		FWN	21.
16	The teacher must clearly state if the use of AI is required to complete the course in the study guide/prospectus. (before start course)	168		FWN	10.

Rank	Unedited rule	Score	Score distribution	Faculty	Edited rule
17	Students should make explicit in their assignments which specific GenAI tools, models, and platforms they used for their assignment	166		FGGA	10.
18	Both teachers and students should always request permission before submitting the work of a (fellow) student to non-licensed GenAI.	162		LUMC	17.
19	Study programmes should implement measures to increase AI literacy and more responsible usage of AI among students.	162		FSW	31.
20	The University must consider the Generative AI model's sustainability standards when deciding which Generative AI tool to approve.	157		FGW	29.
21	Lecturers should make explicit in their syllabus how they assess the potential unauthorised use of GenAI in students' assignments	153		FGGA	20.
22	Teachers must ensure each assessment is in line with the learning goals of that course or the degree irrespective of the GenAI usage	153		FGW	1.

Rank	Unedited rule	Score	Score distribution	Faculty	Edited rule
23	All use of GenAI in assessments by students and teachers must be disclosed	152		FWN	11.
24	When calculating the workload in EC for a given course, course coordinators should calculate workload estimations based on the assumption that students will not make use of GenAI tools to complete the expected output, unless GenAI is required by the course coordinator.	152		FSW	18.
25	The University should provide a list of approved Generative AI services for staff to use when allowed to use AI.	149		FGW	28.
26	The University should provide a list of approved Generative AI services for staff to use when allowed to use AI.	146		FGW	28.
27	If the use of Gen AI is not permitted, teachers must clarify the consequences of such use.	145		FdR	20.
28	Students and instructors should include a statement on the use of AI in their assignment and assessment documents	140		FGW	11.

Rank	Unedited rule	Score	Score distribution	Faculty	Edited rule
29	If GenAI use in assessment is allowed at a faculty, learning about the risks and best practices of GenAI use should be included in the program as a mandatory course	140		FWN	32.
30	The faculty should provide a standardized format for AI disclosure--including both the kind of use as well as the rationale for it	138		LUMC	13.
31	AI use disclosure must include how GenAI was used (structure, grammar, etc.), which model was used and what, if any, verification...	135		FWN	30.
32	If a programme chooses to integrate assessments permitting the use of gen-AI, it must implement a curricular plan to ensure the fundamental skills and learning objectives are substantively taught and addressed throughout without gen-AI	132		FGW	2.
33	When planning assessments and associated workload, teachers should not assume students will use GenAI, unless the learning objective is related to the use of GenAI.	131		LUMC	18.
34	There should be a faculty-specific counsellor whose function is to promote AI literacy, and mediate discussions around Gen AI use and assessments	131		FdA	30.

Rank	Unedited rule	Score	Score distribution	Faculty	Edited rule
35	Use of GenAI and its potential impacts must be disclosed and addressed alongside ethics considerations in assessments concerning internships, thesis, and field research.	126		FdA	12.
36	Individual group member of a group assignment should always remain accountable for their individual parts when students have different views of GenAI use	111		FGGA	15.
37	When GenAI is used to generate questions and/or answer models, the output of the model should be reviewed and approved by at least two human assessors.	110		FSW	26.
38	Students must not be forced to use Gen AI.	107		FdR	8.
39	AI-detectors should not be used as an assessment tool/aid.	106		FdA	27.
40	Examiners should design obligatory assessment "checkpoints" that evaluate if the students can perform the transferrable skills targeted by the study program without the use of GenAI	105		FWN	3.

Rank	Unedited rule	Score	Score distribution	Faculty	Edited rule
41	Lecturers should explicitly include in the learning objectives of the course the mandatory use of GenAI in the assignments on e-prospectus	104		FGGA	21.
42	Course coordinators should provide a GenAI statement, that describes where and in what ways has GenAI been used in preparation of course materials	94		FSW	11.
43	Students should be allowed to opt-out of non-licensed AI use in any assessment products—both their own and the teacher's—for ethical reasons, without penalty.	93		LUMC	8.
44	At the start of the semester/programme, the students should be informed by the board of examiners of the learning objectives, assessment methods, and how the use of GenAI intersects/affects them.	88		FdA	22.
45	In principle, students bear joint responsibility for the use of Gen AI in group work, but in case problems arise, the teacher may apply discretion in assessing each group member.	83		FdR	15.
46	Lecturers should facilitate open and respectful discussions among group members when there are different opinions about GenAI among students	70		FGGA	14.

Rank	Unedited rule	Score	Score distribution	Faculty	Edited rule
47	A course grade should be based on at least one AI-proof assessment method.	66		FSW	5.
48	Lecturers should compose their combination of assessments in such a way that at least 50% of the grade is derived in circumstances where students have no access to GenAI. This part of the assessment should lead to a pass (5.5 or higher, so no compensation is possible). This rule does not apply for assessing course objectives involving GenAI literacy.	46		FSW	4.
49	Students must have the right to decide whether to use GenAI, as long as its use complies with course learning objectives and rules (academic integrity)	30		FGGA	6.
50	Teachers may use GenAI only when it is meaningfully disclosed in the E-prospectus so as to enable students to not pick the course in question should the students wish to not have their work (partially) graded by GenAI.	21		FdR	23.

Appendix II: Similar reports

This appendix lists other existing reports, education policy frameworks or guidelines for GenAI in education and/or educational assessment. The list below is not exhaustive, and it is only meant to provide an indicative sketch of the national and international initiatives for regulating or guiding GenAI access, acquisition or use in higher-education assessment practices.

University of Amsterdam/Teaching and Learning with Artificial Intelligence in Higher Education (TALAI). (2025). ***TALAI Policy Guidelines for Assessment in the Context of GenAI in Higher Education,*** **available at:** https://www.talaiproject.eu/_files/ugd/8fb8f6_d810eeb58ac84a6d8389a2f09d7342e3.pdf

The guidelines put forward seven sui-generis principles for guiding and refining assessment practices at UvA: human oversight in grading, quality control of AI-generated content, alignment with learning objectives and outcomes, revision of existing policies, transparent communication of grading criteria, and transparency in GenAI utilization. The principles are linked to fourteen specific guidelines for GenAI use in assessment.

University of Amsterdam. (2025). ***Policy Framework and Guidelines on GenAI in Education*****, available at:** Policy framework and guidelines on GenAI in education - University of Amsterdam

The framework takes responsibility as the central value for GenAI use inside the university, it puts forward five broad guidelines applying to all GenAI at UvA, and one general guideline theme on GenAI and assessment.

Center for University Education (Brookings). (January 2026). ***A New Direction for Students in an AI World: Prosper, Prepare, Protect*****, available at:** A-New-Direction-for-Students-in-an-AI-World-FULL-REPORT.pdf

The report is mainly targeted at pre-university children, but contains 12 broad recommendations that are relevant for higher education. The report notes that, currently, risks overshadow the benefits when it comes to AI use in education.

Tertiary Education Quality and Standards Agency (TEQSA). (November 2023). ***Assessment Reform for the Age of Artificial Intelligence*****, available at:** Assessment reform for the age of artificial intelligence

The report draws on principles concerning AI use, and presents five propositions for beneficial AI use in the context of assessment. The report also discusses ways to reform higher-education assessment with an eye to (Gen)AI technological developments.

Bassett, Mark. (2025). ***S.E.C.U.R.E. GenAI Use Framework for Staff, Charles University*****, available at:** S.E.C.U.R.E. GenAI Use Framework for Staff

The framework is developed to help university staff to make informed decisions about GenAI software use. It classifies GenAI use risks into 6 categories listed in the acronym of the framework: Security credentials, Ethical use, Confidential information, Use of personal information, Rights protection, Evaluation of outputs.

University of Melbourne. (2025). ***Assuring Learning in the Age of GenAI: Guidelines and Principles for Assessment*****, available at:** Assuring-learning-in-the-age-of-Generative-Artificial-Intelligence-.pdf

The guidelines contain two main assessment principles (Principle 1: "At least 50% of the total mark in a subject must be based on secure assessment types." and Principle 2: "There may be exceptions to the subject-based Assessment Principle 1 where a form of 'programmatic' assessment is adopted within a programme") as central requirements for redesigning assessment in the context of proliferating GenAI use.

Notre Dame University. (2023). *Statement on Generative AI and Academic Integrity*, available at: Statement on Generative AI and Academic Integrity (May 2023) | Undergraduate Academic Code of Honor | University of Notre Dame

The webpage also contains a set of AI Recommendations for Instructors, and an Undergraduate Code of Honor for students, both working as rules and guidelines concerning GenAI use in academic activities.

O'Sullivan, J., Lowry, C., Woods, R., & Conlon, T. (2025). *Generative AI in Higher Education Teaching & Learning: National Policy Framework*, available at: Principles for Ethical AI Adoption

This report proposes five sets of recommendations for GenAI adoption and integration in the Irish higher-education sector, including, among others, teacher and student responsibility for GenAI use, procurement rules for GenAI services, and creating and maintaining a public register of approved AI tools.

Slotnick, Ruth C., Shea, Peter et al. (2025). *GenAI in Assessment: A Practical Guidebook for Massachusetts Public Higher Education*, available at: GenAI Assessment Guidebook.pdf

The guidebook proposes five core principles for GenAI implementation in higher-education assessment. These principles are: using AI to enhance, not replace human judgment, ensure transparency in GenAI use, centering equity at different assessment stages and dimensions, maintaining contextual relevance, and the support of continuous professional development.

Holmes, W., & Miao, F. (2023). *Guidance for generative AI in education and research.* Unesco Publishing.

The report lists a series of policy proposals and measures for how to use GenAI in both education and scientific research, such as ensuring sufficient social interactions among human members of the academic community, auditing GenAI outputs (especially on questions of bias), and using AI tools to reduce exam pressure.

UNESCO. (2025). *AI and the Future of Education: Disruptions, Dilemmas and Directions.* UNESCO, available at: IA_Education_UNESCO_EN.pdf

This anthology contains critical position pieces and advisory notes on how assessments can be redesigned and GenAI used in the post-2022 context, including refocusing on dialogical assessment approaches, design constraints set on GenAI service procurement, and auditing AI use in the educational assessment context.

Kovanović, V., Barthakur, A., Joksimović, S., & Siemens, G. (2025). Why universities need to radically rethink exams in the age of AI. *Nature, 648*(8092), 35-37.

This comment/position paper singles out some strategies for how universities should organise assessments in the context of increasing chatbot use. These strategies include continuous assessment, valuing higher-order skills and developing clear policy guidelines.

Jin, Y., Yan, L., Echeverria, V., Gašević, D., & Martinez-Maldonado, R. (2025). Generative AI in higher education: A global perspective of institutional adoption policies and guidelines. *Computers and Education: Artificial Intelligence, 8*, 10034

This article is a review of university policies for GenAI adoption across forty universities in six global regions. The findings show that universities are emphasizing academic integrity, enhancing teaching and learning practices, and equity in education.

Universiteit
Leiden
Governance and Global Affairs