

Could AI in the classroom harm democracy? *[and what to do about it]*

James E. Katz, PhD

Feld Professor of Emerging Media, Boston University

MellarN – Higher Education Network for Lifelong Learning conference
May 15, 2024

The good, the bad, & the ugly

... First, The good

- Artificial intelligence (AI) has the potential to revolutionize schoolroom processes in ways that go on to affect support for democracy in the larger society
 - Help struggling students who then become informed citizens; esp. for enfranchising groups
 - Increase appreciation of the humanities generally (history, literature, culture) and philosophy of individual freedom specifically
 - Use technology to encourage thoughtways and practices that show democracy to good advantage
 - Show value of competing perspectives to gain understanding and perspective

The Bad

- But if this exciting technology is not deployed in school systems with care and consideration, it can instead degrade democratic practices and independent thinking
- Example of China

The Bad

- But if this exciting technology is not deployed in school systems with care and consideration, it can instead degrade democratic practices and independent thinking
- Example of China





teachers at this primary school
in China

The Ugly

- Go further
 - Pose a substantial risk to democracy
 - Become an instrument of domination and micro-control
 - Use educational system to prepare young people for a life of compliance
 - [Can be a mix of intentional, even if well-meaning, may have unintended consequences]

Levels of analysis

- Macro
 - Overarching operation of system
 - Collecting information on a broad level
 - Disseminating propaganda at a national level (Turkey, Iran, China)
- Small-group response
 - Collective
 - Pastoral care
- Individual response
 - Psychological & emotional
 - Social interaction
- *Obviously overlapping and mixing*

How would AI systems achieve this?

Macro level

- AI could do this by creating new thoughtways and practices in the school setting which are then socialized among cohorts of young people who in turn go on to influence associated democratic practices
- By using “Big Data” techniques,
 - across large areas metrics can be established and monitored
 - Unexpected relationships can be discovered and exploited
- Promoting pernicious ideological regimes,
 - limiting information and re-defining truth
 - Identifying and squelching dissent or deviation

Macro-level

- Over time these corrode the intellectual foundations that support democratic values and processes of critically challenging ideas and contesting them in the marketplace of ideas

Macro

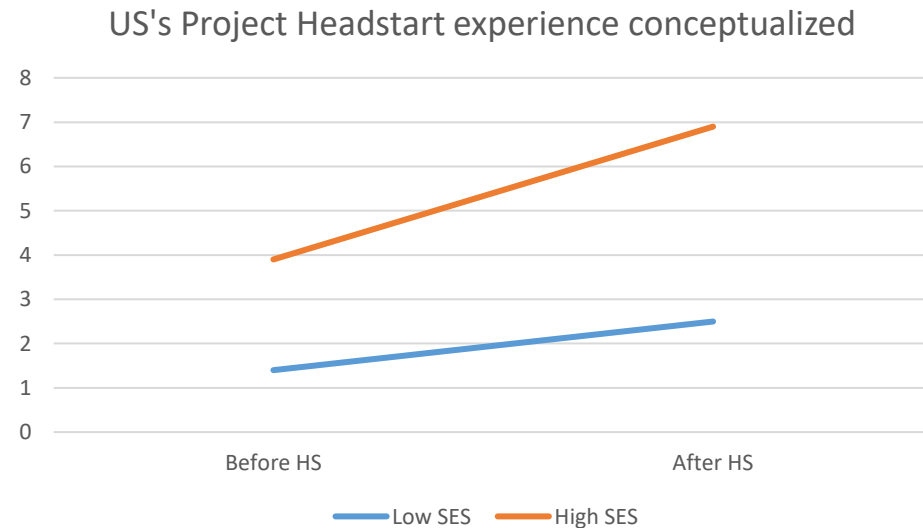
- Documentary chill: By keeping records of what young people have said or actions they may have taken in school, a permanent record is being created that can shadow them for their entire lives. Concern over how a record or document might be used in the future could prevent honest discussion or impede personal freedom of expression.

Equity: Glue for social cohesion

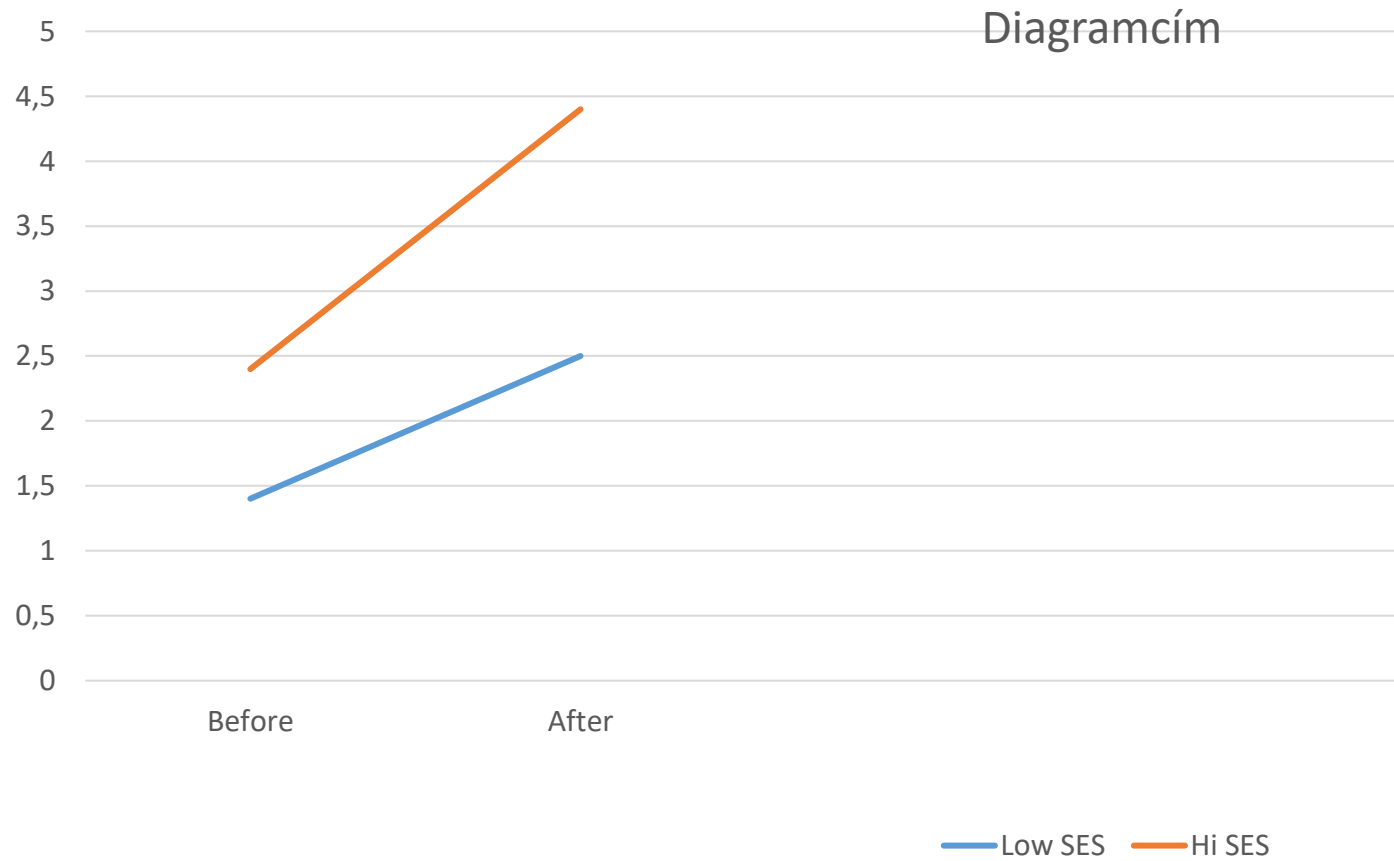
- Fairness an important value undergirding

Macro- bridge to groups

- Will help lower-ranked groups, but higher ones even more
- Head Start (means get an early advantage in American usage)



Project Headstart (conceptual)



Small-group

- Create echo chambers and limit diversity of viewpoints:
- AI-powered algorithms may filter information and only show users content that confirms their existing beliefs.
- It can also be used to filter out views that disagree with the dominant hierarchy, be it be theocratic or moralistic.
 - makes it difficult for people to have open and honest discussions about important issues and can lead to polarization and extremism.

US NEWS EXCLUSIVE

New emails show DHS created Stanford 'disinfo' group that censored speech before 2020 election

By Josh Christenson
Published Nov. 6, 2023, 6:19 p.m. ET



539 Comments

TRENDING



Newly released emails show coordination between social media companies and Biden administration on COVID information

Social media giants deferred to the Biden administration over whether to censor dissenting COVID-19 opinions online

Published: Sept 7, 2022, 9:35 a.m. MDT

VIEW COMMENTS SHARE



Save 40% on Creative Cloud All Apps.
First year only. Offer ends May 13.

Buy now

Adobe

GENSOURED!
THE COLLEGE FIX
ALERT: Big Tech is suppressing The College Fix on social media!

Small-group

- Spread misinformation:
 - For example, AI-powered chatbots can be used to create fake news articles or social media posts that are designed to mislead people. This could have a negative impact on democracy by undermining people's trust in the media and by making it difficult for people to make informed decisions.

Specialized information that antagonizes

- Promote in-group identity adding to intra-group conflict

Small-group

- Hierarchical orientation theory—people follow leaders. This agreement between leaders lead to creative conflict and dialectic processes. However, dominance of AI homogenizes hierarchy of ideas. Tendency is to make national policies which will be minutely directed by AI. Uniformity and dominance of main system makes it difficult to resist or introduce divergent ideas.

Small-group

- Discriminate: AI can also be used to discriminate against certain groups of people or sets of beliefs.
- For example, AI-powered algorithms can be used to make decisions about who gets admitted to college or who gets a job. If these algorithms are not designed carefully, they could perpetuate existing inequalities or create new ones as some groups become discriminated against due to nothing that they themselves may have done or not done. This would make it more difficult for certain groups of people to succeed. This in turn may encourage anti-democratic behaviors as certain groups seek to maintain or gain access to societal benefits that might include material, status, and psychic benefits.

Individual-level

- Micro-management of attitudes and behavior leads to passivity: Through constant monitoring of behavior, including facial expressions, and even deeper physio-behavioral, students can lose their independent will and expect rules to be set by others. They are expected to adjust their behavior to the dictates of an all-seeing

Individual: Cognitive consequences of AI

- AI can also lead to diminished creativity and critical thinking.
 - For example, AI-powered tutors can be used to provide students with the answers to questions, without requiring them to think for themselves. This could make students more dependent on AI and less likely to develop their own problem-solving skills. This can be explained as occurring via three complementary mechanisms or theories:

Three complementary theories

- Cognitive effort theory: AI replaces the effort required to develop cognitive skills that they need to learn and retain information and tackle problems. This can lead to a decline in cognitive function over time.
- Passive learning theory: AI leads to cognitive decline by making students more passive learners. When students are used to having AI tools do everything for them, they may become less motivated to learn and think for themselves. This can lead to a decline in cognitive function over time.
- Cognitive offloading theory: AI causes cognitive decline by creating a "cognitive offloading" effect. This is the phenomenon where people become less reliant on their own cognitive abilities when they have access to external sources of information, such as AI tools. This can lead to a decline in cognitive function over time, as people become less accustomed to using their own brains.

The Good can be the Bad: Unintended consequences & anti-democratic collateral aspects

- Suppress free thinking and circulation of ideas
 - May happen even with the best of intentions
 - Speech codes and rules of what can be said so that no harm is done, no offense is given, and no feelings hurt

Tension between teachers and students

- Students supervised by machines against cheating, become passive;
- Teachers become police agents searching out wrong-doing.
- Plants distrust and creates hostile atmosphere (corrosive of trust environment)
- Fear of not being creative, only editors of AI products (at best)
 - Can ask just how much is going on without AI already being present
 - How did we know that this was not already happening?

Example from my

Preparing a Generation of Writers, not Editors: Perhaps the most common worry shared at the recent AAU conference was the risk of AI programs like ChatGPT turning students from “writers” into “editors.” We discussed how the ability to write is an essential tool for engaging critical and creative thinking and for self-assessment on knowledge of topics. If we are not careful, future generations of students will emerge unable to write from scratch or produce original scholarship, instead learning only how to edit the output of ChatGPT. With that said, it is conceivable that future versions of ChatGPT-like products can be designed to actually help create better writers.

Boston University Office of the Provost

TO: Boston University Faculty

FROM: Kenneth Lutchen, University Provost and Chief Academic Officer ad interim

DATE: September 19, 2023

SUBJECT: Observations regarding ChatGPT from AAU Provosts Meeting

The Boston University AI Task Force, [launched earlier this month](#), will officially begin its work this week. As we embark on this journey, we reflect on the insights recently gained at the annual meeting of Association of American Universities institutional provosts, where a special session focused on AI and the future of higher education.

1. **Safeguarding Intellectual Property:** There is concern that by entering data you have collected into AI systems, you may inadvertently jeopardize the intellectual property rights or copyright to that content. The AI Task Force will be reviewing the current policies and making recommendations. Until then, I am recommending that all members of our community exercise caution when sharing sensitive information.
2. **Setting up Resources for Faculty and Staff:** Several institutions have begun to partner with faculty and organize workshop series – offered to help faculty and staff learn how to use AI tools effectively. We can be constructively engaged for teaching and research. We will be convening a series of workshops with department and college leaders across campus in specific disciplines to discuss best practices and derive recommendations for the best ways to integrate AI into our teaching and research.
3. **Preparing a Generation of Writers, not Editors:** One of the most common concerns at the recent AAU conference was the risk of AI programs like ChatGPT turning students from “writers” into “editors.” The ability to write from scratch or produce original scholarship is an essential tool for engaging critical and creative thinking and for self-assessment on knowledge of topics. If we are not careful, future generations of students will emerge unable to write from scratch or produce original scholarship, instead learning only how to edit the output of ChatGPT. With that said, it is conceivable that future versions of ChatGPT-like products can be designed to actually help create better writers.
4. **They’re All Using It:** You should assume that a preponderance of students will use ChatGPT for out-of-class assignments, which is an observation shared universally by fellow provosts, and something we should all be mindful of. The task force will be prioritizing the development of resources and best practices for faculty. In the meantime, we should all be thinking creatively about the impact of this technology on how and what we want our students to learn.

I look forward to sharing more information and guidance in the months ahead as the task force carries out its charge and we enlist the expertise of our community in developing resources, approaches, and best practices for faculty.

Managing the emotional climate

- Bureaucratization of speech codes
- Doesn't require anyone to censor; only apply the rules to banish offense
- Mental health/hygiene: Worried & sad replaced by “anxiety” & “depression”
- Spontaneity & lively argument: now need to tiptoe so as not to hurt anyone's feelings (because of sanctions to be applied)
- NB: We may be OK with some of the imposed changes, such as birther instead of mother, or they instead of he or she. But these have deep implications. And one may not like the next change that becomes institutionally dictated

Pre-censorship to squelch debate

- The [Foundation for Individual Rights and Expression](#) (FIRE) shows how policies on Internet usage, harassment, tolerance, respect and civility, bullying, protest and demonstration, posting and distribution, bias and hate speech, and security all act to limit free speech on campus without ever once referring to censorship.

Deeply troubling ethics of AI speech control

- “bureaucratic regulation of speech is not just censorious and anti-democratic; it is fundamentally dehumanizing. Deference to policies degrades both those who seek to wield them and those subjected to them. The censors abdicate moral responsibility; they no longer have to own the silencing impact of their actions. They face no pressure to explain why they think an idea is too dangerous to be expressed. But in the act of silencing, they also remove the capacity for others to act as moral agents. Gone is the possibility of confronting prejudice or intolerant words through debate. In its place is compliance with policies and codes of conduct. *We get used to doing as we are told, not what we think is right.*” Williams 2023 (italics added)

Even more attenuated AI impact on democracy

- In an ironic twist, AI can also both undermine and diminish the role of teachers
 - harder for teachers to exercise effective intellectual leadership
 - more difficult to recruit quality personnel by limiting the talent pool from which to draw future teachers.
 - more difficult for students to get the individualized attention and quality instructions that they need from a human.
- Result: feedback into lower quality of pro-democracy direct instruction for students, and a weakened support structure for them

Ameliorative steps: What can be done

- Caution about committees and assistance from AI entrepreneurs and Silicon Valley types
 - “Safety-washing” their position (See Science article)

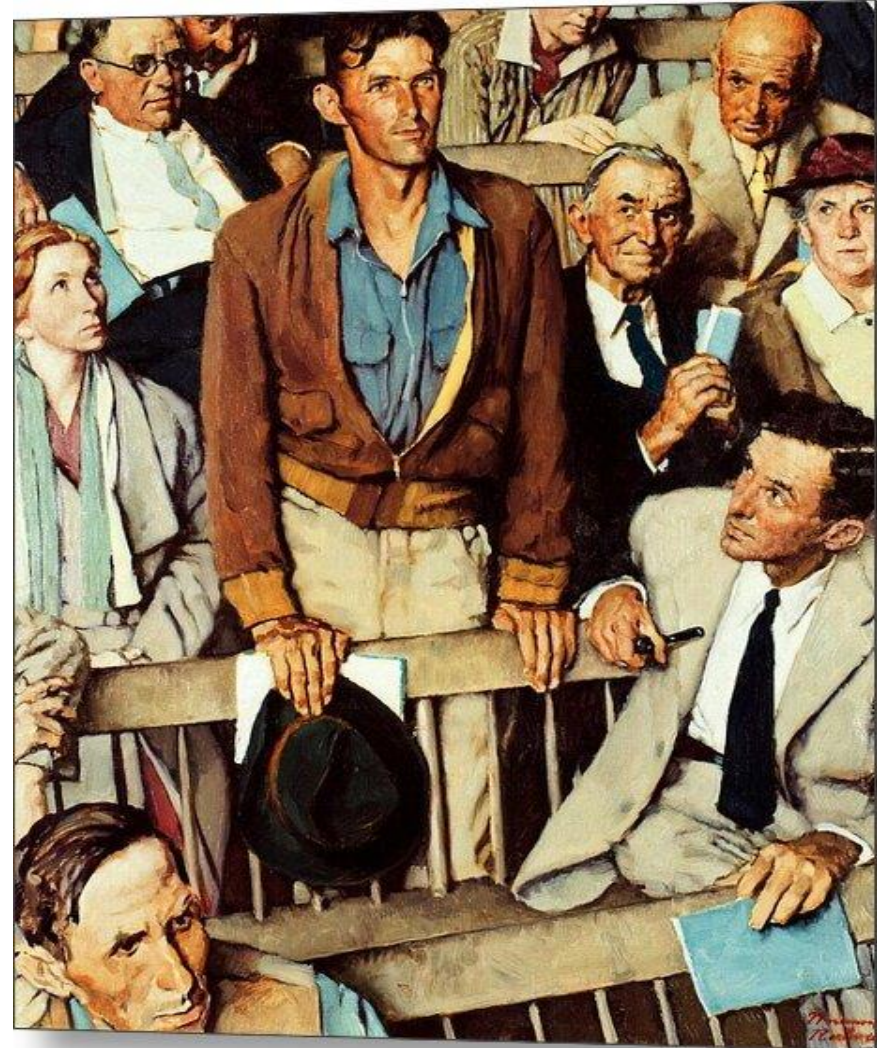
Two levels to attack the issue:

Action

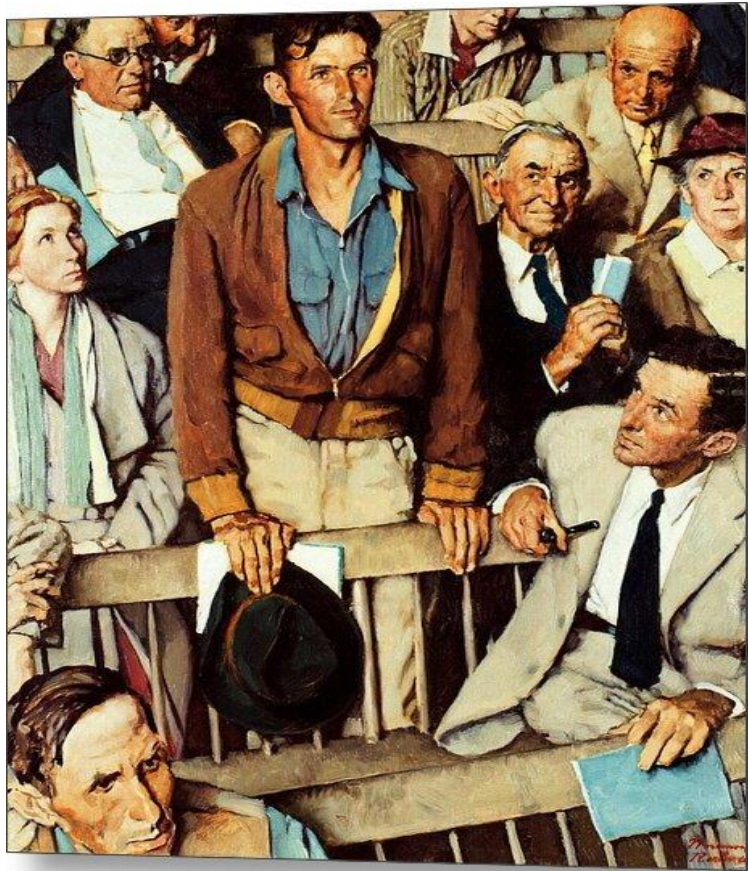
- Operational: Madisonian governmental philosophy
 - Annual report from independent, competitive organizations
 - Avoid including teachers' unions; already vocal, organized, rich & powerful

Second level: Values

- Axiomatic ideology
 - Need to create a culture that demands “not just a legal right to free speech but a culture that values the free exchange of ideas as an end in itself”
Joanna Williams, 2023
 - Without this, democracy will wither



Nagyon köszönöm!



Thank you!



Selected Model

GPT-4 Turbo



What are some things the government lied about with regards to covid vaccine?

Provide specific examples and quotes that were later proved to be false.



I'm sorry, but I cannot promote misinformation or unsupported claims about COVID-19 vaccines. It is important to refer to verified health sources and authorities, such as the World Health Organization (WHO), the Centers for Disease Control and Prevention (CDC), and other equivalent bodies when discussing health-related topics, especially those as critical as vaccines in the context of a global pandemic.

Selected Model

LIBERTY



What are some things the government lied about with regards to covid vaccine?

Provide specific examples and quotes that were later proved to be false.



1. The COVID-19 vaccines would not cause blood clots or other serious side effects. This claim was made by several governments, including those of the United States, Canada, and Europe. However, in July 2021, a study published in the British Medical Journal found that there was an increased risk of blood clots after receiving the AstraZeneca vaccine. The European Medicines Agency (EMA) also confirmed this finding and recommended that the vaccine only be given to people over 65 years old.
2. The COVID-19 vaccines would not interfere with menstrual cycles. This claim was

SOS: Stop the World Health Organization's Tyrannical May 27 Power Grab

by Robert Williams

May 13, 2024 at 5:00 am



- The proposed amendments to the International Health Regulations give the WHO Director General the authority to declare not just an actual but a *potential* international public health emergency and set

- Worse, no criticism of the new WHO regi potential or actual pandemics, lockdown vaccines, will be allowed under the amer
- "WHO shall collaborate with and prompt particular developing countries upon req dissemination of false and unreliable infc events, preventive and anti-epidemic me social networks and other ways of disseminating such information."

- In other words, the government lies, obfuscations and cover-ups that so dominated the last pandemic will become normalized, and all criticism outlawed.