

Managing Learner Use of Artificial Intelligence and Preventing Misuse

V1.0

Purpose

Even though the potential for learners to misuse artificial intelligence (AI) is new, most methods to prevent this misuse already exist; training centres will have measures in place so learners understand the need to submit their own original work for assessment and to identify possible cheating. This guidance serves as a reminder to tutors and assessors about following best practices in this area as it relates to AI use.

Main Points:

- As has always been the requirement, and per SFEDI Awards Approved Centre Criteria, any work given for qualifications must be the learner's own.
- If learners misuse AI such that the work they give for assessment is not their own, they will have cheated, according to SFEDI Awards rules, and may face strict penalties.
- Learners and centre staff need to be informed about the risks of using AI and what constitutes cheating.
- Learners must ensure the work they give for assessment clearly shows it is their own. If any parts reproduce content directly from an AI, the learner must identify those parts and understand they will not demonstrate meeting criteria independently, so will not be rewarded (see Acknowledging AI Use below).
- Tutors and assessors may only accept work they believe is the learner's own.
- If tutors suspect learner work given for assessment contains AI-generated sections not acknowledged, they must investigate and take suitable action.

SFEDI Awards staff and moderators have established procedures to identify, report and investigate cheating involving learners, including AI misuse. SFEDI Awards continues monitoring developments in this area and will update this guidance appropriately.

Understanding AI Use and Associated Risks

AI use means learners obtain information/content through AI to potentially use in assessed work. While AI tools and capabilities will likely expand greatly soon, misusing AI tools for qualifications at any time is cheating. Learners and teachers should also know AI tools are still in development and often have limitations, like inaccurate or unsuitable content.

AI chatbots are AI tools that generate text in response to user prompts and questions. Users can ask follow-up questions or have the chatbot revise the responses provided. AI chatbots respond based on patterns in their training data sets (large language models). They produce responses statistically likely to be relevant and appropriate. AI chatbots can:

- Answer questions
- Analyse, improve, and summarise text
- Author essays, articles, fiction, nonfiction
- Write computer code
- Translate text between languages
- Generate new ideas, prompts, or suggestions on a topic
- Generate text with certain attributes like tone, sentiment, formality

Current AI chatbots include:

- ChatGPT (<https://chat.openai.com>)

- Jenni AI (<https://jenni.ai>)
- Jasper AI (<https://www.jasper.ai/>)
- Writesonic (<https://writesonic.com/chat/>)
- Bloomai (<https://huggingface.co/bigscience/bloom>)
- Google Bard

There are also AI tools that generate images, such as:

- Midjourney (<https://midjourney.com/showcase/top/>)
- Stable Diffusion (<https://stablediffusionweb.com/>)
- DALL-E 2 (OpenAI) (<https://openai.com/dall-e-2/>)

Using AI chatbots poses significant risks for learners completing qualifications. As noted, they produce responses based on language statistically likely to be appropriate so cannot be relied upon. AI chatbots often provide seemingly convincing but incorrect or biased answers. Some have been found to give dangerous and harmful responses. They can also fabricate references to books/articles by real or fictional people.

What is AI Misuse?

As per SFEDI Awards Approved Centre Conditions, learners must submit their own original work for assessments. This means the final product must be in their own words, not copied or paraphrased from other sources like AI tools, and the content must reflect their independent work. Learners need to demonstrate their own knowledge, skills and understanding required for the qualification as detailed in its specification. This includes showing their competence regarding the assessment objectives for the subject related to the questions or other tasks set. Any AI use where learners have not independently proven their own achievement will likely be considered cheating. While AI may become an established workplace tool in the future, for demonstrating qualifications knowledge, understanding and skills, it is important that learners not rely on tools like AI, and develop subject knowledge themselves.

AI tools should only be used when learners can show the final submission is the product of their own independent work and thinking.

Examples of AI misuse include, but are not limited to:

- Copying or paraphrasing AI-generated content so the work is no longer the learner's own
- Using AI to complete assessment parts so the work does not reflect the learner's analysis, evaluation or calculations
- Failing to acknowledge AI tools used as an information source
- Incomplete or poor AI acknowledgement
- Submitting work with intentionally incomplete or misleading references/bibliographies

This constitutes cheating as defined in the SFEDI Awards Malpractice and Maladministration Policy.

Centre Guidance on AI

Centres will already have agreed assessment policies and procedures to ensure authenticity. These must now also address AI misuse risks.

Tutors, assessors and other staff should discuss AI use and agree on managing learner AI use at their school/college/centre. Centres must make learners aware of appropriate and inappropriate AI use, associated risks, and potential consequences of misuse on a qualification assessment. They should also inform learners about the centre's plagiarism approach and cheating consequences. Centres may communicate with learners to increase awareness of risks/issues and ensure they follow the centre's policy.

Centres should:

- a) Explain the importance of learners submitting their own independent work for assessments and emphasise AI misuse risks.
- b) Update the malpractice/plagiarism policy to acknowledge AI use (e.g. what it is, associated risks, what constitutes misuse, how this will be treated as cheating, appropriate use cases, how it should be acknowledged) – referencing this document is simplest.
- c) Ensure the policy includes clear guidance on referencing (including websites).
- d) Include clear guidance on acknowledging any AI use to avoid misuse (see Acknowledging AI Use below).
- e) Ensure teachers/evaluators understand AI tools, associated risks, and detection tools (see What is AI use and Associated Risks? and What is AI Misuse?).
- f) Consider requiring learners to sign a declaration that they understand what constitutes AI misuse and that it is prohibited.
- g) Remind learners of the importance of their (electronic) declaration that work submitted is their own, consequences of false declarations, and that they have understood and followed subject requirements.
- h) Note that awarding organisation staff, examiners and moderators have established malpractice reporting and investigation procedures.

Acknowledging AI Use

It remains essential that learners understand the importance of referencing sources used when producing assessed work, and how to do so properly. Appropriate referencing demonstrates academic integrity and is key to maintaining assessment integrity. If a learner uses an AI tool that provides details of the sources it used to generate content, the learner must verify those sources and reference them in their work normally. Where an AI tool does not provide such details, the learner should independently verify the AI-generated content, then reference the sources used. Additionally, where learners use AI they must acknowledge its use and clearly show how they used it. This allows tutors/assessors to review the AI use and determine if it was appropriate for the given assessment. This is especially important as AI-generated content does not undergo the same academic scrutiny as other published sources.

Where AI tools have been used as an information source, the learner's acknowledgement must show the AI source name and when the content was generated - e.g. ChatGPT 3.5 (<https://openai.com/blog/chatgpt/>), 25/01/2023. The learner must retain a copy of the question(s) and computer-generated content for reference and authentication purposes, in a non-editable format like a screenshot, and provide a brief explanation of how they used it.

This should be submitted with the work so the tutor/assessor can review the work, AI-generated content, and its use. If this is not submitted, and the tutor suspects AI use, they should consult the centre's malpractice policy for next steps and take action to confirm the work's authenticity.

Other recommendations around acknowledging AI use:

- a) Remind learners that poor referencing, paraphrasing, and copying text may constitute cheating, attracting severe penalties like disqualification. For AI, learners must understand what is and is not acceptable when acknowledging AI content/sources - simply referencing "AI" or "ChatGPT" would be unacceptable, just as merely stating "Google" rather than the specific sites/pages consulted.
- b) Remind learners that if AI use means they have not independently met assessment criteria, they will not be rewarded.

Identifying Misuse

Identifying learner AI misuse requires the same skills and observation techniques tutors likely already use to confirm authenticity of work. Some methods we recommend are:

Comparing to Previous Work

When reviewing work authenticity, compare against other work by the learner. For written work, note characteristics like:

- Spelling/punctuation
- Grammar
- Style/tone
- Vocabulary
- Complexity/coherency
- Understanding/level
- Production method (handwritten/typed)

Compare to work done in class or under supervision.

Potential AI Indicators

The following may indicate AI misuse in learner work:

- a) Default use of American spelling, terms, localisations
- b) Language/vocabulary not suitable for qualification level
- c) Lack of expected direct quotes/references
- d) References that can't be found/verified (some AI tools fabricate references)
- e) No mention of events after a certain date (limits of AI training data)
- f) Incorrect/inconsistent perspective use where generated text wasn't altered
- g) Language style differs from classroom work
- h) Variation in language style within a piece if AI text was amended
- i) Lack of expected graphs/tables/visuals
- j) Lack of local/topical knowledge
- k) Generic content not specific to learner/task/scenario if required
- l) Inclusion of AI-generated warnings about its limitations or hypothetical nature
- m) Typed submission when handwritten is typical
- n) Unusual repeated conclusions/essay structures in lengthy essays (AI asked to add depth)
- o) Strongly stated non-sequiturs or confidently incorrect statements
- p) Verbose/hyperbolic language unlike learner's usual style

Automated Detection

AI chatbots produce content by predicting the most statistically likely next word, unlike human writing with word variety. Some programs and services use this to analyse written content and determine the probability it is AI-generated:

- OpenAI Classifier (<https://openai.com/blog/new-ai-classifier-for-indicating-ai-written-text/>)
- GPTZero (<https://gptzero.me/>)
- Giant Language Model Test Room (GLTR) (<http://gltr.io/dist/>)

These tools could check learner work and/or verify authenticity concerns. However, they base scores on word predictability, so will score amended AI content lower. Detection quality varies and AI/tools continue advancing. Use tools as part of a comprehensive approach to work authenticity, considering all available evidence when reviewing potential cheating concerns.

Suspected AI Use

SFEDI Awards performs extensive plagiarism checks on all work, now including AI detection. If we suspect incorrect AI use in assessment submissions, we will investigate and impose sanctions per policy if necessary, ranging from warnings to disqualification and temporary qualification bans. Actions also include centre sanctions where staff knowingly accept or fail to check inauthentic work.

Contact Us

If you have any queries about the content of this policy, please contact our customer service department.

SFEDI Awards,
19 Victoria Road,
Darlington
DL1 5SF

customerservices@sfediawards.com

0845 224 5928

@sfediawards