

Safe and Effective AI for Learning





Greg Nadeau

Ed3 Convenor



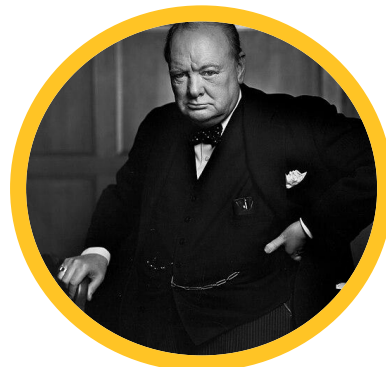
David Sallay

A4L



Steve Smith

A4L



**Sir Winston
Churchill's Ghost?**

England

AI Privacy, Security & Safety

Privacy

- AI is a broad term with seemingly endless uses, whereas privacy law focuses on specificity
- AI edtech has a lot that is still covered by traditional privacy laws with some unique issues (e.g., data retention, what is identifiable)
- Privacy not always the main concern, but it gives parents and students rights
- Major concern/opportunity is with transparency (not just that it is a black box, but that it is being used at all)

Security

- Does AI require additional security controls not typically required?
- If so GESS will be updated to include these
 - GESS Update

Safety

- AI presents new challenges when thinking about students interacting with AI directly
- What safeguards can we create to ensure students remain safe
- Current levels of concern are skyrocketing

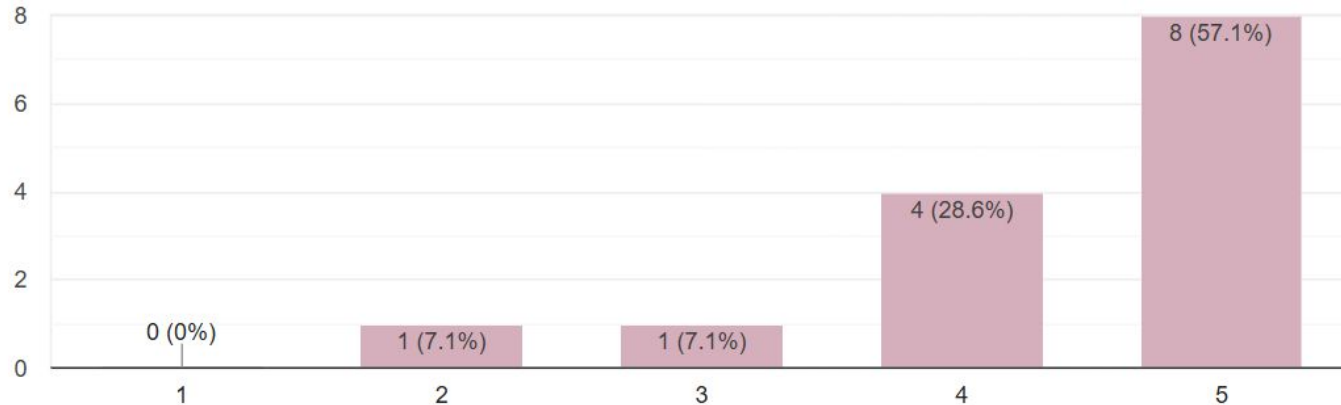
Is AI Safe?

Dario Amodei, CEO of Anthropic, recently published "[The Adolescence of Technology](#)." In it he says:

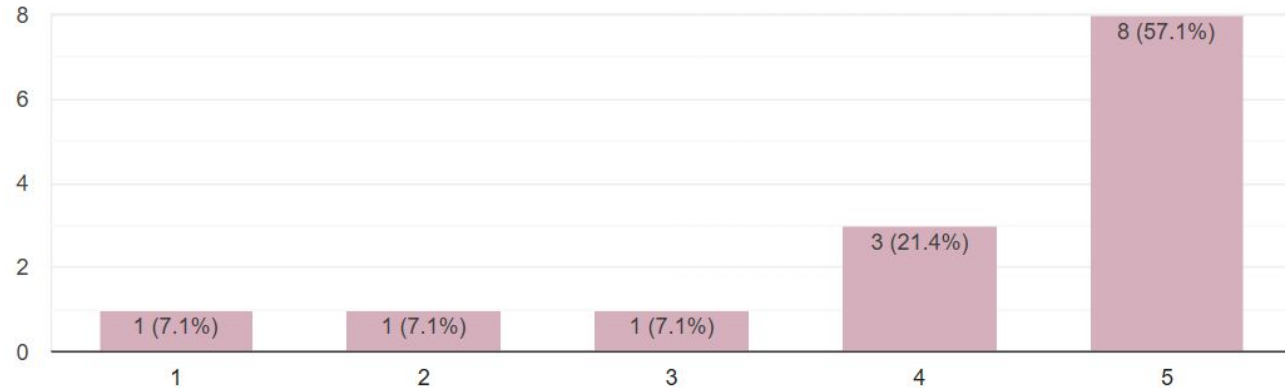
- Even if AI doesn't actively aim to attack humans and isn't explicitly used for oppression or control by states, there is a lot that could go wrong short of this, via normal business incentives and nominally consensual transactions.
- A world with billions of intelligences that are much smarter than humans at everything is going to be a very weird world to live in.
- We see early hints of this in the concerns about AI psychosis, AI driving people to suicide, and concerns about romantic relationships with AIs.
- As an example, could powerful AIs invent some new religion and convert millions of people to it?
- Could most people end up "addicted" in some way to AI interactions?
- Could people end up being "puppeted" by AI systems, where an AI essentially watches their every move and tells them exactly what to do and say at all times, leading to a "good" life but one that lacks freedom or any pride of accomplishment?



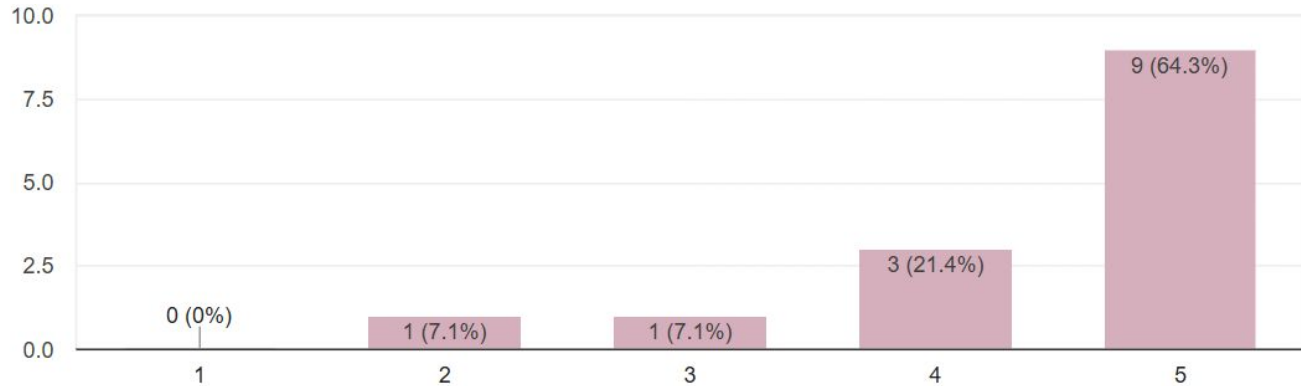
1. The huge potential benefits of daily student-AI learning are matched by massive risks.



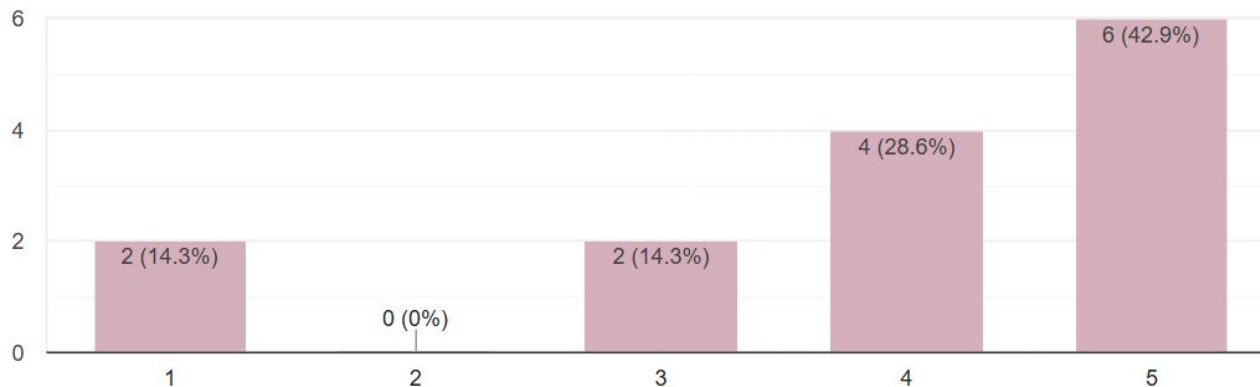
2. Learners, parents, and educators all want to know if AI is both safe and effective.



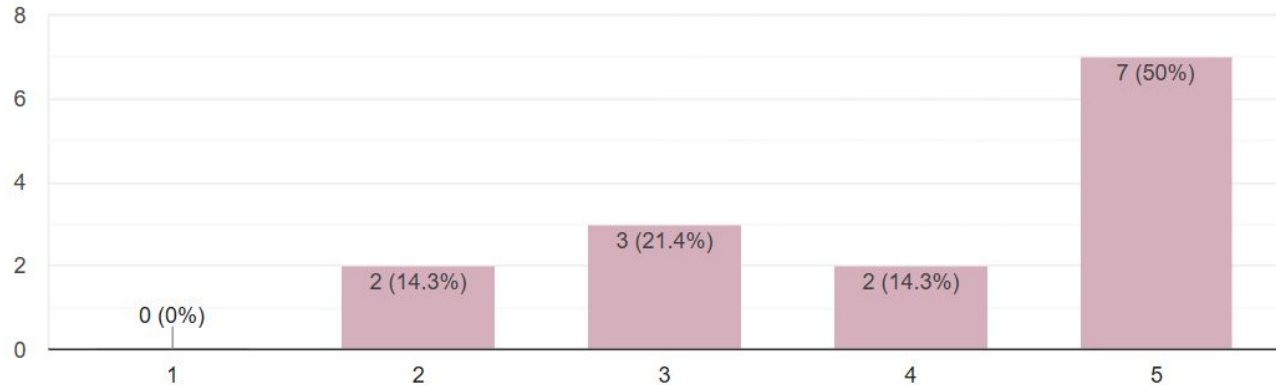
3. The time to act to create safety and efficacy assurances for minors is now.



4. Government AI controls only become bipartisan when focused on minors, but could become applicable for all ages



5. Ed3 and DSU will focus on enabling AI solutions, generating evidence, and improving system feedback loops



Create a tiger team to develop a strawman framework for safe and effective AI learning

	Bronze	Silver	Gold
Effective	Little evidence of significant improved learning outcomes	Some evidence of significant improved learning outcomes	<u>Strong evidence</u> of significant improved learning outcomes
Safe	Little evidence of pro-social safety controls	Some evidence of pro-social safety controls	<u>Strong evidence</u> of pro-social safety controls

Convene 2 panels of experts to review and make improvements to the frameworks and answer questions like:

- Do Khan Academy or Alpha School provide strong evidence of significant improved learning outcomes?
- Does Anthropic's System Card for Child Safety provide additional evidence of safety that OpenAI and Google do not?

AI effectiveness expert team

- **Do Khan Academy or Alpha School provide strong evidence of significant improved learning outcomes?**

- **Tom Kane:** Harvard professor, recognized for quantitative analysis of education data
- **Monica Bhatt,** University of Chicago professor who led research demonstrating the impact of high quality tutoring
- **Angela Duckworth:** Renowned for her research on "grit"—the combination of passion and perseverance—and its impact on student success.
- **Carol Dweck:** A Stanford psychologist famous for her work on "growth mindset," the belief that abilities can be developed.
- **Jo Boaler:** A Stanford professor focusing on mathematics education, promoting a growth mindset in math and advocating for reform in teaching methods.
- **Linda Darling-Hammond:** President of the Learning Policy Institute and a professor emerita at Stanford, focusing on teacher preparation and educational equity.
- **Pedro Noguera:** A sociologist whose work focuses on the ways in which schools are influenced by social and economic conditions.
- **Raj Chetty:** A Harvard economist whose research uses big data to study equality of opportunity and the long-term impact of teachers.
- **John Hattie:** Known for his meta-analyses of educational research, particularly his "Visible Learning" project, which ranks influences on student achievement.
- Facilitated by **Beth Holland** and **Beth Rabbitt,** FullScale Learning

AI safety expert dream team

- Does Anthropic's System Card for Child Safety provide additional evidence of safety that OpenAI and Google do not?
1. **Drew Bent:** Anthropic
 2. **Ben Gomes:** Google
 3. **Irina Jurenka:** Google DeepMind
 4. **James Donovan:** OpenAI
 5. **Yusuf Mahmood:** AFPI
 6. **David Sallay:** A4L
 7. **Julia Freeland Fisher:** Christensen Institute
 8. **Pete Just:** CoSN
 9. **Adam Billen** EncodeAI
 10. Facilitated by **Amelia Vance:** Public Interest Privacy Center and **Michael Klein:** Institute for Security and Technology (IST)

Gather a National Summit on Safe and Effective AI Learning for Minors

- National Parent Union
- Common Sense Media
- EncodeAI
- EDSAFE AI Industry Council
- Digital Promise
- NGA
- CCSSO
- SHEEO
- ISTE
- CoSN
- CGCS
- ACE

Create a challenge to motivate providers to submit initial evidence for review.

Re-convene the expert panels to review the review of evidence

January 2027: Publish first “consumer reports” of safe and effective AI learning

Ray Bradbury, anticipating LLMs in 1996

Suppose and then suppose and then suppose
That wires on the far-slung telephone black poles
Sopped up the billion-flooded words they heard
Each night all night and saved the sense
And meaning of it all.

Then, jigsaw in the night,
Put all together and
In philosophic phase
Tried words like moron child.

Thus mindless beast
All treasuring of vowels and consonants
Saves up a miracle of bad advice
And lets it filter whisper, heartbeat out
One lisping murmur at a time.

So one night soon someone sits up
Hears sharp bell ring, lifts phone
And hears a voice like Holy Ghost
Gone far in nebulae
That beast upon the wire,
Which with sibilance and savorings
Down continental madnesses of time
Says Hell and O
And then Hell-o.

*To such Creation
Such dumb brute lost Electric Beast,
What is your wise reply?*