

Some thoughts on AI

Megan Talbot (met32@aber.ac.uk)

Two things to talk about

- An assessment I designed and how it went
- My observations from running AI authenticity panels

Our AI enabled assessment

- “To what extent should British law recognize prenuptial agreements?”
 - They get the ChatGPT answer.
 - Can you beat the AI?
-
- Answer is judged as normal, not on the extent to which AI was or was not used.

Why?

- Places AI use in a context more akin to how it may be used professionally (as a tool to be supervised + exercising discretion)
- Encourages evaluation of strengths and weaknesses of the text, useful both within and beyond the context of AI.
- Hard to cheat with using AI, it will have a hard time improving on itself.

How was it handled

- “do I have to use the AI output if I don’t want to” –this mostly came from students who were doing well in other modules and ended up doing well in this one.
 - No, its yours to use as much or as little as you want
- We provided guidance on how to improve from the AI output
- “You can submit the unchanged ai answer if you want, we won’t tell you what mark it would get, but you can if you honestly think it represents your best work.”

What changed?

- Average marks for lvl 2 went up, average marks for lvl 3 went down, overall there was no significant average mark change
- The overall percentage who failed the module went from 12.5 to 7
- The percentage of 1sts seems to have gone up from 14 to 19.2
- Please forgive me, the data I had wasn't great, and I'm not used to doing stats, have mercy on a poor humanities lecturer :D

The data

Module code	Module Full Name	No Studs	Marks 1 - 19	20 - 29	30 - 39	40 - 49	50 - 59	60 - 69	70 - 79	80+	Average	Stddev	% Failed
LC26420	Family and Child Law	49	0	2	5	7	23	9	1	0	52.00	10.71	16.33
LC36420	Family and Child Law	15	0	1	0	1	2	3	6	2	64.80	16.66	6.67
2024													
LC26420	Family and Child Law	35	0	1	1	6	9	12	5	0	56.00	11.86	8.57
LC36420	Family and Child Law	22	1	1	0	2	4	7	5	1	59.62	15.84	13.64

Pros / cons?

- It seems to “cut the bottom off”
- It encourages explicit use of AI literacy skills, “is it worth using this”
- Uses “normal essay format”, avoids perturbing the students too much with a new assessment style.
- Despite being warned some students get caught out by hallucinations, so skills may not be being exercised as I would want.

Would I do it again?

- I legit don't know (I welcome your input)
- Evidence is suggesting that using AI in certain ways is not great for learning
 - (if you are interested)
 - Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X. H., Beresnitzky, A. V., ... & Maes, P. (2025). Your brain on chatgpt: Accumulation of cognitive debt when using an ai assistant for essay writing task. *arXiv preprint arXiv:2506.08872*.
- Exercises with these skills are good, but I am not sure if a summative assessment is the place for it?

Authenticity

- All of this is anecdotal, I hope the uni collects the data though
- The number of students using AI improperly doesn't seem to be going down
- Detection varies DRAMATICALLY between staff, possible injustice?
- “marking down” to avoid paperwork / process very common
- Most confess at the first opportunity.

My approach

- “how did you prepare this work”
- Subject expert usually asks for explanation of any phantom references or nonsensical claims.
- Meetings tend to be longer and more educative
- Often connect student to resources / personal tutors

Some trends I have noticed

- The majority have a job + full time course
- Students with low confidence
- special circs often occurring at the same time
- “I knew I needed to improve so I tried something different” – paying attention to marks, but not feedback
- Low use of resources / assistance

Explaining why

- Laziness vs desperation?
- Some students are far too reluctant to talk to us for help
- Some time crunches are real, some are not
- Very poor general skills?

A notion

- The acceptable number of UAP's is 0
- What if we treated every UAP like it is a critical event rather than an inevitability?
 - Not necessarily increasing punishments, research suggests that doesn't work.
 - If preventative measures are introduced its often ad hock by departments and inconsistent with no monitoring to see if they are successful.
 - We don't have a UAP prevention strategy or any organised approach to reviewing cases and implementing preventative measures.
 - Accepting some UAP's as inevitable lets our students down (or rather makes it easier for them to let themselves down) and is more resource intensive than prevention

How can we help maybe?

- A lot of it seems to be caused by general problems
- To help with skills we already ask students to meet their PT regularly, but many don't, and we don't have the resources to see them all if they did.
- Some sort of “skills traffic light” system?
- How can we get students to read / process feedback (we offer meetings to help, but many don't come)
- Why are some students picking full time courses when part time might suit them better?

I would love to hear your thoughts

- I'm not sure if I want to run the assessment again or try something else.
- Do you have any experiences of AI UAP's you want to share?
- I would love to exchange ideas about how we can tackle the causes of AI misuse.