

GenAI and Academic Misconduct: Guidance for Colleagues

1. Context

This guidance provides advice for academic colleagues where they suspect academic misconduct involving the **(mis)use of Generative AI (GenAI) in an assessment**. This aligns to existing guidance – [specifically QA53](#) (Examination and Assessment Offences), and where relevant, [QA 53 Appendix 4](#) (Viva Voce for suspected cases of misconduct). This guidance has been prepared in consultation with colleagues in the CLT, Academic Registry, FARs, Education Managers and Associate Deans (Education). It also draws heavily on wider sector guidance.

Whilst the recent and rapid increase of GenAI tools such as ChatGPT has raised questions for the HE sector around the robustness of current assessment processes and, in particular, issues of academic integrity, it is important to note that the advent of such tools represents a *widening* of our current understanding of academic misconduct, rather than an entirely new form of breach. Certainly, where collusion and plagiarism may have been previously identified, this would traditionally have been in the form of colluding with *other students or individuals* (e.g. via essay mills). This has now broadened to include GenAI.

Moreover, during this academic year (23/24), as the technology has evolved, students may have received mixed messaging across a course around appropriate and permitted usage of GenAI, and this should be borne in mind when informing your academic judgement as to the extent and severity of any academic misconduct. To counter this, the University is moving towards all courses communicating expectations around the use of GenAI in each piece of coursework for next academic year (24/25), and from September 2024 the Academic Integrity Test will contain specific training and questions around GenAI.

Whilst we are in this transition phase, it will be important to **communicate expectations** around GenAI usage to students verbally or, in writing within lectures or seminars, so that students have clarity and transparency around the safe and appropriate use of GenAI across the course assessment portfolio. Students can be reminded of the [academic integrity statement](#) and any [relevant penalties](#) that could be applied. This will help to ensure greater clarity for students around the permitted use of GenAI in assessments, and in turn, help to ensure students cannot claim that they did not receive guidance on this.

2. Cases of Suspected Student (Mis)use of GenAI: Guidance Steps for Colleagues

The process of determining and processing a claim for academic misconduct can be summarised in four main steps:

- (1) Understanding Academic Judgement and Reasonable Grounds
- (2) The Marking Process
- (3) Investigation of an Offence/Determining Originality of Work
- (4) Review of Assessment Design

2.1. Step 1: Academic Judgement and Reasonable Grounds

First, we recommend that:

- You are advised **not to make** claims of misconduct based on “proof” of inappropriate use of GenAI, or statements such as ‘this sounds like it was written by AI’, as [humans have been proven to pretty bad](#) at identifying AI writing and this is often subject to unconscious bias which may single out specific groups of students.
- Instead, focus on *how misconduct evidences itself* in the work before you.
- This is because GenAI (mis)usage often manifests itself in ways that mirror [existing examples of misconduct](#), such as made-up references, paraphrasing without attribution, fabrication of results, and a sudden shift in tone or style.
- Consequently, detecting the misuse of AI necessitates the same skills that you are already using to ensure that student work is genuine.

In short, rather than build a case based on trying to evidence misuse of AI *itself* (almost impossible currently!), instead look for, and base your claim on, where this may have *manifested itself* in the assessment – you will likely find this mirrors existing examples of misconduct.

Second, the basis of any claim is:

- Your **academic judgement** – this is fundamental in forming a decision and formulating a course of action. This recognises [your](#) experience as a subject matter expert and marker.
- In the view of the [OIA](#), academic judgment ‘is a judgment that is made about a matter where **only the opinion of an academic expert is sufficient.**’
- You know your subject, the assessment task and the signs of academic misconduct; in marking work, you have gained great experience in evaluating academic voice, style, if the work is coherent and of the quality you would expect to see at this level.

- In forming your judgement, you will be establishing if you have **reasonable grounds** that the assessment is not the student's own work **and** if there is any evidence to support this in the work itself.
- There is no hard and fast definition of reasonable grounds, but generally this can be considered as **a set of facts or circumstances that would cause a person of ordinary judgment** to believe something beyond a mere suspicion.
- Remember: you are **not** required to prove misconduct **beyond** reasonable doubt/grounds; this is not a court of law, and the student is **not** on trial. **You are not** determining if the student is "guilty" – instead, there is a formal investigation where a case is suspected (the type of investigation depends on if you have textual evidence, or if you need to undertake further questioning. This will be explored further in the course of this guidance).

2.2. Step Two: The Marking Process

To support you in determining and evidencing any claim of suspected misconduct, we set out some tips for **potential indicators of misuse** to look out for when marking work in Appendix 1.

In gathering this evidence, you **must not** submit students' work into a so-called GenAI detector tool – this is potentially illegal, and these detector tools *are fundamentally flawed*, they *are not fully developed*, and have proven to be unreliable in the *specific context of HE*. You **cannot** build a case on reasonable grounds if you rely on these tools, as they produce invalid responses.

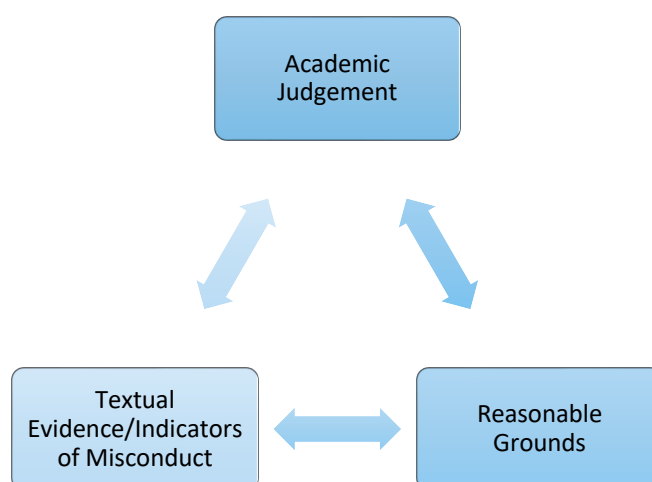
Instead, **you may**:

- Upload the *assessment question* into GenAI tools to gauge the style and quality of output it generates, but it is important to realise that this is not failproof, and the answer it generates will depend on the quality of the actually prompt and the specific tool being used (e.g. ChatGPT 3.5/4/Claude Anthropic/Google Gemini, etc).
- Undertake an online search of well-known essay mills sites and e-cheating sites (e.g. Chegg) for the assignment, or passages of text within the assignment.

2.3. Step Three: Investigation of an Offence/Determining Originality of Work

In order to progress with a formal investigation of academic misconduct, you need to reflect on three interconnected factors (Fig 1):

Figure 1: Forming a Claim of Academic Misconduct

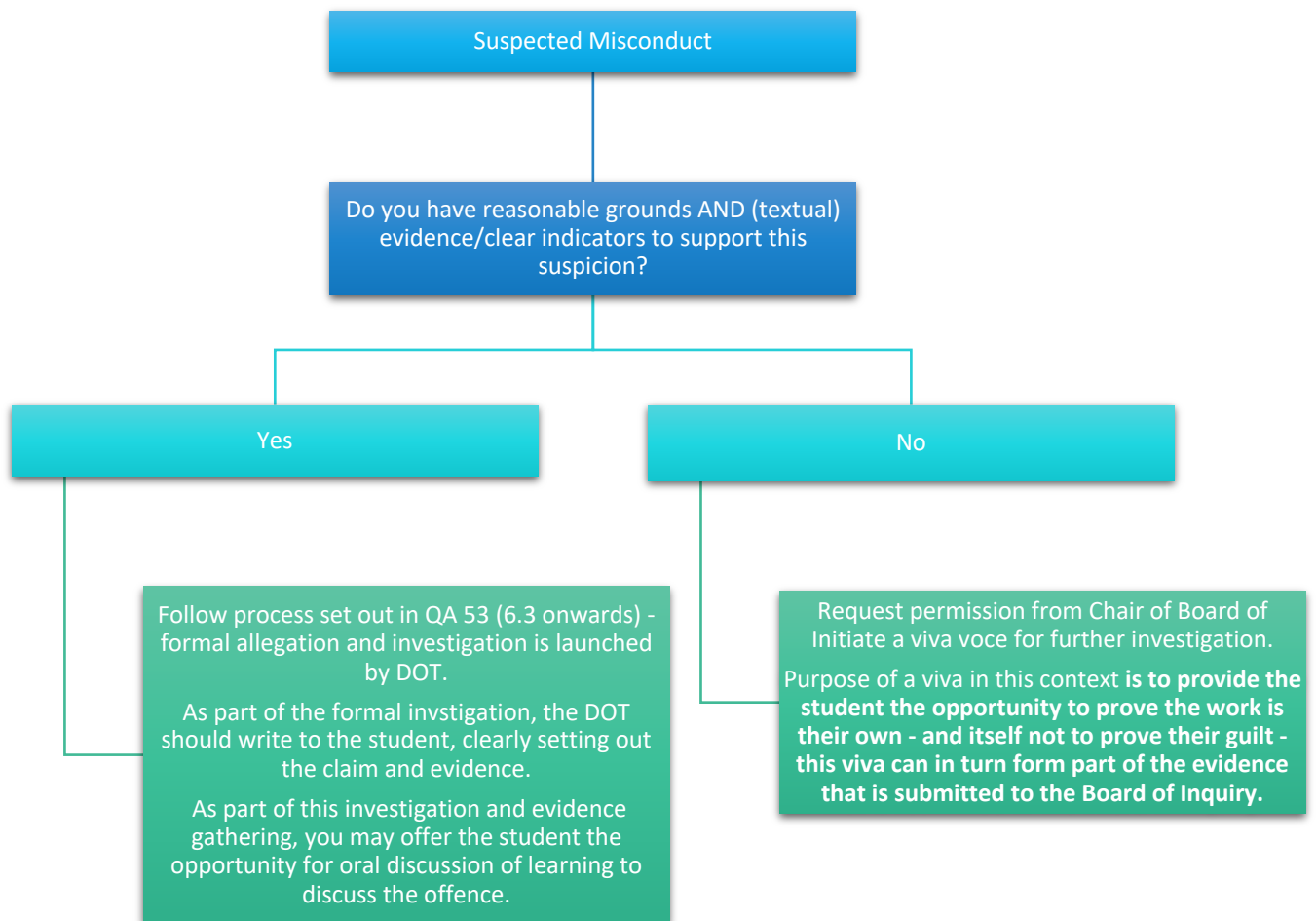


The route you go down will depend on if you have ‘reasonable grounds on the evidence before you.’ The full process for both routes is set out in QA53, but an overview is set out, and depicted in Fig. 2, below.

Generally, this means **can you substantiate your claim** with (textual) evidence – e.g. clear evidence of made-up referencing/lack of references/a watermark for content (a lot of the image generation tools now include some form of watermark in the metadata) in the student’s work that has not been attributed to AI:

- In cases where you **have textual evidence** that meets reasonable grounds in your academic judgement, a formal claim and investigation can be launched, led by the HOD/DOT.
- As part of this process, an **oral discussion** can be offered to the student to discuss the alleged offence. This oral discussion/ presentation of learning can itself form part of the evidence gathering of the investigation. This is similar too, but nonetheless distinct from, the option of a viva voce.
- Alternatively, if you are struggling to find (textual) evidence that would meet the threshold of ‘reasonable grounds’, or need to investigate further in order to fully inform your academic judgement, the University has confirmed that a viva voce may be held, with permission of the Chair of BOS, to determine the authorship of the work (See [Appendix 4 of QA53](#)).
- Tips for conducting a viva voce are set out in **Appendix 2** below.

Fig. 2: Process for Investigating Suspected Claims of Academic Misconduct



2.4. Step 4: Review of Assessment Design

If a number of cases emerge associated with a particular assessment or unit, it is worth considering whether the assessment itself could be adapted or developed to reduce risk and ensure the assessment remains robust. The CLT can support colleagues with the development and enhancement of assessments to help ensure students use GenAI in ways that are ethical and appropriate and align with the graduate attributes associated with your subject discipline.

Appendix 1: Textual Evidence/Indicators of Potential Misconduct

Note: the following indicators are **not proof** a student has cheated, but may lead you to question the authenticity of the work. Depending on your academic judgement, and the extent to which misconduct is evident in the text, you may feel this is enough to raise a formal claim, or that it requires further investigation – see Step 3 above.

- a) Dramatic improvement in the quality of the student’s response compared to earlier assessments (assuming identity of student is known – **see note 1 below**).
- b) Different styles or voices in the text, and/or text that is similar to other students or an AI-generated answer (**see notes 2 & 3 below**).
- c) Lack of references, made-up references, poor paraphrasing, or references that students would not otherwise expected to have access to (**see note 4 below**).
- d) The work is not reasonably likely to be completed by a student at that level due to the high quality of the arguments being made or the sources used.
- e) Instances of incorrect/inconsistent use of first-person and third-person perspective where generated text is left unaltered.
- f) The inadvertent inclusion by students of warnings or provisos produced by AI.
- g) Generic terms or content, and different spellings (UK/USA) or unusual formatting.

Note 1: Dramatic Improvement in Style or Quality of Response

Staff may note where there may have been a dramatic improvement or distinct change in style in the student’s writing capabilities, to the extent that, based on their academic judgement, the student’s work looks very different from earlier submissions. This change would need to be significant and notable as the general expectation is that student work via assessment should demonstrate improvement over time as students learn from and action their feedback.

Note 2: Identical or Significantly Similar Passages of Text Submitted by Another Student

GenAI tools can produce text which is very generic in nature and has a similar bland ‘tone’ (especially where students have poor prompting skills). Whilst such tools can create very different responses based on the same prompt, you may notice nonetheless that similar passages of text appear in multiple student responses. Of course, this may simply indicate a ‘standard’ case of collusion, but that would still represent academic misconduct.

Note 3: Identical or Significantly Similar Text Produced by a GenAI Tool

You may notice that when you type the assessment question into a GenAI tool that it produces text which is very similar in nature to a paragraph or passages of text that a student has included in their response. Whilst this is not definitive evidence of (mis)use of

GenAI, you may wish to discuss these passages of text with the student to determine if this is their own work.

Note 4: Made-up References or Paraphrasing without (Correct) Attribution

Whilst a student including made-up references, or large passages of text without any references, is not evidence *per se* of (mis)use of GenAI tools, it may, depending on the level of falsification and volume/significance of the issue, constitute an assessment offence.

- *Note* – earlier versions of GenAI tools (such as GPT3.5 and earlier releases of GPT4) were particularly prone to generating false references.
 - i. Research suggests this is becoming less of an issue (particularly as LLMs are linked to the live internet and other research tools), but it can still be a feature of such tools.
 - ii. This is not specific to GenAI tools, but you may see evidence of students including significant passages of text which are unattributed. This may be very generic in content and style.
 - iii. In such cases, you may use academic judgement to determine if an offence has been committed (or if it is simply poor academic practice). The offence would **not** be that a GenAI tool has been used, but that text which forms part of their argument is **unattributed and/or falsely attributed**.

Appendix 2: Tips for Conducting a Viva Voce

- **You should only** use the meeting with a student as an opportunity for discussion and exploration in order that they can demonstrate that the work is their own.

It is important to remember, that we are in a transition phase where both students and staff are adjusting and adapting to these new tools. Collusion and/or plagiarism is not always a conscious intent and can be the result of uncertainty, a lack of confidence or clarity. For instance, if it is determined that the student's work **is their own**, you may also wish to use this meeting as a developmental opportunity for the student to develop their learning and academic conduct.

- **You should** draw on the learning outcomes of the assessment in preparing for the meeting.

Use these as a starting point to formulate questions for the meeting which encourage the student to focus on the *process* of the assessment. E.g. you could focus the discussion on higher order thinking skills, such as their ability to reflect or problem solve ("you wrote the following – tell me more about this and how you came to this conclusion. What did you read that informed your approach, and how did you go about addressing any gaps in your knowledge").

In short – do they **understand** what they have written and **how** they came to their conclusions, rather than simply do they **recall** what they wrote.

- **You may** use the opportunity to ask questions which probe the quality of their referencing.

For example, "I noted you cite Anon Smith's work. Can you tell me more about this – where did you access this as it is not in our Library; did you know the work does not exist, and so why did you cite it?"