OCR
Oxford Cambridge and RSA

OCR's response to the DfE call for evidence: Generative artificial intelligence in education

August 2023

OCR (Oxford, Cambridge and RSA) is a leading UK awarding body. We provide engaging GCSEs, A and AS Level and vocational qualifications in a wide range of subjects to equip students with the knowledge and skills they need for their future, helping them achieve their full potential.

Our heritage is long established and provides us with expertise across academic and vocational qualifications. Since becoming established in 1998 our experience, knowledge and skills have enabled us to build a reputation for reliably high standards.

As part of Cambridge University Press & Assessment, we have access to unrivalled expertise and research capability across assessment and examinations.

Cambridge Assessment International Education - the largest provider of international qualifications for students up to 19-years-old, and Cambridge English Language Assessment - experts in English language assessment are our sister organisations.

Questions

Experience with generative AI:

Question 10

Have you or your institution used generative AI tools in an education setting? If so, could you briefly describe the ways it was used, and the specific tools used.

Within OCR and across Cambridge University Press & Assessment, we are developing a range of applications for generative artificial intelligence. Some are still at an experimental stage, with the technology still very much in its infancy. Others, like those in English language learning, build on decades of AI and Machine Learning application. We are continuously developing key principles and practices to govern robust and ethical deployment of AI, and have published some on AI and assessment, and on AI in academic research publishing. Schools, teachers and students face practical questions that can impact on access, equity and integrity.

Question 11

What were the main challenges you faced in using generative AI and how did you address these?

Challenges using such tools include finding ways to ensure that they comply with the five principles outlined in the government's AI Regulation white paper, in particular that Generative AI tools are: safe, secure and robust; transparent and explainable; fair; accountable; and feature sufficient forms of redress. Not all Generative AI tools are currently transparent about the sources of input data or the process used in producing outputs, the other principles are hard to achieve.

Question 12

What was the result of your use of these tools, including any impacts?

This is covered in our responses to Questions 10 and 11. It will be interesting to see responses from schools and colleges who have greater freedom to use generative AI tools in the classroom and beyond.

Opportunities and benefits:

Question 13
How do you think generative AI could be used to improve education?

Our society and economy will need people who are able to use AI discerningly. More important than the latest large language models themselves will be the critical skills and thinking that underpin their use: from data science, public health and education to politics, culture and green tech.

AI can be used in a variety of ways, including:

- To support tasks relating to a wide range of administrative and management activities within schools and colleges
- To create live assessment questions/papers
- To allow automation of marking or to assist human marking
- To improve quality of mark schemes in offering alternative responses
- To allow students to spend more time on higher order skills (interpreting rather than generating)
- To assist teachers in generating class materials and tests
  To provide alternative research avenues for students
- To support students in refining and improving work
- To support teachers with workload, with auto marking, and auto generation of learning material

Question 14
What subjects or areas of education do you believe could benefit most from generative AI tools?

AI can bring benefits to all subjects; it can provide huge numbers of exam questions, especially for STEM subjects, but also has the potential to provide highly reliable marking for discursive subjects. Computer Science can benefit from AI interpreting code and performance subjects could use AI to interpret sound recordings, videos and other media. It can support personalised learning and assessment. Our extensive conversations with teachers indicates that there is little or no correlation between subjects and whether or not teachers feel confident in using AI.

Computer Science
We have seen students use various forms of code generation using tools like ChatGPT in their classroom work already. Of course, students have used code repositories, such as GitHub, in such ways too – albeit engaging in different ways with the tools and resources. This can be for creating whole projects or just for individual parts of projects. When ChatGPT first came out, it would make many mistakes that would need more thorough checks, but it has reached a point where it's very good at creating and analysing reliable code in some situations. In computer science, we are already seeing quite a lot of AI used or wanting to be used within the programming requirements.

STEM subjects
In STEM subjects, generative AI can generate high volumes of multiple-choice questions (MCQs), but it struggles with complexity. Sometimes MCQs require context. If you ask for the MCQ to be made more difficult or challenging, generative AI finds this difficult and it struggles to make more complex or secure examples. So generative AI is useful for creating simple knowledge-recall types of MCQs, but it is less useful (at the moment) for the more complicated questions that students face in our current assessments. As with any other system, generative AI still needs its outputs to be monitored and approved by human operators. We should view AI tools as a potential means to augment what teachers and educational professionals can achieve, but with human expertise at its centre.

English
At an English Stakeholder Forum convened by OCR, participants talked about how they were initially taken by surprise with how AI has been used in more creative subjects, for example, being able to be used for creative writing prompts. Teachers should be able to ask essay questions that are of a good enough quality that they challenge anyone, so that even if they were going to put in a prompt or parameters to AI to help them generate an essay, they would still require skills or knowledge to be able to do decide on those parameters in the first place. Some people are going to be able to use AI tools better than others. Teachers are keen to

understand the appropriate use of AI, particularly for essays or longer responses, and how to cite references that show how AI has been used.

Creative subjects
There are options for the use of AI within performance-based scenarios or creative based subjects as well as STEM. This could include question generation, but also, prompts, data sets for questions, or stimulus materials.

Students do not have access to AI when they are writing their answers to essay questions under controlled exam conditions (at least that is the case currently). But, in assessing student work where AI has been used to generate essays as homework, or coursework, it is important to be able to identify those aspects which show what the student knows and can do and what they're thinking. This requires knowledge of the student and may sometimes require direct discussion with the student.

Question 15
What are your main concerns about using generative AI in educational settings?

a. **Unauthorised disclosures**
   Inputting confidential data into the AI tool could result in the data becoming publicly available to the general public. Added concerns arise in relation to data consisting of personal information, for example, in relation to security and compliance with data protection laws.

b. **Intellectual Property (IP) risks**
   Copyright risks could attach to the AI tool's output. As a result of the way some AI tools are trained, it is possible that responsive outputs could contain elements of copyrighted material. If that material is used in a public facing manner, the user could therefore run the risk of infringing a third party's copyright. A similar risk arises in respect of any third-party data which the user might input into the AI tool for training purposes if the data was licensed for one specific use only or derived from publicly available materials online.

   Whether copyright protection is available for the output itself remains uncertain. If there is no such protection, there may be nothing to stop widespread copying of the outputs of generative AI.

c. **Inaccuracies and bias**
   Gen AI Tools are trained on data from the internet which is not always fair or balanced. As with other AI, generative AI is dependent on the quality of its training data, and therefore susceptible to the introduction of errors and bias through the training and development process. This could create liability issues for users and seem discriminatory to specific groups of individuals. There is also a worrying level of trust from some people in the accuracy of AI, when, in fact, it often creates 'hallucinations' (When Gen AI tools provide incorrect or made-up information). Users need to be aware of the fallibility of AI and moderate its use accordingly.

d. **Security**
   Where the use of generative AI involves sharing data with a publicly available AI system or a private instance on a third-party cloud-based platform, the risk of cyber-attacks is likely to be greater. Where that data consists of personal data it can also lead to data protection violations.

e. **Environmental**
   Foundation models such as large language models have high computational demands which can result in high energy consumption. Businesses should consider the use of energy-efficient hardware and shared (e.g., cloud) infrastructure based on renewable energy in order for their operation to be environmentally sustainable.

f. **Issues of detection**
   A lot of work has been carried out on detecting where AI has been used. The success of such detectors remains mixed and is something of a moving target as AI tools become increasingly sophisticated at an alarming rate.

g. Impact of AI on non-examined assessment (NEA)
NEA includes coursework and a wide range of performance-based activities which include participating in a team sport, dramatic and musical performances, making something, conducting a process or task, interacting with customers, demonstrating oracy skills or conducting independent research and projects.

The impact on coursework and project work has generated concerns that AI can be used to create quality work that is not the authentic work of the student.

The Joint Council for Qualifications (JCQ) is currently developing its position on controlling the risks of malpractice in the use of AI by students within their Non Examined Assessments (coursework). While the number of malpractice cases Awarding Organisations (AOs) have had to manage in relation to the misuse of AI in NEA has been small, there is no guarantee that this will be the case in 2023/24. Chat GPT 3 launched late in the assessment cycle this academic year, the technology continues to improve and the detection software is not wholly reliable. Although currently in draft, the paper begins by stating that there are four broad strategies for managing the risk AI poses to schools and colleges and AOs being able to detect NEA work that is not a student's own work:

- Improving the JCQ guidance and AO support for centres in detecting the misuse of AI
- Tightening task-setting controls
- Tightening task-taking controls
- Changing the assessment

Question 16
If at all, have these concerns impacted your use of generative AI? Please explain how.

These concerns have inevitably impacted on our use of generative AI, particularly for operationally critical, regulated or user-facing processes - the risks are more acute or less so depending on the context. The use of AI to support teaching and learning, or for formative assessment purposes carries less risk than if it were to be used to generate or mark an assessment that contributes to the awarding of a high stakes qualification.

Question 17
Are there specific subjects or areas of education where you believe generative AI should not be used? Why?

Yes - Where it replaces the fundamental constructs of the qualification (e.g., doing calculations, writing essays etc.). There are arguments for reviewing if these constructs are still relevant going forward, but in some cases such as civil engineering, you would always expect an expert to be able to do manual calculations to understand and second check those generated.

In society we will continue to need people who can synthesize a range of complex data in order to make recommendations for action, evaluating risks/costs/benefits – it seems unlikely this could be made obsolete as a skill. We do need to reflect on how these skills are effectively developed in young people, and whether the current curriculum and assessment models continue to be appropriate.

Question 18
If any, what are your views regarding ethics, data privacy and security when using generative AI in education?

As an awarding body operating in a highly regulated market such as that for Ofqual- recognised qualifications, we have to exercise great caution when considering the use of generative AI, particularly in areas involving high impact decisions for learners. Examples of such areas include the development of test papers for 'high stakes' qualifications and marking.

Errors in connection with such decision-making can have adverse political, legal, regulatory and/or reputational consequences for the industry.

Generative AI presents major areas of ethical concerns for the education sector, including data privacy and non-discrimination. The risks in these areas include the following (amongst others):

a. The risk of data breaches. The security risks are arguably more pronounced when using generative AI. This is because of the vast amount of data being processed by such systems.

b. The lack of consent and transparency. Data subjects have a legal right not be subject to decisions based solely on automated processing, which go on to affect them significantly. In the context of education, these decisions might relate to such things as grading test papers or identifying cheating. One way organisations can overcome this hurdle is by obtaining data subjects' explicit consent. However, this will often prove challenging in relation to such decision-making. Even informing data subjects about how AI systems work will be difficult given their complexity.

c. The potential for bias and discrimination. The use of generative AI in education might be viewed as a means of overcoming subjectivity or bias. However, this view ignores the fact that generative AI might be trained on data which is itself riddled with biases.

There are wider uncertainties around the application of data protection law to generative AI, for example, in relation to the application of the fairness principle (which is touched on above), the purpose limitation principle and the data minimisation principle. All of these heighten the challenge for businesses operating in this sector.

Data policies from AI companies should be clear and be in the best interest of candidates. For instance, will what the user is typing be used to support the AI development?

Future predictions and enabling use:

Question 19
How do you see the role of generative AI in education evolving in the future?

Generative AI is likely to have a dramatic impact on automating non-teaching tasks, for example, learning resource creation, timetabling, lesson planning, and report writing. This has the potential to allow teachers more time to concentrate directly on teaching.

For higher impact activities, such as the development of test papers for 'high stakes' qualifications and marking, we have to proceed with greater caution, bearing in mind the above uncertainties and concerns. Of added relevance is the fact that we operate in a highly regulated industry in which maintaining public confidence in the products we offer is a legal necessity.

Central to many public concerns regarding generative AI is the deeper question concerning the role of human judgment in connection with decisions made using such systems. It is arguably the role of law makers to proactively address this question for businesses generally. AI is both too promising and too risky for governments to take a hands-off approach.

AI is already having a massive impact on education, work and wider society. It is incumbent on the education system to provide young people with the skills they will need to use AI in their future lives. We need to consider revision of the current curriculum, especially at key stage 4, to make it better reflect and prepare people for the AI revolution. This process should recognise that Maths and English may become more important than ever, even as subjects like computing evolve.

Generative AI increases potential to personalise learning to individual student needs and provide one-to-one support. It also seems likely that Non-Exam Assessment (NEA) will have to become controlled in some way or the constructs changed to encourage and utilise the use of AI in NEA whilst still providing distinct evidence of the individual skills and knowledge demonstrated by the student.

Question 20
What support do education staff, pupils, parents or other stakeholders need to be able to benefit from this technology?

OCR conducted a series of interviews with schools and colleges to gather insight on how they are responding to AI, the issues they face, and what further support they might want from exam boards. Some stated they feel very confident using AI, having already worked with staff to establish best practice. Others acknowledged its forthcoming importance and agreed on the need to upskill.

There was consensus on the need for the provision of a reliable AI detection tool. Schools who responded did not feel confident in reliably identifying AI usage and how to differentiate this from the students own, authentic performance.

Question 21

What activities would you like to see the Department for Education undertaking to support generative AI tools being used safely and effectively in education?

Schools need quality digital infrastructure: high-speed internet access with enough laptops and tablets to teach and test well. This is not just for generative AI, but to take advantage of other technological opportunities, including digital assessment. There is a role for the DfE in ensuring adequate and equal access to such technology.

The DfE should set out plans for a review of current subject content and assessment constructs ahead of the next round of qualification and curriculum reform. The DfE needs to ensure any subject content criteria for qualifications reflect the impact and broadening use of AI in education in terms of what knowledge and skills students are required to have for each qualification. There should be a review of current criteria to ensure they are still fit for purpose, and consideration given for when future criteria are developed.

The DfE could also play a role in encouraging the capture and dissemination of best practice in the use of generative AI.

Question 22

Is there anything else you would like to add on the topic of generative AI in education?

The government needs to look carefully at which technologies work today – and are working today in classrooms – and decide which of those we need more of. It must be recognised that there are opportunities to enhance learning and assessment with tried and tested means that already exist. At the same time, we cannot put generative AI back in its box. We have to embrace the opportunities for education while being clear-sighted about the limits.

OCR
Oxford Cambridge and RSA

Find out more
www.ocr.org.uk

or call our Customer Support Centre on **01223 553998**
Alternatively, you can email us on **support@ocr.org.uk**

CAMBRIDGE
UNIVERSITY PRESS & ASSESSMENT