



Ethical Use of AI in STEM Education Research

May 15, 2024 | 2-3:30 PM ET

Learn more at go.edc.org/learning-series-AI



Ilana Horn
Vanderbilt University
(Moderator)



Tiffany Barnes
North Carolina State
University



Joshua Danish
Indiana University



**Samantha
Finkelstein**
Carnegie Mellon University



Ole Molvig
Vanderbilt University

Agenda

- Introductions
- Presentations by Panelists (Barnes, Danish, Finkelstein, Molvig)
- Discussant comments (Horn)
- Q & A with audience

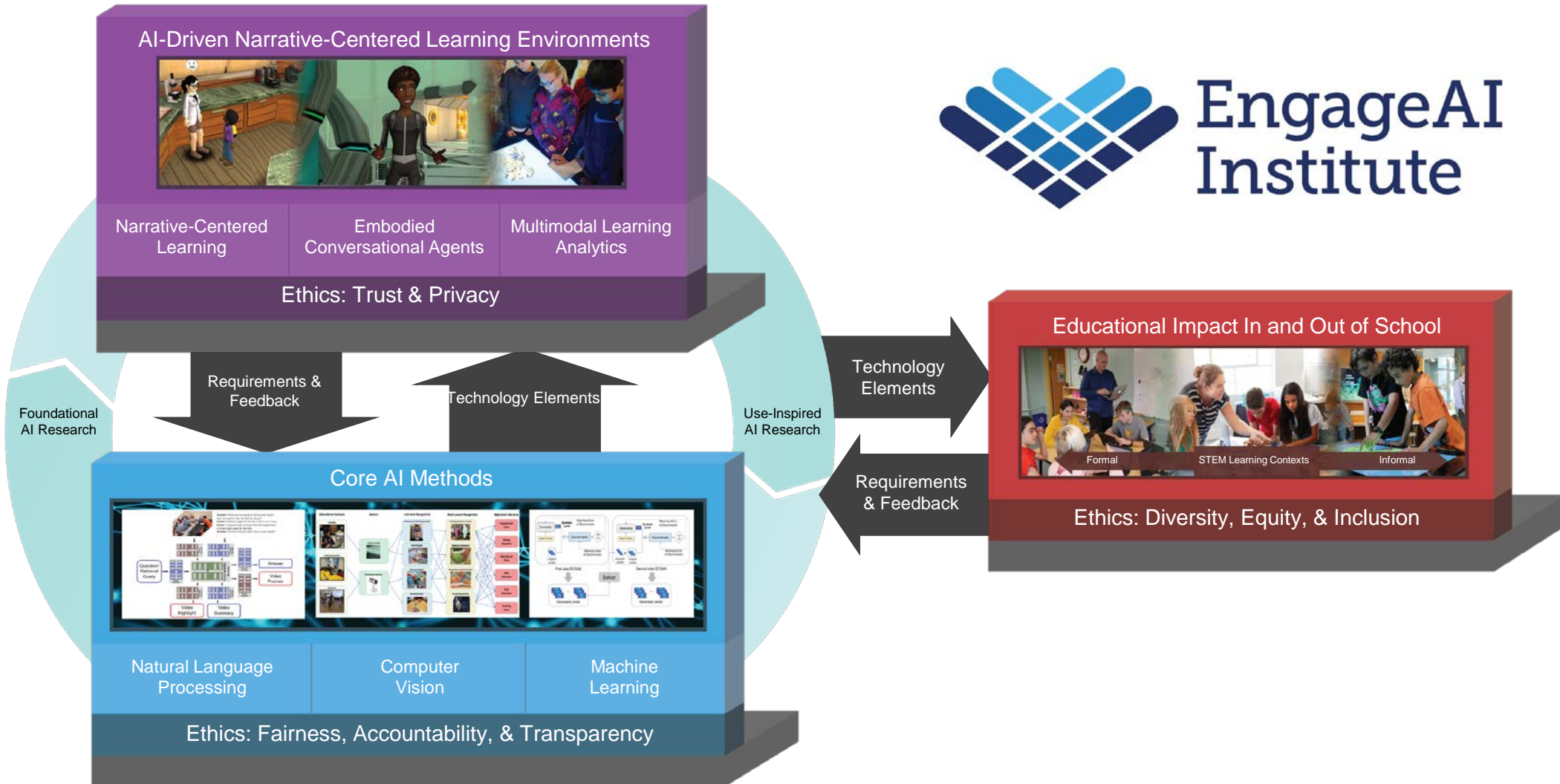
NC STATE UNIVERSITY

Panelist - Tiffany Barnes

North Carolina State University



NSF AI Institute for Engaged Learning



STARS Computing Corps

Developing leaders in broadening participation research & practice



Annual events & programs

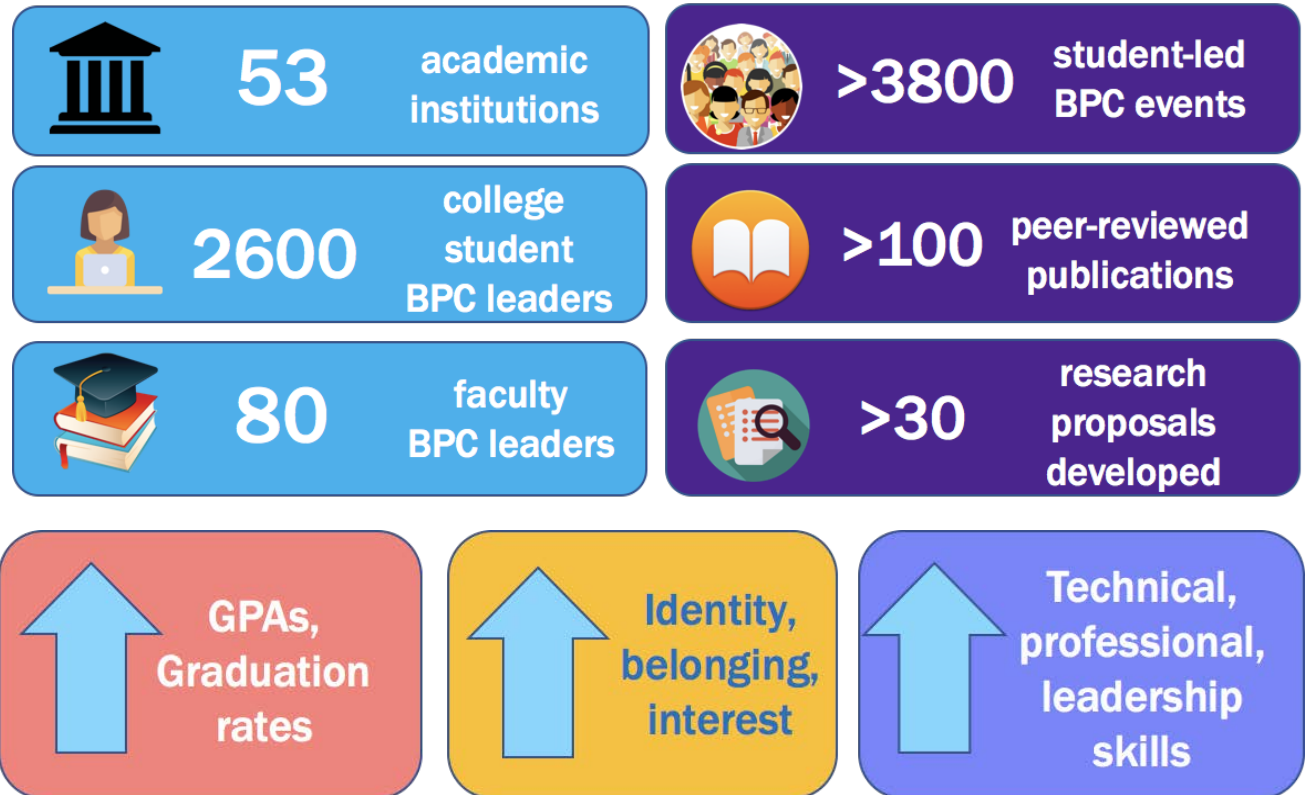
RESPECT research conference

STARS Celebration conference

Faculty & student webinars

STARS student chapters

STARS AI Scholars co-sponsors:



People



Projects

AI & Data-driven Learning Environments and Analytics, especially for problem solving help and progress support

The interface displays a logic puzzle with the following components:

- Truth Tree:** A diagram showing logical steps:
 - 1: $A \rightarrow C$
 - 2: B
 - 3: $C \rightarrow E$
 - 4: $D \wedge \neg E$
 - 5: D (Simp)
 - 6: $A \rightarrow E$ (HS)
 - 7: $\neg A \vee E$ (Impl)
 - Subgoal: $\neg E$
 - Goal: $\neg A \wedge B$
- Rules Menu:**
 - MP (Modus Ponens)
 - MT (Modus Tollens)
 - DS (Disjunctive Syllogism)
 - Add (Addition)
 - Simp (Simplification)
 - Conj (Conjunction)
 - HS (Hypothetical Syllogism)
 - CD (Constructive Dilemma)
 - DN (Double Negation)
 - DeM (DeMorgan's)
 - Impl (Implication)
 - CP (Contrapositive)
 - Equiv (Equivalence)
 - Com (Commutative)
 - Assoc (Associative)
 - Dist (Distributive)
- Problem Info:** Problem Code: 2.1, Level: 2/7, Problem: 2/4
- Buttons:** Get Suggestion, Try to derive $\neg E$

The interface shows a 3D maze environment with a red robot and blue blocks. The command list on the left includes:

- Command Count: 13
- Fire Bot
- Set Variable $a = 0$
- Turn Right
- Repeat While $a < 3$
- Move Forward (multiple instances)
- Turn Right
- Turn Left (multiple instances)
- Function A (multiple instances)

The action menu on the right includes: Move Forward, Turn Right, Climb Up, Move Backward, Turn Left, Climb Down, Pick Up, Put Down, Wait.

The interface shows a programming environment with the following elements:

- Code Blocks:** A 'repeat 3' block containing 'move 10 steps' and 'turn 120 degrees'.
- Objective 2:** Draw something correctly. Instructions: 1- Pen down (under Pen Category): puts the pen on the stage. 2- Move (under Motion Category): makes the sprite move. A visual example shows a pen down block followed by a move 10 steps block resulting in an arrow.
- Progress Dashboard:**
 - Objective 1: Interact with User
 - Objective 2: Draw something correctly (marked with a red X)
 - Objective 3: Draw using Repeat blocks (marked with a green checkmark)
 - Objective 4: Draw Polygons angles correctly

Programming Games & Supports

Framing Common AI Ethics Principles

Justice	Respect	Beneficence																
Responsibility to distribute burdens & benefits equitably	Responsibility to protect human rights & dignity	Responsibility to benefit people & minimize harm																
<table border="0"> <tr> <td>Accountability</td> <td>Equity</td> </tr> <tr> <td>Fairness</td> <td>Inclusion</td> </tr> <tr> <td>Sustainability</td> <td>Diversity</td> </tr> </table> <p><u>Justice / Anti-Oppression by:</u></p> <table border="0"> <tr> <td>Culture</td> <td>Gender</td> </tr> <tr> <td>Place</td> <td>Race</td> </tr> <tr> <td>Language</td> <td>Disability</td> </tr> <tr> <td>Economic class</td> <td>Identity</td> </tr> <tr> <td>Social class</td> <td>Role</td> </tr> </table>	Accountability	Equity	Fairness	Inclusion	Sustainability	Diversity	Culture	Gender	Place	Race	Language	Disability	Economic class	Identity	Social class	Role	<p>Transparency Explainability</p> <p><u>Protect Rights of:</u></p> <p>Privacy Autonomy Freedom AI Literacy Human Dignity Social Relationships</p>	<p>Non-maleficence (do no harm) Pedagogical Appropriateness</p> <p><u>Beneficence for:</u></p> <p>Students Parents Teachers Classrooms Society Environment</p>
Accountability	Equity																	
Fairness	Inclusion																	
Sustainability	Diversity																	
Culture	Gender																	
Place	Race																	
Language	Disability																	
Economic class	Identity																	
Social class	Role																	
<u>Who</u> & <u>Where</u> are people and places with benefits?	<u>How</u> are people prioritized & how are data and decisions handled?	<u>Why</u> will the work improve STEM education?																

Commonly cited AI Ethics Principles in bold

Khan, A. A., Badshah, S., Liang, P., Waseem, M., Khan, B., Ahmad, A., ... & Akbar, M. A. (2022, June). Ethics of AI: A systematic literature review of principles and challenges. In Proceedings of the 26th International Conference on Evaluation and Assessment in Software Engineering (pp. 383-392).

Panelist - Joshua Danish

Indiana University

Ethical AI Design Reflection Map

Embodiments that leverage AI in the design

Anticipated processes in the learning context

Measurable outcomes

Overarching project design

- A specific form of “intelligence” will help lead to measurable benefits for some participants.
- AI Needed qs
 - AI Environmental cost qs
 - AI equitable access and justice qs

- Interface for learners
- AI sensing & privacy qs
 - AI training and bias qs
 - AI text or art qs
 - AI labeling / decision qs

- Interactions between learners and each other or the software
- AI sensing & privacy qs
 - Real-time usage qs

- Benefits for learners
- AI literacy
 - Personalized support
 - Content learning
 - AI labeling / decision qs

- Interface for teachers
- AI sensing & privacy qs
 - AI training and bias qs
 - AI labeling / decision qs

- Interactions between teachers and software
- AI sensing & privacy qs
 - Real-time usage qs

- Benefits for teachers
- AI literacy
 - Support for differentiation
 - Support for orchestration
 - Assessment & decision-making guidance

- Interface for researchers
- AI sensing & privacy qs
 - AI training and bias qs
 - AI labeling / decision qs
 - AI data / model sharing qs
 - AI text or art qs

- Other interactions observed by AI tools (e.g., video recordings)
- AI sensing & privacy qs
 - AI data / model sharing qs
 - Real-time usage qs

- Benefits for researchers
- New forms of understanding about learning
 - New models or data
 - AI privacy, bias, and data sharing qs (all of them!)

Other **consequences** for participants, their community, the environment, creators, etc.

Panelist - Samantha Finkelstein

Carnegie Mellon University

- “The Purpose of a System is What It Does”

- Maps | Classrooms | AI applications

*Adrift in a world in which everything and anything is possible, **thinking** is the only activity standing between ourselves and the most heinous of evils.*

- Design represents and reveals ideologies

- Hannah Arendt

- The ceiling of how ethical an AI application can be is set by the structure in which that application will be deployed
- Talking meaningfully about educational equity requires talking honestly about which existing *status quos* we are and are not currently addressing
- The real ‘best practice’ is *be very honest* about your premises and outcomes.
 - What data do I have? What interpretive leaps am I making about people from that data?
 - *What am I doing to people based on my interpretation of that data?*

Panelist - Ole Molvig

Vanderbilt University

Model Cards For Educators

Prototype for a model card for education, leveraging ethics, technical, and pedagogical expertise, to be included in 2 upcoming publications

Model Card For Education Template		
Model/Tool Title		
<p>Pedagogical Evaluation</p> <p><i>This section uses learning goals and contexts as an entry point for thinking about the function and application of a model.</i></p> <p>Learning Goal Alignment: How does this technology support my learning goals for students? What are the affordances and risks of using this tool compared to others?</p> <p>Learning Context: How/where will the model be used? Does this match with the developers' intended use? Does the training data reflect the learner population? If historical training data was used, how might this perpetuate bias for some learners?</p> <p>Logistics & Material Requirements: How easy is it to log in? Is this blocked by your district? Are there age restrictions? What resources (internet, devices, plugs, etc.) are required to use it?</p> <p>History of (educational) use: Has this tool been used and/or researched in educational settings? What have others reported about it? Are any sample use policies available?</p> <p>Technical Education: What should students and teachers know about how this tool works to use it critically?</p>	<p>Ethical Evaluation</p> <p><i>This section uses the Ethical AI/ED Framework to evaluate models for justice, respect for persons, and beneficence.</i></p> <p>Justice: Have adequate measures been taken to reduce bias (racial, linguistic, ability, etc.) in the performance of the tool? What plan is in place for mitigating harm from bias? How might the use of this model privilege certain groups and/or marginalize others?</p> <p>Respect: Have students, families, and teachers consented to using the model (including any surveillance necessary for its operation)? Do they understand important information about how it works and why it's being implemented? What data does the model need to operate, and how is it stored? How are student surveilled, and where does this information go?</p> <p>Beneficence: Is the tool safe and effective for students and teachers? Do the benefits of using the tool for the learning goal outweigh the risk of harm? How can students and educators opt out of using the tool or challenge its results/predictions?</p>	<p>Technical Evaluation</p> <p><i>This section reports on technical functioning and benchmarks relevant to ethical and effective implementation in educational settings.</i></p> <p>Training: What data was used to train the model? How was the data obtained? Does the training data reflect the population using the tool?</p> <p>Version Information: Is the tool in Beta (trial)? Is this a rebranded version of a foundational model?</p> <p>Cost: How does the tool make money (e.g., licensing, subscriptions, advertising), if applicable? Does the payment structure or access change over time (e.g., free trials, free for use but not download)?</p> <p>Developer: Who made the tool, and why? Did they consult educators and students in the design process?</p> <p>Benchmarking: What measures of model performance, especially related to bias/fairness, are available?</p> <p>Explainability: How does the model work? Can humans explain its results?</p> <p>Environmental Impact: What are the environmental costs of training and running the model?</p>
<p>Card compiled by: _____</p> <p>Date created: _____ Date last updated/checked: _____</p> <p>Sources/Further Reading: _____</p>		





Discussant Comments

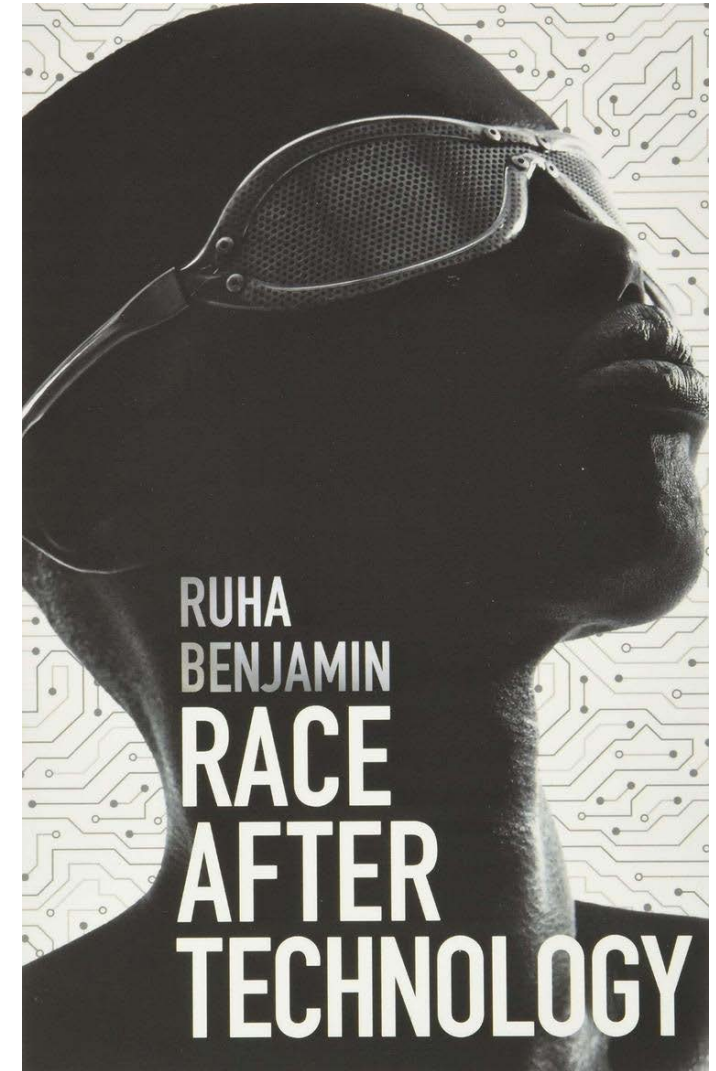
Appreciations

- The Belmont Report is a widespread and familiar ethical framework for researchers
- The brief offers **concrete tools** to investigators and users of AIED
- The authors aim to **center justice** in their framing of the ethical issues

Questions

“Jim Crow practices feed the ‘New Jim Code’ – automated systems that hide, speed, and deepen racial discrimination behind a veneer of technical neutrality.”

– *Ruha Benjamin*



Lessons from Biomedical Ethicists

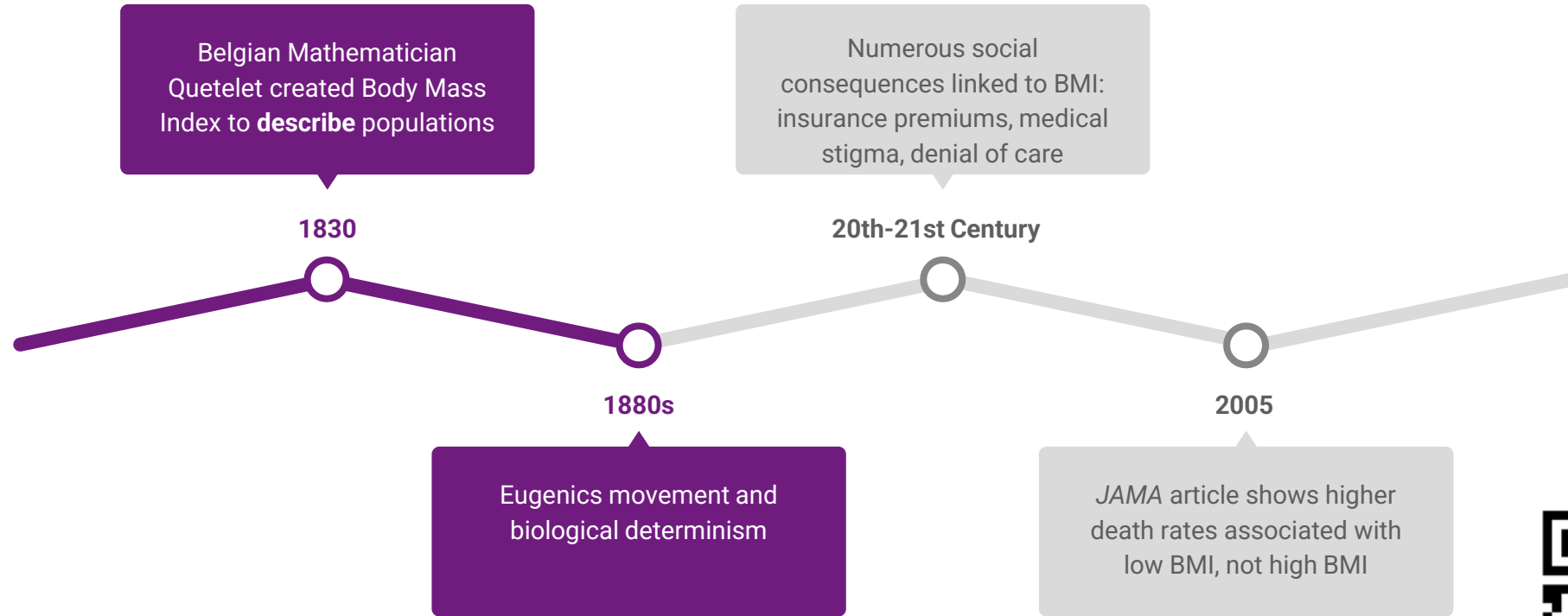
“While the Belmont Report was an impressive response to the ethical issues of its day, the field of research ethics involving human subjects may have outgrown it.”

Friesen, P., Kearns, L., Redman, B., & Caplan, A. L. (2017). Rethinking the Belmont Report? *The American Journal of Bioethics*, 17(7), 15–21. <https://doi.org/10.1080/15265161.2017.1329482>

Belmont report places too much emphasis on individual choice, does not consider potential harms to nonparticipants, and does not account for new modes of research with human subjects.

Brothers, K. B., Rivera, S. M., Cadigan, R. J., Sharp, R. R., & Goldenberg, A. J. (2019). A Belmont Reboot: Building a Normative Foundation for Human Research in the 21st Century. *Journal of Law, Medicine & Ethics*, 47(1), 165–172.
doi:10.1177/1073110519840497

Example from Biomedical Ethics: BMI Index





Q & A

Please use the Q & A function in Zoom to
post your questions

CADRE Resources for You | cadre@edc.org

- 1 [CADREK12.org](https://cadrek12.org) | Access NSF Proposal Toolkit, solicitation webinar recordings, project descriptions and products
- 2 [CADRE Newsletter](#) | Subscribe to keep up-to-date with DRK-12-related news and events
- 3 [@CADREK12](#) | Stay engaged with the DRK-12 community



Visit cadre.edc.org

Learn more about DRK-12 work
and explore our free resources!



Community for Advancing
Discovery Research in Education

This project is funded by the National Science Foundation, grant # 0822241, 1449550, 1650648, 1743807, 1813076 and 2100823. Any opinions, findings, and conclusions or recommendations expressed in these materials are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.