Systemic Approaches to GenAI in Mathematical Sciences Education

Stuart King s.king@ed.ac.uk

THE UNIVERSITY *of* EDINBURGH

Outline

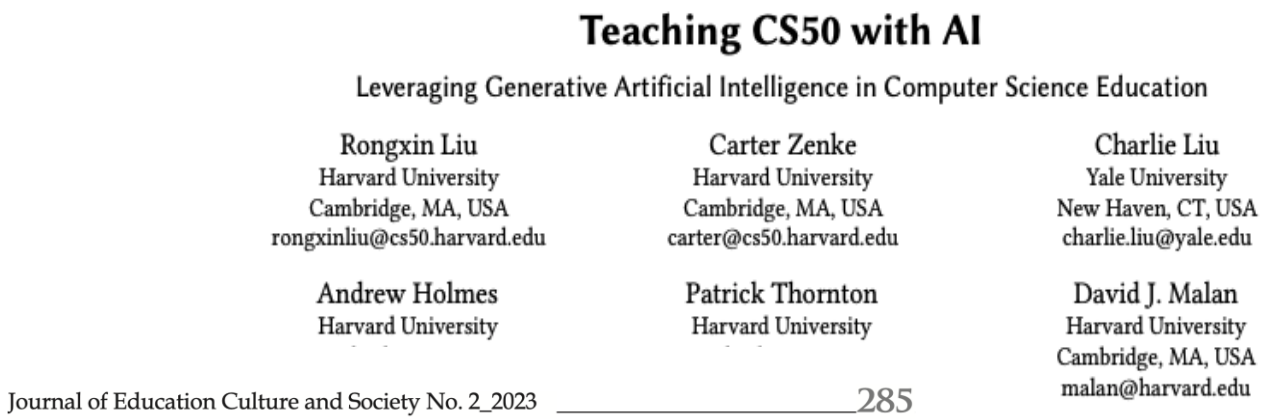
- From particular to general – individual use to collective response
- Student voice, (Staff voice, press?)
- Disciplinary norms – why we need expectations around GenAI use
- How should academic departments or universities be supporting/scaffolding changes?
- What next for us as we lecture our courses?

Particular – how is GenAI being used?

- A lot of the initial response to GenAI (read LLMs throughout) came/comes from those keen on the tech
- I was one of those teaching on a course in introductory programming to arts and humanities MSc students – using creative writing as a way into that
- LLMs were unavoidable in that context which pulled me into working with this since GPT2 era
- This context useful for thinking about the AI in Education (AIED) literature as it stands

Particular – how is GenAI being used?

- Literature has quite a lot of ‘I tried this’ and it worked/didn’t/caveats (Good!)
- But not a lot of analysis and guidance of how to move beyond a single course or intervention



Teaching CS50 with AI

Leveraging Generative Artificial Intelligence in Computer Science Education

Rongxin Liu
Harvard University
Cambridge, MA, USA
rongxinliu@cs50.harvard.edu

Carter Zenke
Harvard University
Cambridge, MA, USA
carter@cs50.harvard.edu

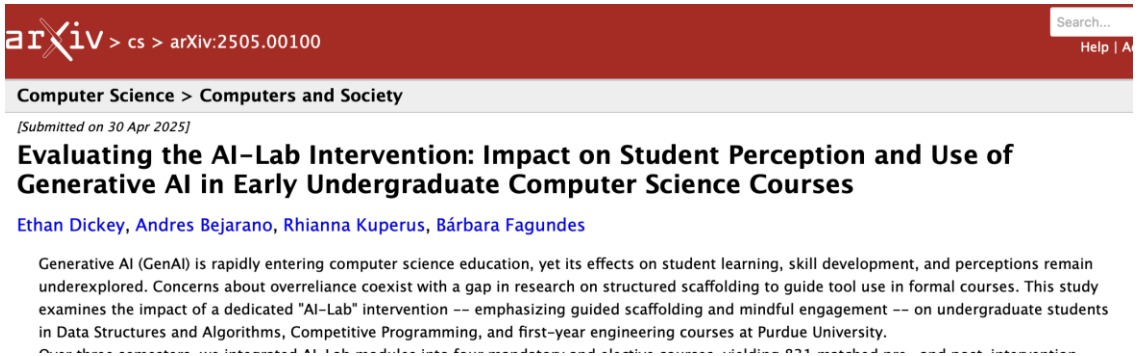
Charlie Liu
Yale University
New Haven, CT, USA
charlie.liu@yale.edu

Andrew Holmes
Harvard University

Patrick Thornton
Harvard University

David J. Malan
Harvard University
Cambridge, MA, USA
malan@harvard.edu

USING AI CHATBOT FOR MATH TUTORING



arXiv > cs > arXiv:2505.00100

Search... Help | A

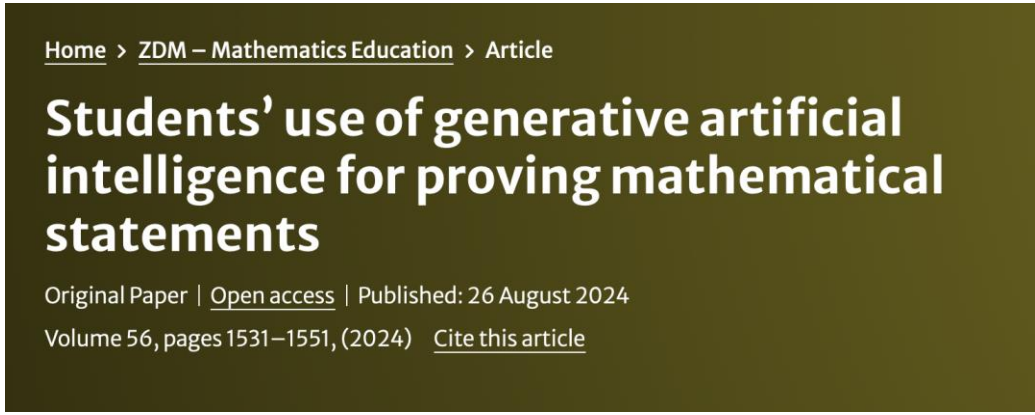
Computer Science > Computers and Society

[Submitted on 30 Apr 2025]

Evaluating the AI-Lab Intervention: Impact on Student Perception and Use of Generative AI in Early Undergraduate Computer Science Courses

Ethan Dickey, Andres Bejarano, Rhianna Kuperus, Bárbara Fagundes

Generative AI (GenAI) is rapidly entering computer science education, yet its effects on student learning, skill development, and perceptions remain underexplored. Concerns about overreliance coexist with a gap in research on structured scaffolding to guide tool use in formal courses. This study examines the impact of a dedicated "AI-Lab" intervention -- emphasizing guided scaffolding and mindful engagement -- on undergraduate students in Data Structures and Algorithms, Competitive Programming, and first-year engineering courses at Purdue University.



Home > ZDM – Mathematics Education > Article

Students’ use of generative artificial intelligence for proving mathematical statements

Original Paper | Open access | Published: 26 August 2024

Volume 56, pages 1531–1551, (2024) Cite this article

Particular – how is GenAI being used?

- Starting to see meta-analyses of these cropping up
- But starting with this very positive skewed literature suggests unwarranted positive approach!

Expert Systems With Applications 252 (2024) 124167

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Artificial intelligence in education: A systematic literature review

Shan Wang [a,1], Fang Wang [b,2,*], Zhen Zhu [c,3], Jingxuan Wang [c], Tam Tran [a], Zhao Du [d]

[a] Edwards School of Business, University of Saskatchewan, 25 Campus Drive, Saskatoon, SK, S7N 5A7, Canada
[b] Lazaridis School of Business & Economics, Wilfrid Laurier University, Ontario, N2L 3C5, Canada
[c] School of Economics and Management, China University of Geosciences, Wuhan, Hubei, 430078, China
[d] Sport Business School, Beijing Sport University, Beijing, 100084, China

Particular – what are some good uses?

- Ongoing teaching innovation project in UoE exploring this from teacher-led perspective
- Some relevant to maths:
  - Summarising content
  - Making quizzes
  - Organising time, revision schedule
  - Coding assistant
  - Sounding board and ideation
  - Tutor – are we there yet though?

AI FOR TEACHING INNOVATION

How might generative AI help us develop new and interesting ways to teach? What potential AI applications might academics want to use, but don't have the skills or time to build? How might generative AI be used to develop creative and practical ways of doing teaching differently?

This project is about supporting and enabling creative teaching innovation through generative AI, by providing course teams with learning design and app development support to build prototypes. It aims to develop our capacity as an institution to think widely and creatively about human/machine partnerships in education innovation, while supporting staff skills development and capacity for working with AI in their teaching, assessment and feedback.

It's part of the University's AI adoption programme of work, and a partnership between the Edinburgh

RESEARCH AREAS

Cultures and Futures

PROJECT TEAM

Professor Siân Bayne; Javier Tejera

KEY CONTACT

Professor Siân Bayne

https://edin.ac/3XyiDyy

Lots of focus on assessment

- A lot of the general discussion in HE is around assessment
- Anxiety driven by the recent HEPI report https://edin.ac/4lv0prJ and media coverage (92% of students use it, 18% in submitted work)
- General feeling from colleagues that usage isn't penalised enough, often doesn't feel it is 'provable' or worth it
- Universities report low levels of academic misconduct cases for the most part – but probably under-reported
- Certain types of assessment feel more 'at risk'
- Disconnect from previous 'keen' usage – feels like an emerging gap engagers/avoiders

What are the bounds on one course?

- Feels somewhat like a perfect storm for academic course leaders to cope with
  - Berated if we don't innovate and change
  - Pressure externally and from students around AI
  - Limited time and workload to adapt
  - Unclear or limited guidance at institutional level
- Lurking danger that all falls on course leads to do
- Somewhere department/institutional thinking has to start in as well
  - Where are GenAI skills taught – wasteful to all do it
  - Balance of assessments – how many 'safe' vs 'at risk' assessments do students do
  - What gets assessed by exam say?
  - Ethos of one course affects expectations for others

Moving towards collective action

- Any directives at course level about assessment security set a path towards too many exams – the appropriate unit feels bigger than that
- Can answer many short direct questions or arguments
- Deprecates using very transactional forms of assessment, overloading
- Poorer at citations – not a knowledge engine (probably won't be an issue for much longer)
- Hasn't changed the type of threat to assessment (could have asked someone else to write it), but changes ease of access, and may feel different morally to students
- Students are already using it extensively

Moving towards collective action

- Requires more thinking across years/programmes/dept
- Thinking about assessment
  - Need to test assessments in GenAI
  - Strategies to avoid GenAI doing well are diminishing returns
- Balance of assessment security across portfolio of assessment
  - Use of exams or Viva
  - Knowing student's work and standard across a course
  - Assess process rather than output – how they would solve a problem rather than just the solution
- Less assessment (or at least less summative)
- Develop more ways to incorporate using GenAI in what we do
- Keep assessing student/staff views – begin to get a handle on expectations for use of GenAI in each discipline – what's ok and what isn't

Different voices – public, students, staff

- I’ve been involved with a a number of surveys, focus groups of students and staff, and one student’s assembly
- Those are across different student and staff groups (not all maths!)
- Rather than give results of one of those I’ll attempt to synthesise from across them and draw out some points - with a focus on the main focus group session
- A lot of the feedback probably isn’t very surprising – but maybe useful to know that intuition is probably correct
- A lot of shared anxieties and views and some differences – generally students are maybe more trusting and positive about the tech than staff

AI & HE press

- Majority focussed on cheating
- Perception that Universities caught napping and don't have a plan
- Continual questioning of integrity of our students

Revealed: Thousands of UK university students caught cheating using AI

Guardian investigation finds almost 7,000 proven cases of cheating – and experts says these are tip of the iceberg

'I received a first but it felt tainted and undeserved': inside the university AI cheating crisis

Student AI cheating cases soar at UK universities

Figures reveal dramatic rise in AI-related misconduct at Russell Group

sed by sector's 'patchy record-keeping'

BBC Sign in Home News Sport Weather iPlayer

NEWS

Home | InDepth | Israel-Gaza war | War in Ukraine | Climate | UK | World | Business | Politics | Cultur

Scotland | Scotland Politics | Scotland Business | Edinburgh, Fife & East | Glasgow & West | Highland:

Alba

Are Scottish students using AI to cheat their way to a degree?

Latest Newsletters

The Atlantic

TECHNOLOGY

AI Cheating Is Getting Worse

Colleges still don't have a plan.

By Ian Bogost

Student workshop views

- Students use GenAI for generating explanations, creating outlines, and developing questions.
- The vast majority of students report having used GenAI in their studies. Very few students have not used GenAI.
- Acts as a grammar and tone checker, helpful for non-native English speakers; saves time in organizing thoughts.
- More used when knowledge gaps exist; accuracy issues noted, especially in source-finding.
- Assists in rephrasing, structuring, topic exploration, math tutoring, and simplifying jargon.
- Concerns about reliance on GenAI for assignments and academic integrity; peer pressure noted. (Maybe a lot is around perception of use)
- Usage often naïve; students struggle with effective prompting.
- Preference not given to ELM (UoE tool); students use various AI tools and are unconcerned about data privacy.

Student workshop views

- Lack of awareness and unclear GenAI guidelines for students (and staff?); many unaware of assignment specifics.
- Need for explicit AI use instructions and citation requirements in course materials.
- Banning AI unrealistic; call for workshops and course-specific guidance. Concerns about transparency in marking.
- Group work discrepancies in AI use affect grades.
- Concerns about staff transparency using GenAI for marking.
- AI hallucinations and biases impact critical thinking and outputs.

Other interesting student points

- Majority of students report at least some level of anxiety about AI (reported high levels of anxiety similar to staff 35% students - 47% staff)
- Github copilot and ChatGPT are the well-liked tools
- Majority of students ‘agree that AI enhances learning efficiency’
- Students have a higher adoption rate of the tech compared to staff
- There is some concern among students about the environmental impact of GenAI

Staff views

- More skepticism from staff about the effect of AI on learning (indeed reflecting mounting evidence to that effect)
- Staff more split on sentiment – a significant number less positive
- What about online students where exams aren’t as common
- Concern about university provision of invigilated assessment
- How do we address the range of staff skills in GenAI – where does the time come from to upskill
- Ultimately does this endanger our jobs – existential concerns - what does it do to the discipline?

Disciplinary norms

- One helpful thing would be a clearer idea of emerging norms in each discipline – what do we care students use GenAI for
- How about what we are allowed to do?
  - Can I write my grant using GenAI? – how about using it to sketch out a plan? – how about fill in the boilerplate stuff?
  - Can I write that internal documentation using GenAI?
  - How about marking? Could I use it to speed up? What about using it for formative feedback?
- Varies across discipline (art vs maths care about image generation differently) – but students hate this variation in regulation
- Might help to have some scenarios for us to discuss and respond to

Disciplinary norms - principles

- Assessments should be authentic, so we should integrate use of AI into assessment
- Consistency but not uniformity (across Dept/Depts/University)
- We enable and trust students to behave with academic integrity
- Policy should be co-created with students
- Even if AI can do it we might want students to do it (and be clear with students what we want and why)
- We should assess/certify students' understanding (it's about what students can do not AI) – 'I don't want to mark an AI output with no student effort'

How should academic departments or universities be supporting/scaffolding changes?

- There are some structural things that sit at Uni level
  - Ensuring equity of access to LLM tool
  - Setting up sensible regulatory frameworks – dealing with everything as misconduct is probably a bad idea
  - Provision of invigilated exams
  - Workload allocation for reworking assessment
  - General training students/staff
  - Collating examples of GenAI in use in courses - ‘cataloguing the possible’
- Clearly there is also a role for prompting rethinking of assessment at department level – both across programme, and in terms of amount and ‘security’

How should academic departments or universities be supporting/scaffolding changes?

- The big tech companies are pushing their platforms – we should be wary of the tech rather than teacher led development
- The predictions for this tech sector is still further fragmentation – we are likely to see LLMs more tuned to academia, maybe more directly aimed at cheating assessments
- Chatbots are being built in all over – and can be useful for helplines etc – seems a matter of time before they start to pop up in VLEs etc – we maybe need to get more used to this idea
- Important that we document innovation, but in ways that recognise what works beyond an individual invested course lead

Practical steps – all hope lost?

- Test questions; maybe adapt and change form where you can
  - Changing form for some things
  - More emphasis on reflection and writing about process
  - Process over outcome (to take a leaf from our creative arts colleagues)
- Communicate what you want to happen to students, supply the ‘why’ part
  - Why is this assessment useful to their learning?
  - Why they shouldn’t use GenAI for a task if you stop them?
- Model for them what is and isn’t ok with GenAI
- Discuss it, demonstrate the good and bad of its usage, encourage students to engage **critically** with it

What are our options when GenAI challenges our assessments?

- Need time to think and discuss in discipline and to build new norms
- Important to keep talking therefore
- Think again about our assessments, our Learning outcomes/objectives and the balance of what we do
- Balance of assessment ‘security’ across programmes (course is probably too small a unit)
- We could begin now:
  - Audit assessments across programmes and review those most threatened or susceptible to undermining by GenAI tools
  - Discuss in the discipline what we use it for, begin to understand normative use

What are our options when GenAI challenges our assessments?

- Keep assessment ‘as is’ – learning objectives (learning outcomes etc) unchanged
  - Ban GenAI (hard)
  - Change mode to avoid GenAI use (many exams?)
  - Don’t care (or use and cite)
- Modify assessment – learning objectives unchanged
  - Ban GenAI, modify to make it impossible to use (figleaf solution)
  - Make use of GenAI (significant change?)
- Modify assessment – change learning objectives
  - Change substance - Ask something different incorporating GenAI, add to LOs
  - Change form – Ask something different and change style of assessment (incorporate or prevent)

Case study – Research skills

- Paper contains a review of a method and a range of applications so can't summarise everything.
- Student answers were not always strong and the most basic LLM answer is a pass (not a strong one).
- LOs aimed at preparing students for research projects

Task

Write a 2 page report (minimum font size of 10) of the above research article. The report must offer a clear summary of the key ideas, concepts, and methods in the paper. In writing your report you may wish to consider for example:

- What are the key concepts, ideas, and methods from the article? You need to do this *in your own words*; however, *do* follow the same notation from the paper for any mathematical expressions you may use.
- What alternative approaches are there to that described?
- What are the strengths and limitations of the proposed methods?
- Can you think of any other novel applications and/or extensions to this research?
- How does the proposed approach perform on any (simple) examples or compare to other techniques?

Some applications of nonlinear and non-Gaussian state–space modelling by means of hidden Markov models

Roland Langrock*

Institute for Statistics and Econometrics, University of Goettingen, Platz der Goettinger Sieben 5, 37073 Goettingen, Germany

(*Received 07 December 2010; final version received 16 March 2011*)

Nonlinear and non-Gaussian state–space models (SSMs) are fitted to different types of time series. The applications include homogeneous and seasonal time series, in particular earthquake counts, polio counts, rainfall occurrence data, glacial varve data and daily returns on a share. The considered SSMs comprise

Case study – Research skills

Some options:

- Just don’t care about LLMs – reassess ‘baseline’ performance, re-emphasise parts LLM does poorly here (unique voice, thinking about new ideas and applications, focusing summary, unsaid information).
- Could give the LLM response and ask for a critique and get students to reflect on what is left out, or what they would change (but LLM could help even with this?).
- Incorporate a second paper and focus the assessment more on how the paper could be applied to a different topic area.
- Ask for more reflection – how would it change their approach to an example problem. Harder to fake this reflective mode.

What are our options when ChatGPT challenges our assessments?

- Keep assessment ‘as is’ – learning objectives unchanged
  - Ban GenAI (hard)

    Look for ways to incorporate non-use of AI – viva, exam etc

    Bans in coursework aren’t enforceable (and the situation is not getting better).
  - Change mode to avoid use (many exams?)

    This will work for some things, but the danger is a return to all exam, which is something most of us don’t want.
  - Don’t care!

    Easy to justify in some cases – maybe accompanies a move to less overall assessment, and more formative work where this suits well.

What are our options when ChatGPT challenges our assessments?

- Modify assessment – learning objectives unchanged
  - Ban GenAI, modify to make it impossible to use (figleaf solution)

  Hard to do well, the GenAI tools are becoming more effective, so pits us against them.
  - Make use of GenAI (significant change?)

  Again, some examples out there but this is the best option for many things – the drawback is it needs time and space to explore and those are in short supply!
- Modify assessment – change learning objectives
  - Change substance - Ask something different incorporating GenAI add to LOs

  Again, another preferrable solution perhaps, but same comment about needing time and space – maybe even more so here as more thought needed to adjust the LOs.
  - Change form – Ask something different and change style of assessment (incorporate or prevent)

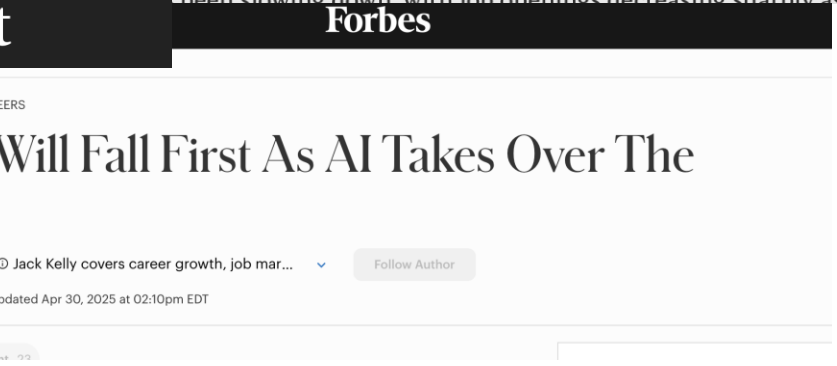
  A lower effort version of the above maybe

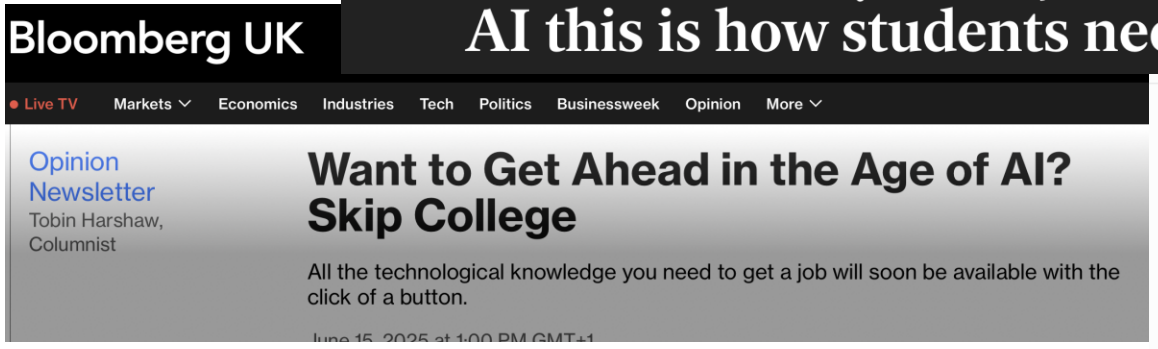
What next?

- It feels like there is a lot to do – and at a time of increasing pressure within the HE sector
- The next set of entrants already have experience in using and relying on GenAI – so we have more work to do in affecting that
- There is something in the ‘universities not coping with this narrative’
- From quick survey across RG universities
  - Not all had external pages explaining their approach (not clear if there were password protected pages)
  - A divide between traffic light style, and just ‘cite if used’ – the latter feels like it cedes quite a lot of ground
  - Nearly all have given up on the idea of detecting unauthorised use
  - Continued questions over provision of invigilated assessment (recent joint IMA/RSS/LMS statement)
  - A lot (most) still using MS Copilot as main system

AI and recruitment?

- Lots of noise about reduction in entry level jobs ‘because of AI’ – changes perceptions at least
- Knock on to recruitment? – adding to the storm facing the HE sector





Final thoughts

- Test your questions on LLM of choice etc – continue thinking about assessment changes
- Supply the ‘why are we doing this’ part of the assessment more.
- Provide more guidance at each assessment
- Discuss it, demonstrate the good and bad of its usage, it isn’t perfect!
- Encourage students to engage **critically** with it.
- Students and staff need to be equipped in using it and how it works
- Dept level - Discuss how GenAI sits in our discipline; Audit assessments across programmes

Thanks for your attention

Always interested in hearing more from people on this topic!

Stuart King s.king@ed.ac.uk