Supporting the ethical and equitable use of Generative AI (GenAI) for Formative and Summative Assessment: a transnational perspective

Methodology

As one of the main aims of the project was to gain a profound understanding of the perceived views of university students and their usage of GenAI in their assessments, a qualitative approach was adopted, and data were collected using focus groups.

The sample population included all students (e.g., UG, PG and PhD level) studying at one of the partner institutions at the time when the focus groups took place. We adopted purposive sampling method, as the essential criteria was that the participants must be using/or have used GenAI tools for formative and/or summative assessments. The participants were recruited via email and VLE, using a universal template. As an incentive, each student was given an amazon voucher (£15) upon the conclusion of their participation. To keep consistency, the template was produced collaboratively by partners. In total, 80 participants took part in our research. The table below lists the detailed breakdown among the partners.

Data collection took place between March to September 2024. With regard to ethics, the initial plan was for the project lead applying for an ethical approval at their institution on behalf of all partners. However, due to a lack of existing institutional agreement covering all partner institutions, as well as the time constraint for getting one arranged in time, all partners ended up gaining their ethical approval locally, and each partner was responsible for planning, organising and running their own focus groups within their institution. To ensure consistency and thematic consistency, one partner (Huddersfield) drafted a list of initial principal interview questions. They were checked, modified and approved subsequently by all partners.

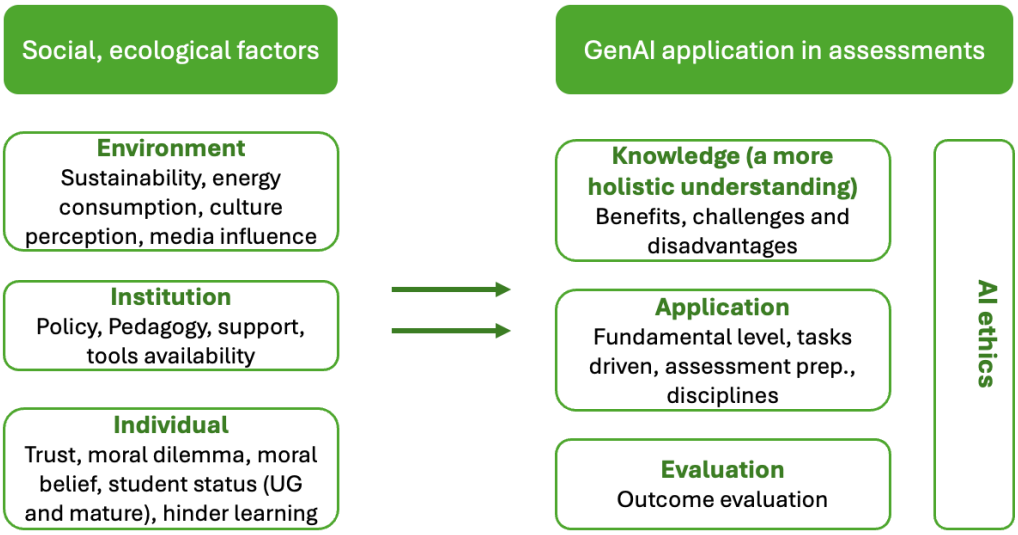
Permission was asked to record the focus group discussions and was granted each time. In total, we have conducted 20 focus groups across all partners.

Partners	Focus groups conducted	Participants in total
King's College London	3	15
Imperial College London	7	27
University of Huddersfield	5	22
University of Birmingham	5	16
Total	20	80

Data analysis

As the focus groups took place across multiple months, partners conducted transcriptions and subsequently prepared for data analysis alongside the data collection. The actual data transcription and analysis were carried out primarily by our student helpers, under the supervision, and also with the support of the project leads at the individual partner institutions. Each partner institution recruited 2 (or 4) student partners (e.g., undergraduate, postgraduate taught and PhD students) to work on the QAA project.

The partners then adopted a thematic analysis method to identify and code themes (Braun & Clarke, 2006) individually. This method aimed to complete six steps: data familiarization, data coding, theme searching, thematic review, defining each theme, and naming themes. Once the initial themes were generated locally, the partners shared the themes, and subsequently compared, summarised and/or synthesised the themes together (fig.1).



Findings and themes

We grouped our themes largely into two categories: social, ecological factors and GenAI application in assessments. The first category contains themes that impact students' understanding and applications of GenAI in their assessments. The second category describes the themes relating to their actual usages. Please note that the AI ethics is integrated into other themes and is not presented as a standalone theme.

Environment

Energy consumption: environmental awareness / concerns about energy usage

Media influence: students see AI as embedded in wider society, economy and their future careers, so note that the institution embracing its use in academia reflects the age in which we live and prepares students for future work including activities that use AI. The discourse around AI in media (e.g. news reporting) can negatively affect student views of ethical AI use, and this discourse is shown to feed into student perceptions of ethical AI use in academia.

Institution

Policy, training and support: Students seek more clarity at all levels. Policies and guidance are mainly available at the macro level (institutional), there is a lack of attention and detailed instructions (including resources and examples) at the meso (faculty) and micro levels (departmental). Resources should be in one single place, rather than being dispersed across various locations. Training with detailed case studies and examples needs to be provided to students and staff.

- "It's not often told to students, […] from my understanding of what one of my lecturers said, saying that we're allowed to use it. However, they didn't tell us in depth how to use it and where it's just not appropriate. […] More guidance should be provided by the university in terms of how to use it, where to use it, and where not to." (UoB - P2)

Need for practical guidance: High demand for practical examples of ethical AI use, participants suggested that universities offer workshops on using GenAI tools, focusing on areas like reference management, grammar, and content structuring. This could help normalize AI usage and equip students with the necessary skills to use AI responsibly.

- "I think the policy is you have to acknowledge it to some extent, but we're not told how to do it properly. So I just don't. I just don't claim any use of it, if everybody else is not claiming it. If I claim it, I fear it may probably get me into some trouble if I do it improperly. So yeah, you know, that's why I just don't say that I use the tool at all." (UoB - P3)

Influence of departmental and institutional beliefs: departmental and institutional perceptions of AI affect student attitudes towards AI. Disciplines that encourage AI as a useful tool for student learning through incorporating specific assessment tasks are noted by students as encouraging ethical and responsible student attitudes towards AI use. Simultaneously, students describe vagueness around departmental and institutional AI use and guidelines, leading to student confusion around what is ethical and what is in line with academic integrity.

Some students believed staff AI use is unethical as it is staff responsibility to provide information directly, such as from their own bank of knowledge rather than using AI to help with tasks. This suggests beliefs of learning as a transfer of content from teacher to student. It may also suggest less explicit student understanding of the rationale of academic tasks in how the task furthers student learning. Other students believed equitable use means parity for students and staff, so staff should use AI with the same criticality and transparency that students do. Students appear to value respect and transparency between expectations for their own AI use and staff's AI use. In this way, equitable AI use means clear communication between students and staff on AI literacy and use, transparency and parity.

Individual

Hindering learning: turning to AI to help develop their work stifles learning, as they are handing off the work to AI instead of doing it themselves. Specifically, they are resistant to using it in preparation of the assessment itself because they see it as presenting a barrier to their own learning/learning things for themselves.

"We have some nice quotes here along the lines of "I'm paying £9000 a year to learn, so why would I get AI to do all the work for me"

- *"It kind of takes away the real purpose of university, especially for the students that just use it [GenAI] as an answer. It doesn't allow them to add any like thought or depth into the assignments." (UoB - P9)*
- *"So I believe the assessments that are given to us are for us. We are here to develop our own thinking and thought process. So if we just give this the pass to the AI and just get it to generate the output and give it back or submit it, then what are we doing here? That's the point where I draw a line. So anything that comes between you and your personal development, you should draw a line over there. Nobody will know if we have used AI or not. That's a different thing. But we are here for something bigger, and we should develop our own thinking." (UoB - P12)*
- *"The easy way to make content with AI can make some students avoid real learning and not develop their skill, which hurts their education and breaks university rules about honesty." (UoB - P13)*

Trust and reliability issues: the findings seem to cover multiple themes, including trust, knowledge and application. Participants expressed a lack of trust in GenAI-generated content, noting that it often requires extensive manual revisions. There is particular concern about GenAI's handling of quantitative data, where accuracy is perceived to be lower compared to qualitative outputs.

Perhaps because of the criticality needed during AI use that was noted, trust of AI and others (staff, peers) is shown to be complex and at times hierarchical for students. At times students trusted academic staff and peers more than AI, but other students trusted AI the most. Some students noted that AI use has a direct effect on their self-perception. This includes questioning their own knowledge, needing to refer to course notes or scientific journals to verify AI-provided information, or ending up wasting time pursuing inaccurate information provided by AI and feelings of frustration. The student consensus is a general distrust of AI to do more complex, higher-order tasks (such as more creative production) and more trust in AI to do lower-level

tasks (such as fact recall). This is supported by cases where students use AI for initial idea generation or summarising information for background research, before continuing with more productive tasks independently.

- "Sometimes it's scary that the future generation will not be as smart as we are, like street smartness, or like having the presence in the room and awareness of what is going around us. Right now, if you see the GenZ, they are on their phones all the time to search for everything. It's not that they hit onto the libraries or textbooks first to search that. Is it reliable enough? No, because I think the age is gonna change soon […] we are entering the age of the world wide web 2.0." (UoB - P6)
- "I've never used it to generate text. I basically don't trust it to give me the right answer". (UoB - P12)
- "What I see there is always it's trying to hallucinate and come up with answers which has nothing to do with my question that I have asked and it gives me reference links which goes nowhere just gives error. 404, not found. So that's the problem that I've seen […]." (UoB - P1)

Student status: the former mainly focuses on assessments themselves, whist the latter focus on knowledge/learning gain.

Societal and ethical implications, accessibility: AI use is illustrated as equitable by some students as it can be accessed freely by anyone on any device. AI is also approachable as its conversational style is informal, showing that AI can lower the affective barrier to asking questions, which students may experience with academic staff. However, paid AI may only be accessible to those from specific socio-economic backgrounds who can afford to spend money on more sophisticated versions of GenAI, resulting in inequitable use.

Knowledge

Efficiency and time-saving: the overall theme which captures all usage purposes mentioned by users is the technology's time-saving benefit. GenAI is frequently used for repetitive tasks such as generating research questions, creating reading lists, structuring essays, organizing thoughts, and drafting content. Participants also use GenAI to summarize articles, note-taking, revision, and idea generation for various academic tasks, allowing users to quickly generate ideas and streamline tasks.

- "I use it every day. And I feel like my life is basically dependent on this. [...], maybe my brain is not coming up with the correct words, perhaps because I come from another country. Sometimes when I email, when I write a birthday card or [...] something, you know, I always use Generative AI to help me to paraphrase [...] being professional, being nice or polite or something. But it has actually really improved the way I write. And it's way much better. So basically, that's the reason why I use it every day." (UoB - P5)
- "When I came here, it was a bit difficult. I have changed my career path, and taking things from the university, understanding each and everything is very difficult. Asking questions to the professor is sometimes not possible, because there might be some silly questions or really tough questions. So, when it comes to AI tools, it is really helpful to understand in each and every question and I can ask the tools like whatever I like." (UoB - P13)

Data privacy: subject related, and differentiations between UG and PG - trust on openAI re data privacy, and unaware of the consequences of uploading their own writing. Subjects such as law, and psychology have more awareness than other subjects.

- "I don't have any privacy or security concerns." (UoB - P8)
- "I might be a bit hesitant to put my personal info into a generative ai tool." (UoB - P6)
- "I don't usually trust any apps or websites at all. I think they might somehow steal your data. But for me still I don't really mind." (UoB - P5)

- "I think it's kinda concerning to be honest, like the amount of data they store from these chatbots. [...] So that hasn't been properly addressed as of yet. I think there needs to be more effort to protect privacy. These chatbots need to be made more private and more secure in terms of data." (UoB - P10)

Ethical use and awareness: students are aware of the ethical implications in principle. They understand they should not be copying content directly into assessments. However, they are unclear about the meaning of equitable use, that is, what they are permitted to do relating to how they apply GenAI in assessments. Some students also connected AI use to less diversity in student academic work, as more assignments have a 'generic AI voice' and make similar points that AI commonly generates. AI use may also be less ethical if we are aware of potential biases in AI algorithms (such as privileging information about specific groups of people) and choosing to nevertheless use AI.

- "I feel that university should not endorse the use of generative AI this way because of ethical concerns. If educational institutions start trusting this type of AI, then the quality of education might fall. The educational institutions are already supporting us with different programming tools to work with like data analytical software […]. We trust our professors to give us the guidelines to work with. We trust them to teach as well, so we trust them to give us the resources we need. So I don't think that institution should actually endorse this type of license, because then there would be no necessity to do a master's degree for me. Because, yeah, then anybody like a 4th grader, who is still at school, can actually do the work if he has generative AI in his hand. You could like actually search how this econometrics analysis work, or what this article is taking about, or like how to do other complicated tasks. Yeah, that's simple. So if the institution actually start doing this, then the quality of education would actually fall. So it should depend upon the students whether they want to use it or not. But institutions should not provide it." (UoB - P6)

Prompt crafting (apriori assumptions): this is a common issue raised. The inputs often do not provide the desired outputs expected. Students seek more guidance and training in this area. The most needed area is clear and detailed examples, across different subject disciplines.

Application

A key part of ethical AI use and AI literacy is thus criticality of its application and use in various academic tasks.

Usage purposes: students use GenAI tools mainly for summative assessments – both at the fundamental (unsophisticated) and advanced level: help develop ideas at an early stage in project development; literature research, comprehending paper analysis, proof reading and brainstorming. The former seems to take more weight.

- "I use Gemini to help me to paraphrase." (UoB - P5)
- "I've also used it [GenAI] for coding. I've used it for just some simple code generation using Python." (UoB - P2)
- "During my assignments and during my general studies, I have to read lots of research papers. I have to go through business articles which are sometimes very lengthy. So what I use these days is ChatPdf. I can copy and paste the entire Pdf, and I then have a conversation about the context of this paper, for example, what are the insights or the findings." (UoB - P11)
- "I use it for summarizing articles, especially when there is not much time left before deadlines." (UoB-P10)

Relationship between AI use and type of assessment: views were mixed regarding ethical AI use and type of assessment. Students agreed that ethical AI use depended more on the type of task rather than the type of assessment (such as formative or summative). Some students were less concerned by issues surrounding ethical use or academic integrity in formative assessment and felt if work is ungraded, it is acceptable to use AI

as there are fewer external implications of AI use in formative assessment (such as unfair grading in summative assessments, which is a more significant problem) This can reflect more externally motivated or instrumental approaches to Higher Education experiences and student learning journey (i.e. where the learning process itself is less valued). Other students disagreed with this point, and felt using GenAI for formative assessment defeated its purpose, which is to experience the learning process and the 'doing' involved.

AI as a beneficial study tool: AI-supported tasks varied significantly in student accounts, from using AI to simplify academic concepts, for repetitive tasks to reduce time spent, or lower-order tasks that are systematic and require less student interpretation. Students note it can be useful for increasing productivity and lowering their cognitive load when doing repetitive tasks. This points to AI academic use as a tool for optimising cognitive processes, meaning to aid students in focusing on higher-order thinking and optimising their time for these higher-order tasks. These tasks are illustrated by students to be flexible as they can be personalised to own learning goals and study preferences, such as to create study plans or aid self-assessment. Multiple students note the efficiency of AI for re-drafting, such as changing the register or written text, writing documents for job applications or condensing writing. However, some students did highlight that they would take a critical approach to ensure their own voice came through in their work over the generic or neutral 'AI voice'.

AI use as a collaboration: ethical AI use is stressed by students as something to be built on and developed individually, as shown in the theme above on AI literacy and criticality. AI use is complementary in academic work as an additional source of information, rather than fully generative and agentic in its role. In this way, AI is seen almost as an additional member of the team, alongside academic staff, information sources (journals, course notes) and peers. AI was for example described by students as sometimes filling in for less academic staff contact time to generate answers to questions. Alternatively, students took on more of a trainer role that guides AI to become familiar with their assignment rubrics and generate feedback that is criteria-referenced.

AI as a mathematical support tool: students also mentioned AI as a tool, drawing on the invention of the calculator as a parallel. With the calculator there was a moral panic over it being the detriment to people's mathematical ability, when in reality it enabled faster working through the arithmetic grind allowing more time for the actual mathematical reasoning. These students were saying that AI is just a tool which might have the same impact, in that it handles the linguistic grind.

- "Sometimes it's difficult for me to understand mathematical concepts. So, I use ChatGPT to generate mathematical insights and answer some questions. That gives me the process step by step. It sometimes helps me solve complex integration and differentiation which is sometimes difficult to solve by myself." (UoB - P11)

Evaluation

Outcome evaluation: the data reflect varying perspectives across different local contexts. Regarding false information / hallucination. Another significant challenge is the generation of false or non-existent references, which necessitates manual verification, adding to the workload rather than alleviating it. However, data also show that some students recognise that AI has limitations (e.g. AI hallucinations, generating inaccurate information) in multiple academic tasks, such as coding or explaining scientific concepts. For this reason, students note AI generation should be used in conjunction with other information sources, compared, and used critically for ethical use in academic work, although they noted this process is a challenge of incorporating AI into your educational toolkit.