

Teachers' Interview Findings and Results

Jesper Bruun, Aswin Rangkuti
Department of Science Education
AIMSC Radar Meeting

UNIVERSITY OF COPENHAGEN



Background and Research Questions

Research on GAI in physics education ask what the technology can do.

Physical Review Physics Education Research

Highlights Recent Accepted Collections Authors Referees Editorial Policies Press About Editorial Team

OPEN ACCESS

Performance of ChatGPT in solving wave problems

[Muhammad Aswin Rangkuti](#) ^{1,2,*}, [Wawan Bunawan](#) ^{1,2}, [Hendra Y. Agustian](#) ³, [Tuti Hardianti](#) ^{1,2}, [Dewi Wulandari](#)¹, and [Nazwa Hudaiby](#) ¹

Show more

Phys. Rev. Phys. Educ. Res. **22**, 020113 – Published 6 August, 2026

DOI: <https://doi.org/10.1103/4dgy-1dg1>

 2

Show metrics

Abstract

The ability of ChatGPT, especially its latest versions, attracts significant attention from educators and researchers in fields such as physics. Several studies on this topic explore the performance of ChatGPT in answering physics questions across different versions of ChatGPT. In this study, we evaluate the development of ChatGPT in answering physics questions using the Mechanical Waves Conceptual Survey (MWCS) to examine whether its capabilities improve across different versions. A total of 75 responses per item are generated using ChatGPT versions 4, 4.o, and 5.2, resulting in 4950 responses for comprehensive analysis. Based on the comparison between the three versions of ChatGPT, we find that some difficult types of questions that are answered incorrectly by previous versions, as reported in earlier studies, can now be answered correctly, although there are still minor mistakes, especially when examining the reasoning. A generalized linear mixed

Physical Review Physics Education Research

Highlights Recent Accepted Collections Authors Referees Editorial Policies Press About Editorial Team

OPEN ACCESS

Exploring generative AI assisted feedback writing for students' written responses to a physics conceptual question with prompt engineering and few-shot learning

[Tong Wan](#) ^{*} and [Zhongzhou Chen](#) [†]

Show more

Phys. Rev. Phys. Educ. Res. **20**, 010152 – Published 13 June, 2024

DOI: <https://doi.org/10.1103/PhysRevPhysEducRes.20.010152>

 0

Citations 117

Export Citation

Show metrics

Abstract

Instructor's feedback plays a critical role in students' development of conceptual understanding and reasoning skills. However, grading student written responses and providing personalized feedback can take a substantial amount of time, especially in large enrollment courses. In this study, we explore using GPT-3.5 to write feedback on students' written responses to conceptual questions with prompt engineering and few-shot learning

Physical Review Physics Education Research

Highlights Recent Accepted Collections Authors Referees Editorial Policies Press About Editorial Team

OPEN ACCESS

Can ChatGPT support prospective teachers in physics task development?

[Stefan Küchemann](#) ^{1,*}, [Steffen Steinert](#) ¹, [Natalia Revenga](#)¹, [Matthias Schweinberger](#) ¹, [Yavuz Dinc](#)¹, [Karina E. Avila](#) ², and [Jochen Kuhn](#) ¹

Show more

Phys. Rev. Phys. Educ. Res. **19**, 020128 – Published 11 September, 2023

DOI: <https://doi.org/10.1103/PhysRevPhysEducRes.19.020128>

 2

Citations 114

Export Citation

Show metrics

Abstract

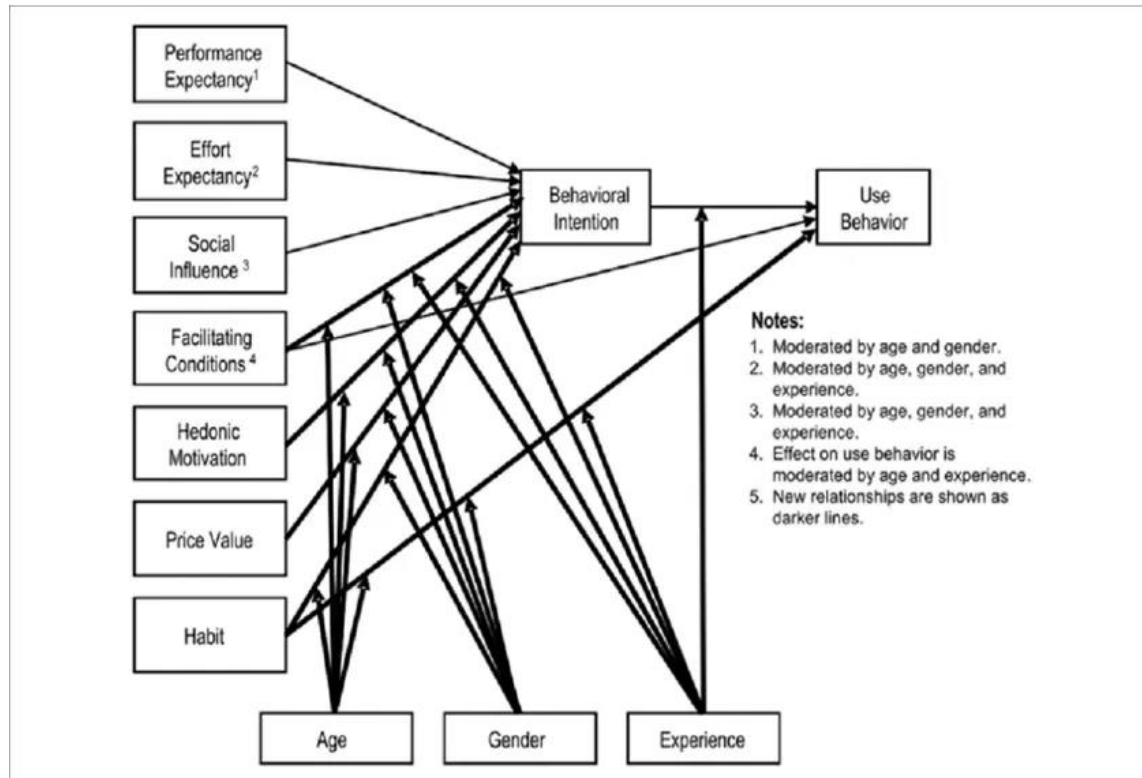
The recent advancement of large language models presents numerous opportunities for teaching and learning. Despite widespread public debate regarding the use of large language models, empirical research on their opportunities and risks in education remains limited. In this work, we demonstrate the qualities and shortcomings of using ChatGPT 3.5 for physics task development by prospective teachers. In a randomized controlled trial, 26 prospective physics teacher students were divided into two groups: the first group used ChatGPT 3.5 to develop text-based physics tasks for four different concepts in the field of kinematics for 10th-grade high school students, while the second group used a classical textbook to create tasks for the same

Far less is known about how teachers actually use and experience it

What we wanted to find out

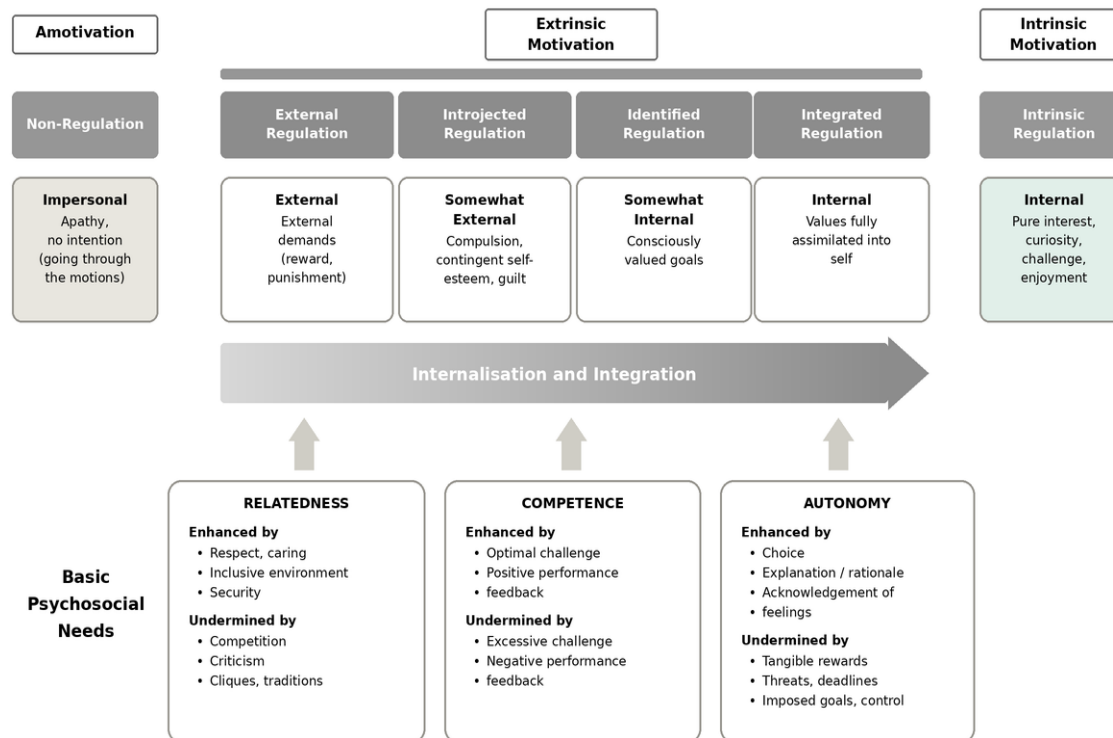
- How do teachers adopt and use GAI in their teaching?
- What are their concerns?
- How does it affect them as teachers?

Methods (Framework to develop the questions)



- Performance expectancy: Can GAI improve teaching?
- Effort expectancy: How much effort required to learn and operate GAI?
- Social influence: For example, pressure to GAI.
- Facilitating conditions: Do teachers get help in using GAI?
- Hedonic motivation: Do teachers enjoy using GAI?
- Price value: Trade-off between benefits and financial costs.
- Habit: Do teachers use GAI routinely?

The Unified Theory of Acceptance and Use of Technology (UTAUT2) (Vakantesh, et. Al, 2012)



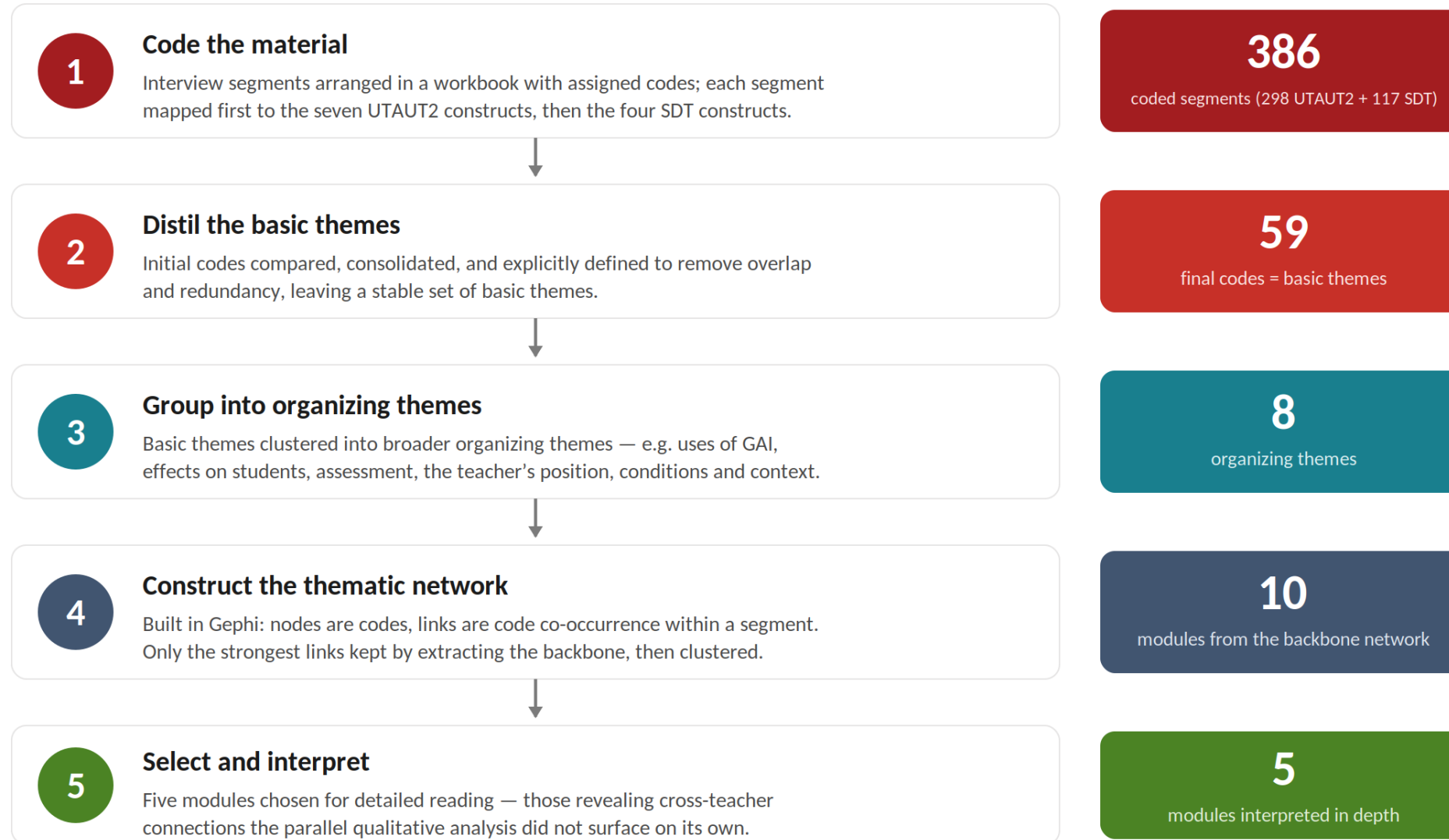
- **Autonomy:** Is use of GAI is a self-endorsed choice that fits their own values and judgement or is it more imposed on them.
- **Competence:** Do teachers feel more or less capable in their work as physics teachers with GAI?
- **Relatedness:** GAI's influence on teachers' relationships with their students and colleagues.

Self Determination Theory (SDT) (cook & Artiono Jr, 2016)

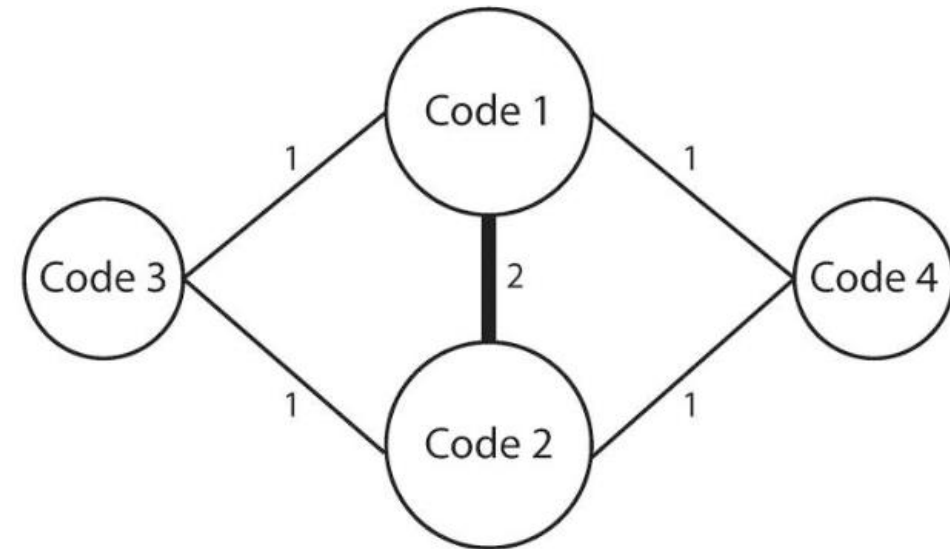
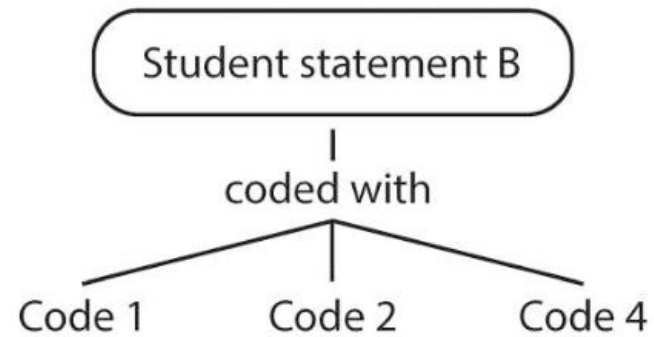
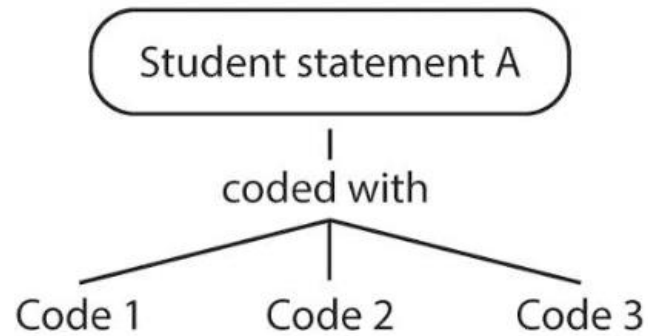
Participants

Teacher	School	Gender	Experience Using GAI	Type of GAI
Teacher 1	School 1	Male	Frequent	ChatGPT, Claude
Teacher 2	School 2	Male	Moderate	ChatGPT
Teacher 3	School 2	Male	Light	Copilot
Teacher 4	School 3	Female	Moderate	Copilot
Teacher 5	School 3	Male	Light	ChatGPT, Copilot
Teacher 6	School 4	Female	Light	Copilot
Teacher 7	School 4	Male	Moderate	Copilot
Teacher 8	School 5	Female	Light	Copilot, ChatGPT
Teacher 9	School 6	Female	Light	Gemini
Teacher 10	School 7	Female	Light	Copilot
Teacher 11	School 7	Male	Moderate	Copilot

Stages of thematic network analysis



How to read





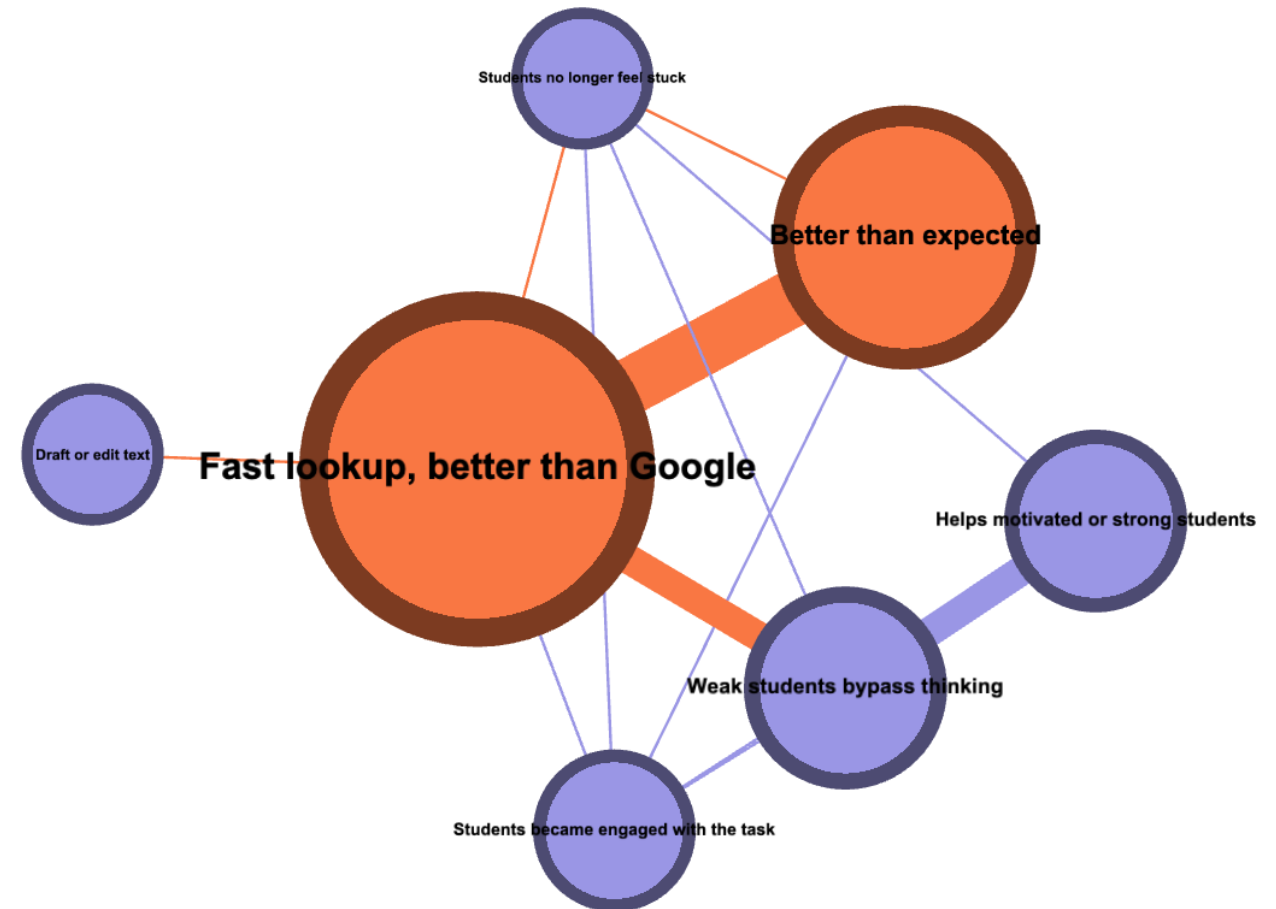
Results

The strong-weak divide

- Motivated/strong students use GAI to go further, while weaker students use it to avoid the thinking.

“

Academically strong students most often come from academically strong homes — a mum or dad who's an academic, help over dinner. The students who don't have that opportunity are left missing it. That's what I'd realised when these AIs came.

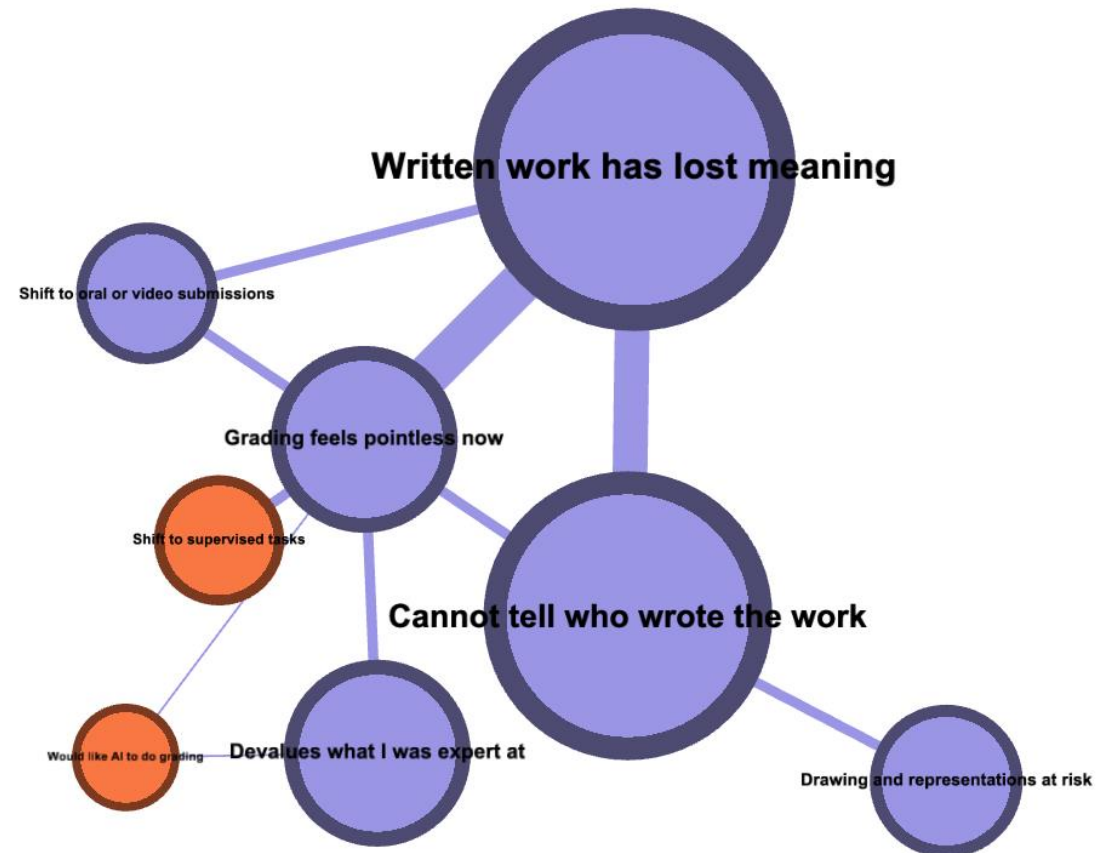


Assessment and trust crisis

- The strongest feeling in the whole dataset because it reaches past the tools and into the relationship between teacher and student.

I feel ridiculous as a teacher if I have to sit and read a task they didn't write themselves... we have to rethink the exam forms and the way we grade.

Some produce surprisingly detailed answers with technical terms they wouldn't normally use — seemingly to appear knowledgeable rather than to understand.

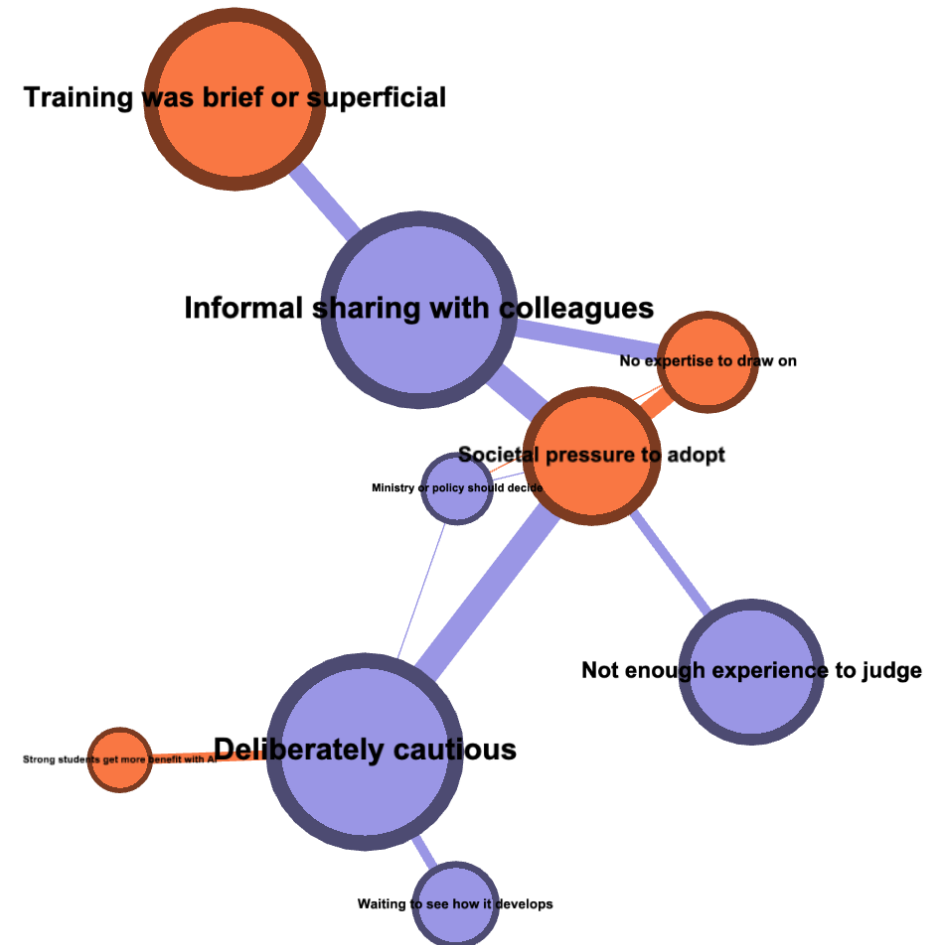


A crossroads with no framework

- Teachers' unease was not about GAI itself, but about being left to work it out alone.

I feel there's a societal pressure. No one has told me I have to use it. But I can feel it... like it's come past as a wind. And if I'm to help my students as well as possible, then I also have to.

You do get breathed down the neck. But who's coming to save us?... There must be someone who sets frames for our teaching and exams. But there just isn't. So we stand at a crossroads now.



Learning physics still needs friction

- Understanding develops through struggle and GAI may remove from this the learning process.

It simply can't provide the friction that learning requires. And if it tries to, there's resistance from the students. They'd rather have resistance through my teaching than through the machine.

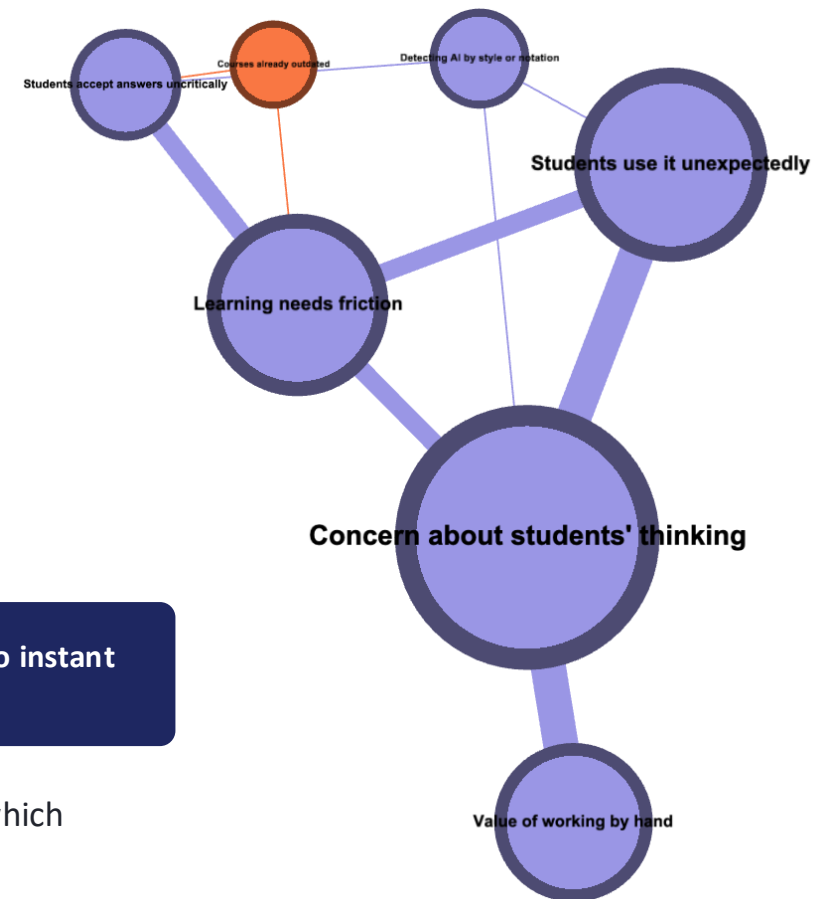
WHERE FRICTION LIVES IN PHYSICS

Drawing diagrams

Deriving relationships

Problems with no instant answer

A tool that supplies the answer directly may remove exactly the struggle through which understanding is built and consolidated.

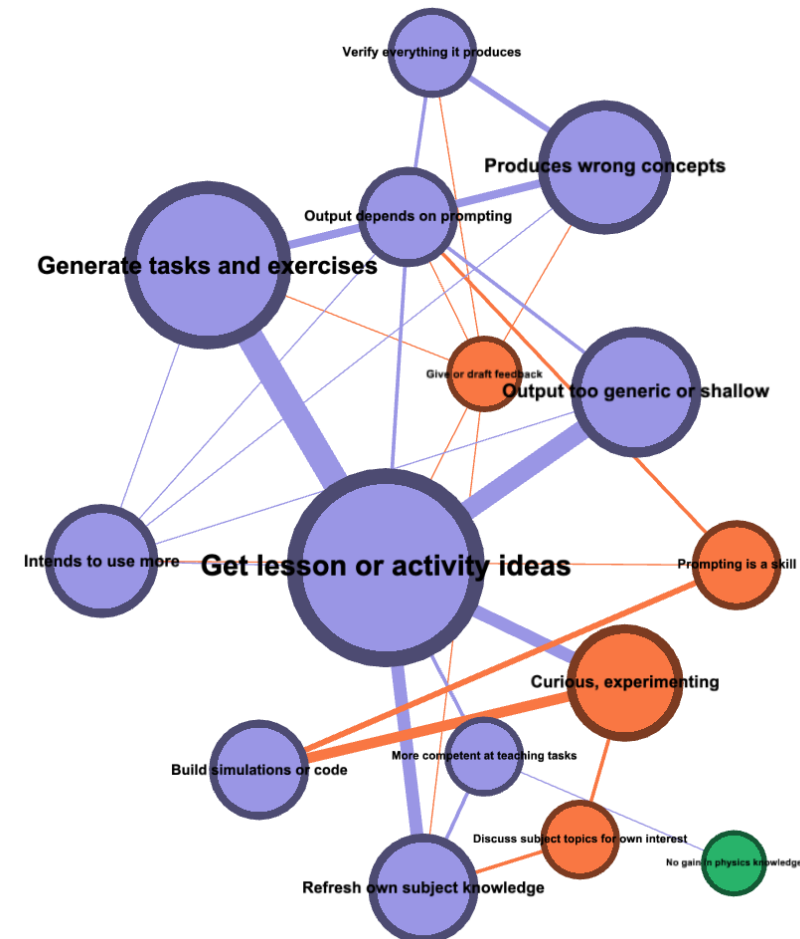


A machine without mind

- When asked what would make the teachers rely on GAI more, they didn't want a more powerful tool but they wanted a more trustworthy one.

I have a problem with that thing where it can't distinguish whether it's understood something or not. It doesn't have a concept itself. It just generates.

I used this AI to say: what if we have some radium, what does it decay to? Even though it claimed it could, it always went wrong. So it fell apart.



Conditional Autonomy

- On the surface, teachers made the pedagogical calls. On closer look, that control had limits.

Teachers choose:

- whether to use the tool
- what to use it for
- how to respond — reviewing, correcting, or rejecting outputs

“ I was about to say it was me. I think I set it in motion, but where it ends up, that I can't control. Of course, if I get it prompted well enough, then it's probably me who controls it.

Do teachers pick up the tool?

- The choice was largely made for them. Most used Copilot because leaders recommended it on data-protection grounds.
- Only three teachers had ever used more than one GAI system.

Divided Views on Competence

- Moderate users felt more capable at planning their teaching but no more knowledgeable in the subject itself.

“

I also feel more competent in that you can use it to develop your teaching. I don't think you become physics-subject smarter.

PEDAGOGICAL COMPETENCE ↑

- Faster access to teaching ideas
- Alternative ways to explain concepts
- Recall of content not taught for a while
- Less time on grading & admin

SUBJECT EXPERTISE → UNCHANGED

- Content felt shallow and unreliable
- Some tried it for physics, then stopped
- Supplies material, doesn't deepen understanding
- Teachers kept the two carefully apart

- A few teachers worried GAI devalued competence they had valued and grey lines they used to hold firm.

“

*I don't think it's very exciting to grade their tasks, especially not when I know they've gotten help from the chat. So, I'd like it if AI could solve that task for me... **when does it become okay that I use AI to grade the students' tasks?***

- Some human roles cannot be replaced by AI.

“

The part called teaching, what should we [work with], and how should you work with it, is not threatened. But that more roaming reference work is more threatened.

THREATENED

Being the trusted source of disciplinary knowledge.

DISTINCTLY HUMAN

The relational core of teaching

- Presence
- Attentive listening
- Motivating and responding to individual students

Suspicion and the loss of bond

- Relatedness was the SDT need teachers felt as most vulnerable where GAI strained their trust in students.

“ I'm clearly more skeptical toward what they say. There's become a shorter path for them to say something that sounds right, but which they haven't thought about themselves... I'm more reserved toward whether they thought their way to it.

WHAT TEACHERS NOW WATCH FOR

Facial expressions

Questioning
unexpected answers

Signs understanding
actually occurred

Rather than accepting contributions at face value, teachers now look for evidence of genuine thinking.

Teaching depends fundamentally on interaction.

- Knowledge develops through the reciprocal relationship between cognition and social interaction.
- When students can produce plausible answers without understanding them, that relational foundation starts to erode.

Conclusion

THROUGH UTAUT2

- Main concern: negative effects on students, who use GAI in unproductive ways
- Used mainly for low-stakes tasks: refreshing knowledge, teaching ideas, curiosity
- Little freedom to explore tools: schools directed them to Copilot for data protection
- Poor impression of output quality, while acknowledging it depends on the prompt

THROUGH SDT

- Shallow use: autonomy largely preserved; no sense of becoming a stronger physics teacher
- Teacher–student relation disturbed by loss of trust
- A distance gradually opening between teacher and students