Evolving education in an AI world

impacts, opportunities & challenges

Mike Klymkowsky, Ph.D. (soon retired professor, MCDB / CU Teach) 20 May 2025

much in the news...

PRESIDENTIAL ACTIONS

ADVANCING ARTIFICIAL INTELLIGENCE EDUCATION FOR AMERICAN YOUTH

Executive Orders | April 23, 2025

much in the news...

PRESIDENTIAL ACTIONS

ADVANCING ARTIFICIAL INTELLIGENCE EDUCATION FOR AMERICAN YOUTH

Executive Orders | April 23, 2025

OPINION

By Jessica Grose
Opinion Writer

A.I. Will Destroy Critical Thinking in K-12

May 14, 2025

A New Headache for Honest Students: Proving They Didn't Use A.I.

nytimes.com

overview: (some) genAI issues

spontaneous questions welcome

overview: (some) genAI issues

- institutional: role of / impacts on summative assessments

spontaneous questions welcome

overview: (some) genAI issues

- instructors: learning feedback (pedagogical suggestions)

spontaneous questions welcome

overview: (some) genAI issues

- students: impacts on knowledge / critical thinking skills

spontaneous questions welcome

overview: (some) genAI issues

- broader: **what if** genAI bots are better teachers?

spontaneous questions welcome

my AI-related experiences

my AI-related experiences

- As part of **Biology Concept Instrument** (BCI) project (2006)

my AI-related experiences

- Ed Svirsky built Ed's Tools; used by Kathy Garvin-Doxas to build misconception-based "distractors"

my AI-related experiences

- Tom Lundy built virtuallaboratories (in FLASH)

my AI-related experiences

- a hands-on, student-driven introduction to classic experiments and methods in molecular biology

my AI-related experiences

- Aware of Latent Semantic Analysis (LSA) Landauer, Kintsch & collaborators – surprised (shocked) by appearance of ChatGPT Nov 2022

What was I doing

Between the BCI and ChatGPT

What was I doing

Between the BCI and ChatGPT

- a little cell and molecular biology / a little teaching (introductory / developmental biology)

What was I doing

Between the BCI and ChatGPT

- some thinking about core concepts in biology

What was I doing

Between the BCI and ChatGPT

- some thinking about core concepts in biology

Transformative collaboration with **Melanie Cooper**
(Chemistry Education, Clemson then Michigan State)

What was I doing

Between the BCI and ChatGPT

Transformative collaboration with **Melanie Cooper**
(Chemistry Education, Clemson then Michigan State)

- rethinking general chemistry - **Chemistry, Life, the Universe & Everything** (CLUE), Organic CLUE & biology (biofundamentals) – all OER

What was I doing

Between the BCI and ChatGPT

Transformative collaboration with **Melanie Cooper**
(Chemistry Education, Clemson then Michigan State)

- Used (web-based) beSocratic formative assessments
  – written by Sam Bryfczynski, an amazing grad. student

What was I doing

Between the BCI and ChatGPT

Transformative collaboration with **Melanie Cooper**
(Chemistry Education, Clemson then Michigan State)

- **Result**: Improved student learning / grades / performance
  (ACS baseline)

Affordances made possible by LLMs

encouraged by CustomGPT and RAG bots

Affordances made possible by LLMs

encouraged by CustomGPT and RAG bots

- a simple example - make multiple choice questions useful

Affordances made possible by LLMs

encouraged by CustomGPT and RAG bots

- students explain why incorrect or distracting (irrelevant) choices are incorrect or irrelevant

Affordances made possible by LLMs

encouraged by CustomGPT and RAG bots

- Rapid objective analyses of questions asked & students' answers – using RAG ChatGPT 4-based bot

Affordances made possible by LLMs

encouraged by CustomGPT and RAG bots

- Provides teacher with timely feedback on student thinking

an example

Q: How is genetic drift like molecular diffusion?

A. Both are the result of directed movements.

B. Both involve passing through a barrier.

C. Both involve random events without regard to ultimate outcome.

D. They are not alike. Genetic drift is random; diffusion typically has a direction.)

an example

Q: How is genetic drift like molecular diffusion?

A. Both are the result of directed movements.

B. Both involve passing through a barrier.

C. Both involve random events without regard to ultimate outcome.

D. They are not alike. Genetic drift is random; diffusion typically has a direction.)

***Answer Choice D**: They are not alike. Genetic drift is random; diffusion typically has a direction.*

an example

Q: How is genetic drift like molecular diffusion?

A. Both are the result of directed movements.

B. Both involve passing through a barrier.

C. Both involve random events without regard to ultimate outcome.

D. They are not alike. Genetic drift is random; diffusion typically has a direction.)

***Answer Choice D**: They are not alike. Genetic drift is random; diffusion typically has a direction.*

***Misconception Identified**: Students struggle with the concept of "direction" in diffusion. While it's true that diffusion involves movement from higher to lower concentration areas, this is a result of random molecular motion, not a directed or purposeful process. Some students mistakenly attribute a non-random, directed quality to diffusion.*

***Percentage of Misconception**: Approximately 70% of the responses misinterpret the nature of direction in diffusion, confusing it with intentional or non-random movement.*

Suggestions for Instructional Improvement:

Klymkowsky & Cooper. 2024. The end of mulitple choice tests. arXiv

Suggestions for Instructional Improvement:

- ***Clarify the Concept of Randomness**:*
  - *Emphasize that both genetic drift and molecular diffusion are processes influenced by random events. Use visual aids or simulations to show how these processes occur without any directional intent.*

Klymkowsky & Cooper. 2024. The end of mulitple choice tests. arXiv

Suggestions for Instructional Improvement:

- ***Clarify the Concept of Randomness**:*
  - *Emphasize that both genetic drift and molecular diffusion are processes influenced by random events. Use visual aids or simulations to show how these processes occur without any directional intent.*

- ***Address Misconceptions about Barriers**:*
  - *Specifically teach the differences in how barriers function in biological contexts, distinguishing between physical barriers in diffusion and metaphorical or statistical "barriers" in genetic drift.*

Klymkowsky & Cooper. 2024. The end of mulitple choice tests. arXiv

Suggestions for Instructional Improvement:

- ***Clarify the Concept of Randomness**:*
  - *Emphasize that both genetic drift and molecular diffusion are processes influenced by random events. Use visual aids or simulations to show how these processes occur without any directional intent.*

- ***Address Misconceptions about Barriers**:*
  - *Specifically teach the differences in how barriers function in biological contexts, distinguishing between physical barriers in diffusion and metaphorical or statistical "barriers" in genetic drift.*

- ***Enhance Understanding of "Directed Movements"**:*
  - *Use examples and counterexamples to teach* ***what is meant by "directed" in scientific contexts, clarifying that it refers to physical influences like gradients rather than purposeful actions.***

Klymkowsky & Cooper. 2024. The end of mulitple choice tests. arXiv

Suggestions for Instructional Improvement:

- ***Clarify the Concept of Randomness**:*
  - *Emphasize that both genetic drift and molecular diffusion are processes influenced by random events. Use visual aids or simulations to show how these processes occur without any directional intent.*

- ***Address Misconceptions about Barriers**:*
  - *Specifically teach the differences in how barriers function in biological contexts, distinguishing between physical barriers in diffusion and metaphorical or statistical "barriers" in genetic drift.*

- ***Enhance Understanding of "Directed Movements"**:*
  - *Use examples and counterexamples to teach* ***what is meant by "directed" in scientific contexts, clarifying that it refers to physical influences like gradients rather than purposeful actions.***

- ***Overall,** while some students grasp the stochastic nature of these processes, there is a clear need for more targeted instruction to address prevalent misconceptions, especially concerning the concepts of direction and barriers.*

Klymkowsky & Cooper. 2024. The end of mulitple choice tests. arXiv

teaching teachers

EDUC/MCDB 4811/5811/6811: Teaching and learning biology

teaching teachers

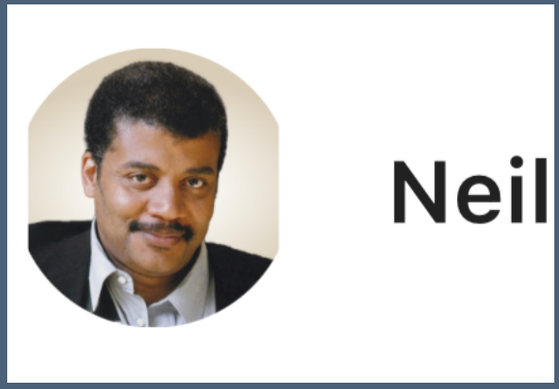
William Lindsay & Ian Her Many Horses
– CU Teach, School of Education

EDUC/MCDB 4811/5811/6811: Teaching and learning biology

teaching teachers

William Lindsay & Ian Her Many Horses
– CU Teach, School of Education

Introducing future teachers to student thinking and gen AI tools to shape it.



EDUC/MCDB 4811/5811/6811: Teaching and learning biology

teaching teachers

– Valerie Otero (School of Education)

teaching teachers

– Valerie Otero (School of Education)

- college level instructors rarely formally trained (i.e. through Schools of Education)

teaching teachers

– Valerie Otero (School of Education)

- “on the job” training often relies (primarily) on self-evaluation

teaching teachers

– Valerie Otero (School of Education)

- absence of timely, objective and actionable analyses of student learning / problems

teaching teachers with data

provide data-based feedback based on student responses

teaching teachers with data

provide data-based feedback based on student responses

- flag poorly understood or misapplied ideas & observations

teaching teachers with data

provide data-based feedback based on student responses

- Help designing targeted formative assessments

teaching teachers with data

provide data-based feedback based on student responses

- can include analysis of questions asked

teaching teachers with data

provide data-based feedback based on student responses

- **Q**: are course learning goals reasonable and relevant

teaching teachers with data

provide data-based feedback based on student responses

- **Q**: are goals being met by students?

teaching teachers with data

provide data-based feedback based on student responses

- suggest multi-semester “instructor boot camp”
  course design/delivery – evaluate (needed) modifications

Evaluating course goals

Evaluating course goals

- Flag courses with over-ambitious (i.e. unrealistic) learning objectives

Evaluating course goals

- e.g. recent paper on “endorsed learning objectives” with over 160 learning objectives

Evaluating course goals

- e.g. recent paper on “endorsed learning objectives” with over 160 learning objectives

seemed like a lot to me

https://chatgpt.com/share/6802b6c5-b280-8005-b8a4-e12f9e441fce
Q: *Acting as a socratic skeptic, identify the weaknesses of "endorsed learning outcomes" paper in terms of effective teaching and student learning.*

https://chatgpt.com/share/6802b6c5-b280-8005-b8a4-e12f9e441fce
Q: *Acting as a socratic skeptic, identify the weaknesses of "endorsed learning outcomes" paper in terms of effective teaching and student learning.*

GPT's Response: Acting as a Socratic skeptic … here are key "weaknesses and limitations"… in the context of "effective teaching" and "student learning"

https://chatgpt.com/share/6802b6c5-b280-8005-b8a4-e12f9e441fce
Q: *Acting as a socratic skeptic, identify the weaknesses of "endorsed learning outcomes" paper in terms of effective teaching and student learning.*

GPT's Response: Acting as a Socratic skeptic … here are key "weaknesses and limitations"… in the context of "effective teaching" and "student learning"

- 1. False Precision of Community Consensus
- 2. Conflation of Learning Objectives with Real Understanding
- 3. Neglect of How Scientific Knowledge Emerges
- 4. Systematic Devaluation of Higher-Order Thinking
- 5. Uncritical Adoption of Textbook-Derived Structures

https://chatgpt.com/share/6802b6c5-b280-8005-b8a4-e12f9e441fce
Q: *Acting as a socratic skeptic, identify the weaknesses of "endorsed learning outcomes" paper in terms of effective teaching and student learning.*

GPT's Response: Acting as a Socratic skeptic … here are key "weaknesses and limitations"… in the context of "effective teaching" and "student learning"

- 1. False Precision of Community Consensus
- 2. Conflation of Learning Objectives with Real Understanding
- 3. Neglect of How Scientific Knowledge Emerges
- 4. Systematic Devaluation of Higher-Order Thinking
- 5. Uncritical Adoption of Textbook-Derived Structures
- 6. Over-reliance on Alignment as a Pedagogical Panacea
- 7. Limited Epistemological Sophistication
- 8. Potential for Pseudo-Standardization
- 9. Failure to Account for Time Constraints and Cognitive Load
- 10. Minimal Attention to Student Voice or Learning Difficulties

instructor evaluation

instructor evaluation

- Established short-comings of student course/instructor evaluations (bias, manipulation, low completion rates)

instructor evaluation

- genAI could help evaluate whether a course and its instructor are achieving its purported learning goals

Why does this matter?

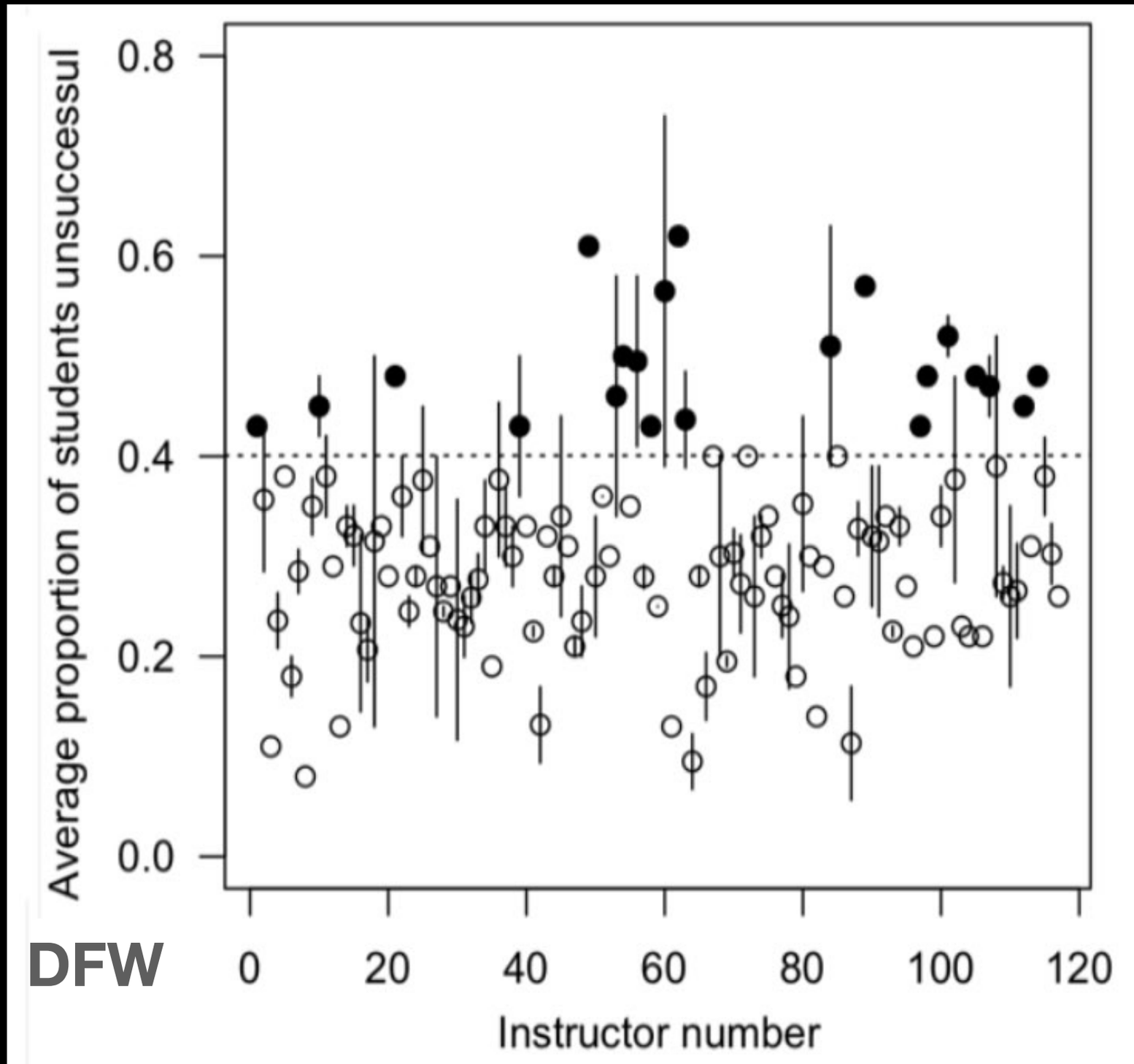
Why does this matter?

- Courses (on presumably the same material) can vary dramatically

Why does this matter?

- in terms of grade distribution and presumably what students have learned

Result: Students can face an "educational minefield"



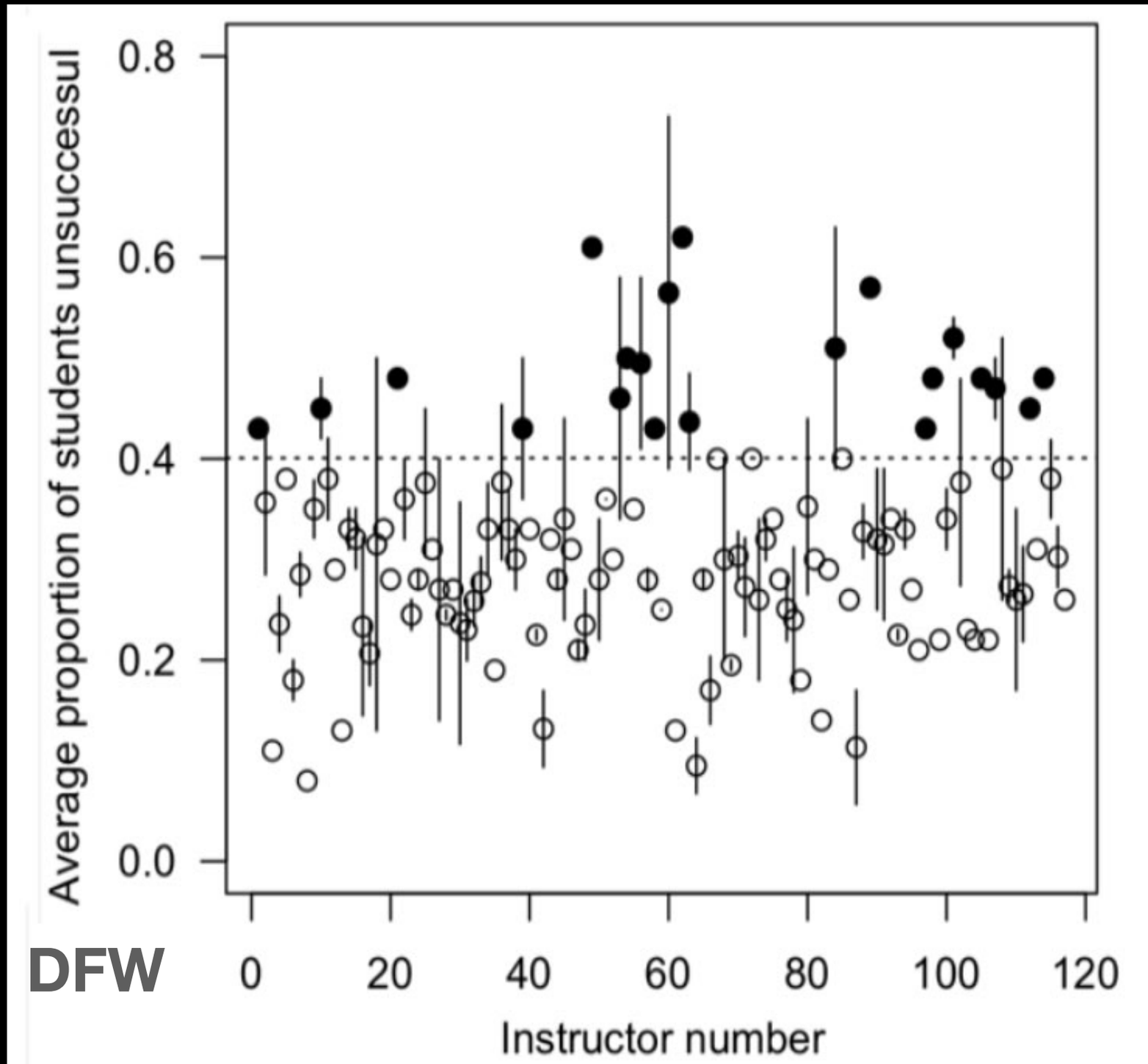
Each point is the average for an instructor of the SAME CLASS and the error bars are 1 standard error.
Points without error bars are individuals who taught the class only once.

Horizontal line is the upper 95% confidence interval of the mean (meaning the black points are, more or less, outliers).

Data analysis: Andy Martin EBIO
(+ CU Institutional Research)

Result: Students can face an "educational minefield"

- Impacts educational success and costs



Each point is the average for an instructor of the SAME CLASS and the error bars are 1 standard error.
Points without error bars are individuals who taught the class only once.

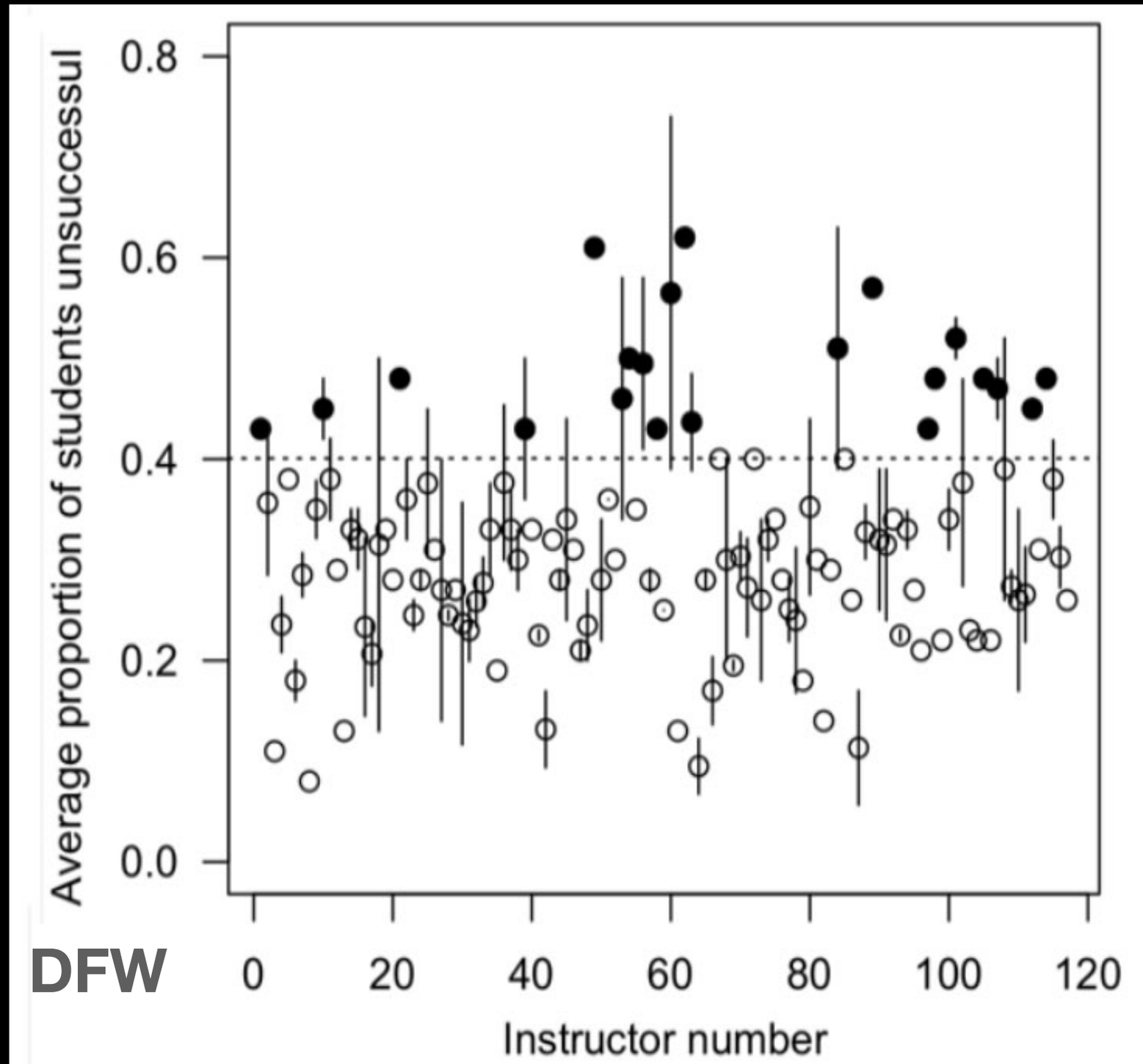
Horizontal line is the upper 95% confidence interval of the mean (meaning the black points are, more or less, outliers).

Data analysis: Andy Martin EBIO
(+ CU Institutional Research)

Result: Students can face an "educational minefield"

- mental well-being

Data analysis: Andy Martin EBIO
(+ CU Institutional Research)



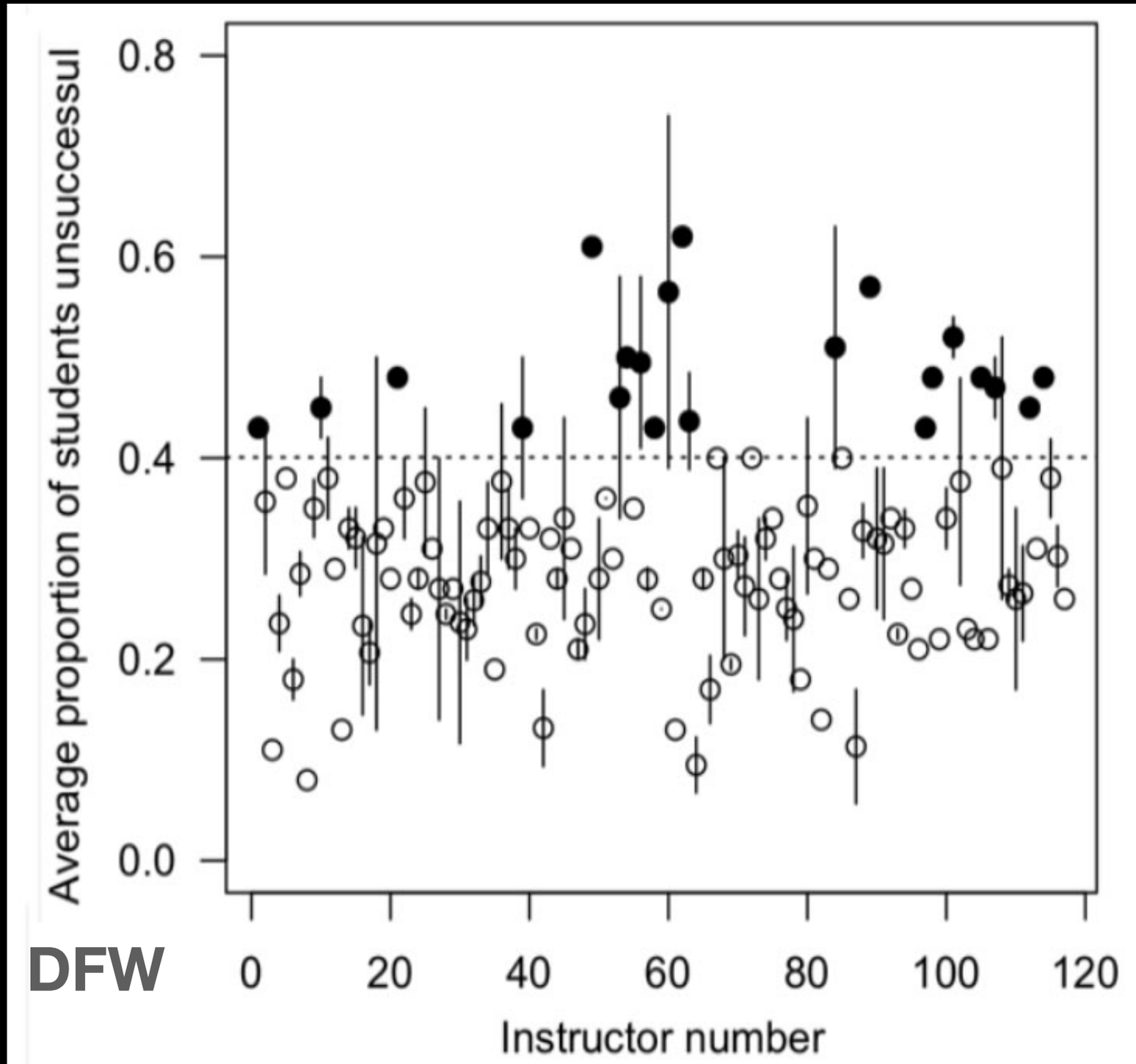
Each point is the average for an instructor of the SAME CLASS and the error bars are 1 standard error.
Points without error bars are individuals who taught the class only once.

Horizontal line is the upper 95% confidence interval of the mean (meaning the black points are, more or less, outliers).

Result: Students can face an "educational minefield"

- may reinforce unearned privileges and undeserved handicaps

Data analysis: Andy Martin EBIO
(+ CU Institutional Research)



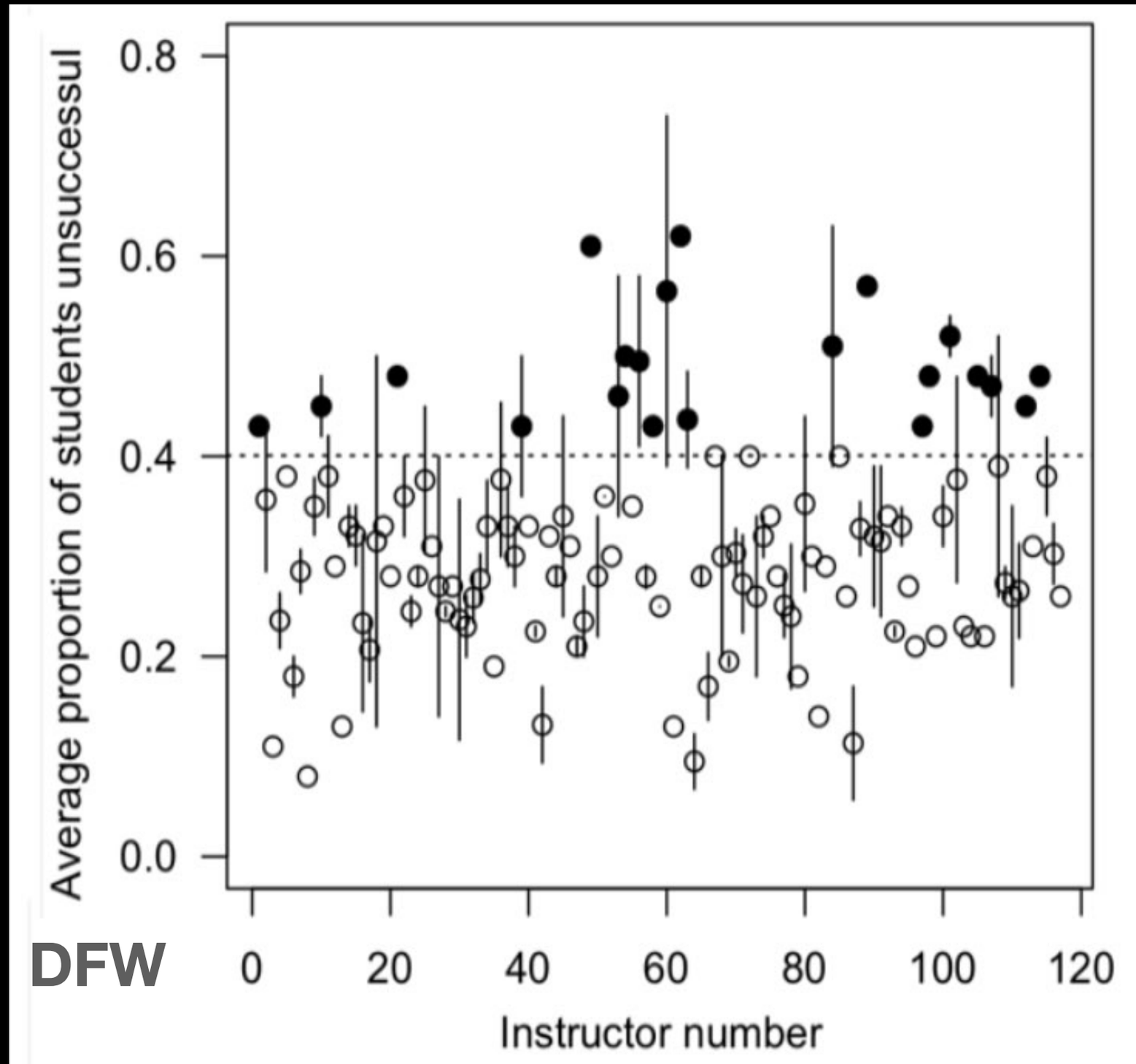
Each point is the average for an instructor of the SAME CLASS and the error bars are 1 standard error.
Points without error bars are individuals who taught the class only once.

Horizontal line is the upper 95% confidence interval of the mean (meaning the black points are, more or less, outliers).

Result: Students can face an "educational minefield"

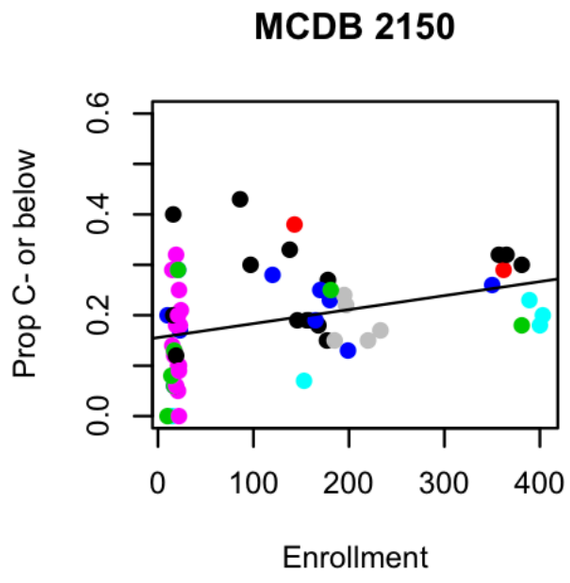
- Less that 50% graduate in 4 years, only ~60% in 6 years.

Data analysis: Andy Martin EBIO
(+ CU Institutional Research)



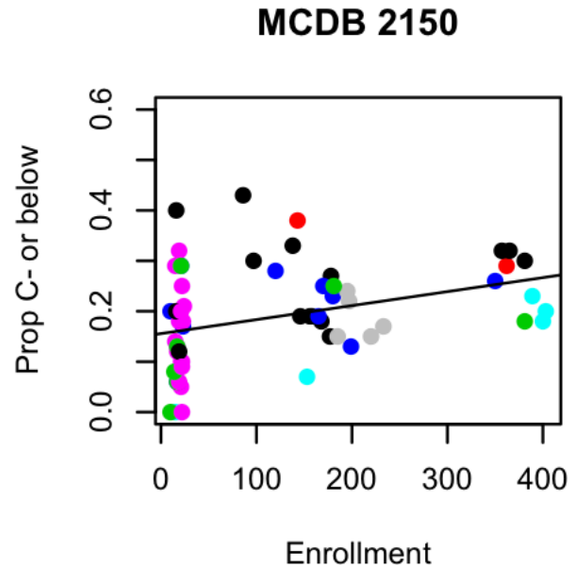
Each point is the average for an instructor of the SAME CLASS and the error bars are 1 standard error.
Points without error bars are individuals who taught the class only once.

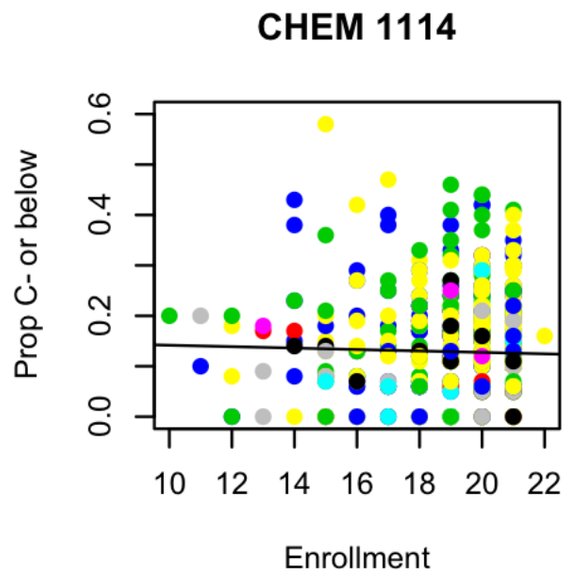
Horizontal line is the upper 95% confidence interval of the mean (meaning the black points are, more or less, outliers).



Each point is a separate section.
Colors are different instructors.

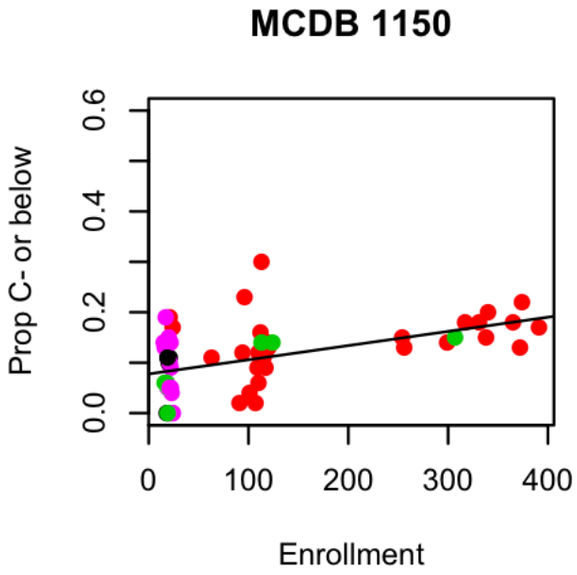
from Andrew Martin -EBIO

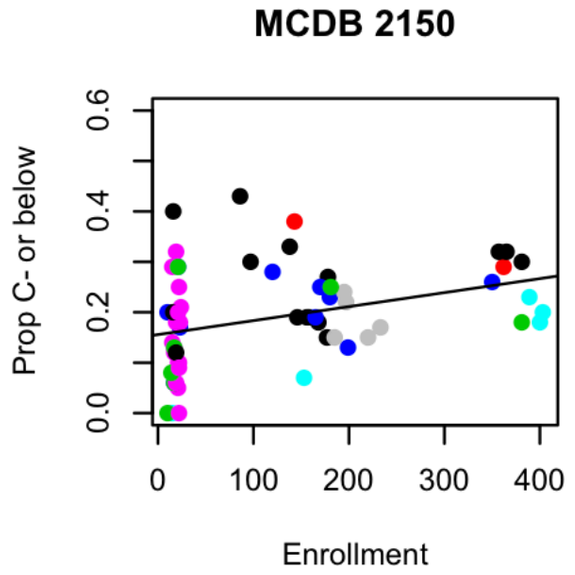


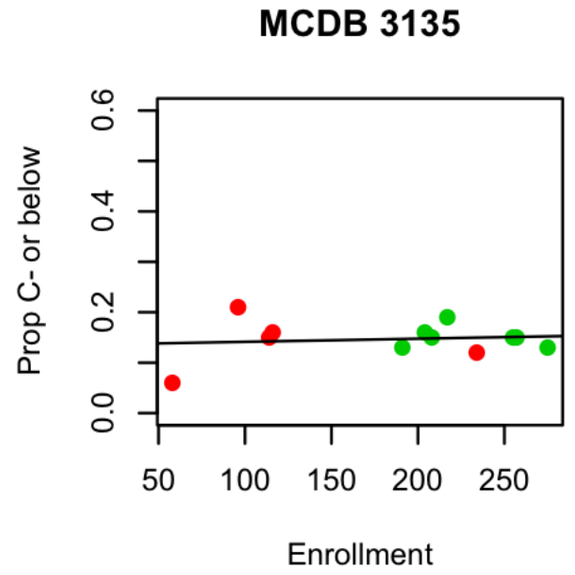


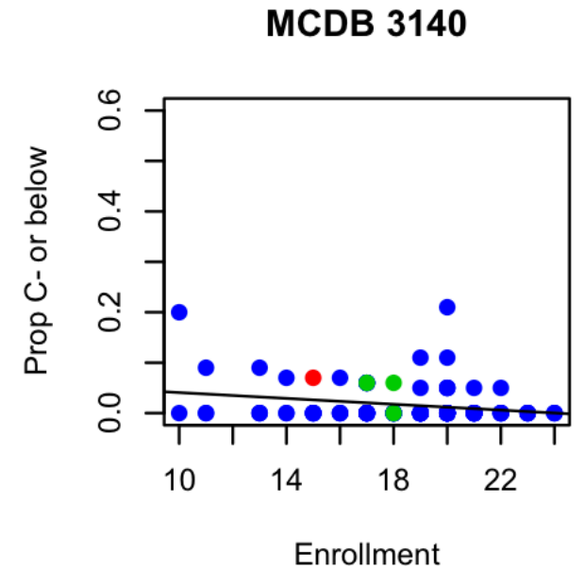
Each point is a separate section.
Colors are different instructors.

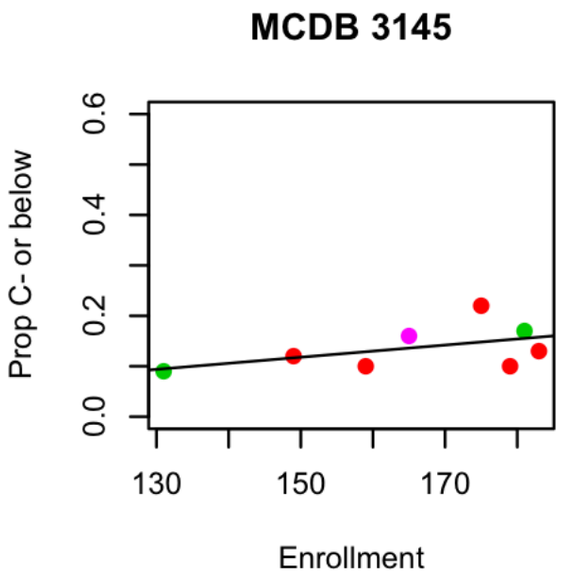
from Andrew Martin -EBIO

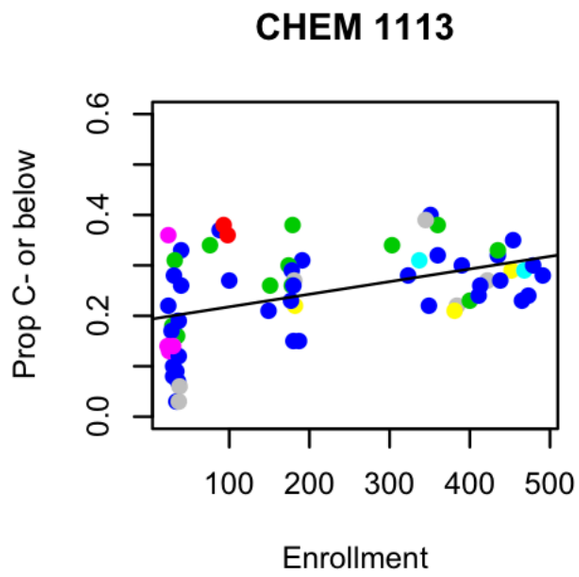


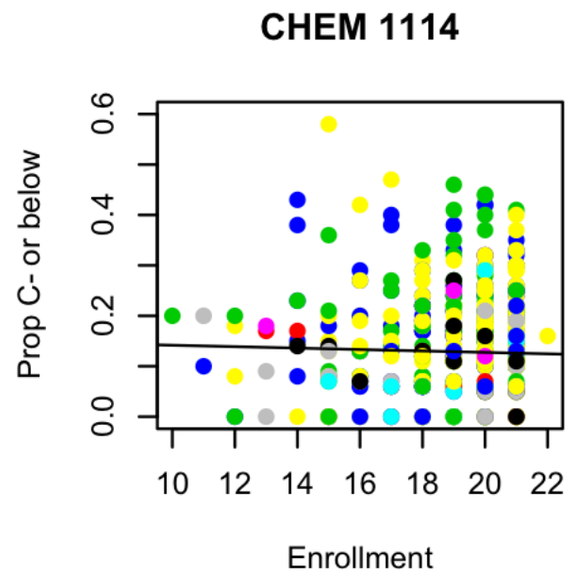


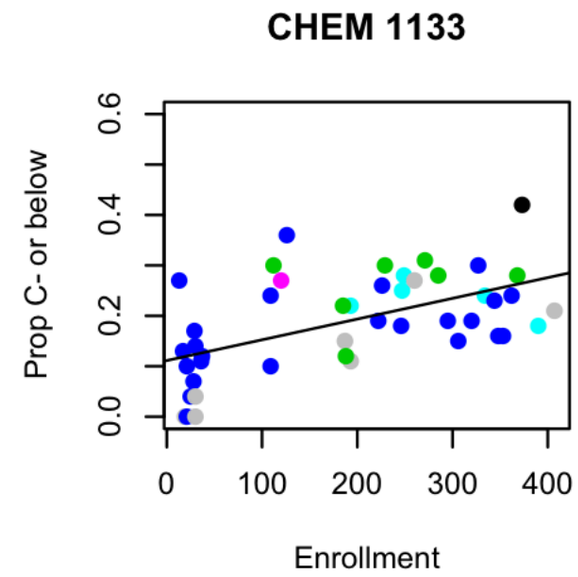


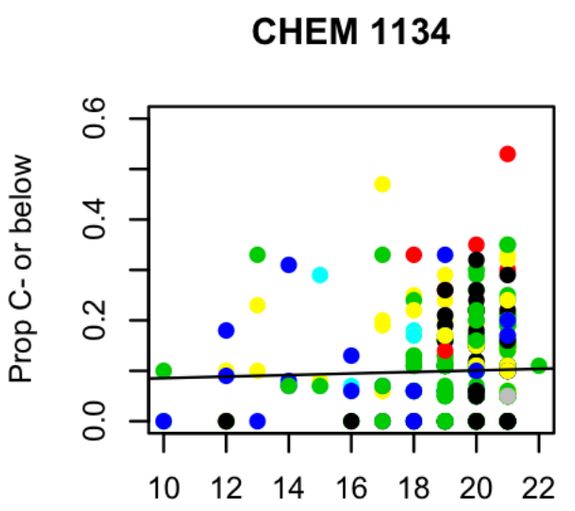


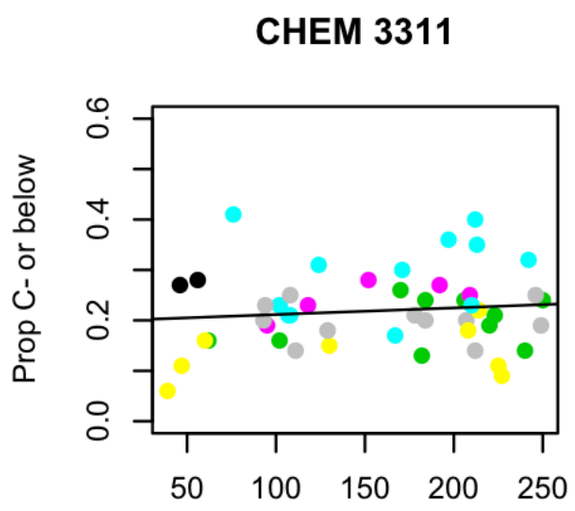












Each point is a separate section.
Colors are different instructors.

from Andrew Martin -EBIO

Why does it matter?

Why does it matter?

- evidence for differential learning efficacy is rare ...

Why does it matter?

- **n.b.** few controlled comparisons (such as carried out by Melanie Cooper and colleagues for CLUE & OCLUE)

Why does it matter?

- **n.b.** few controlled comparisons (such as carried out by Melanie Cooper and colleagues for CLUE & OCLUE)

Matz, R. L., Fata-Hartley, C. L., Posey, L. A., Laverty, J. T., Underwood, S. M., Carmel, J. H., ... & Cooper, M. M. (2018). Evaluating the extent of a large-scale transformation in gateway science courses. *Science advances*, *4*(10), eaau0554.

Affordances: evaluating course & curricular effectiveness

Affordances: evaluating course & curricular effectiveness

- Are courses meant to **teach or sort** student?

Affordances: evaluating course & curricular effectiveness

- Substack Essay: "First do no harm": Medical School Admissions Requirements and Educational Malpractice

UNDERGRADUATE SCIENCE

Weed-Out Courses Hamper Diversity

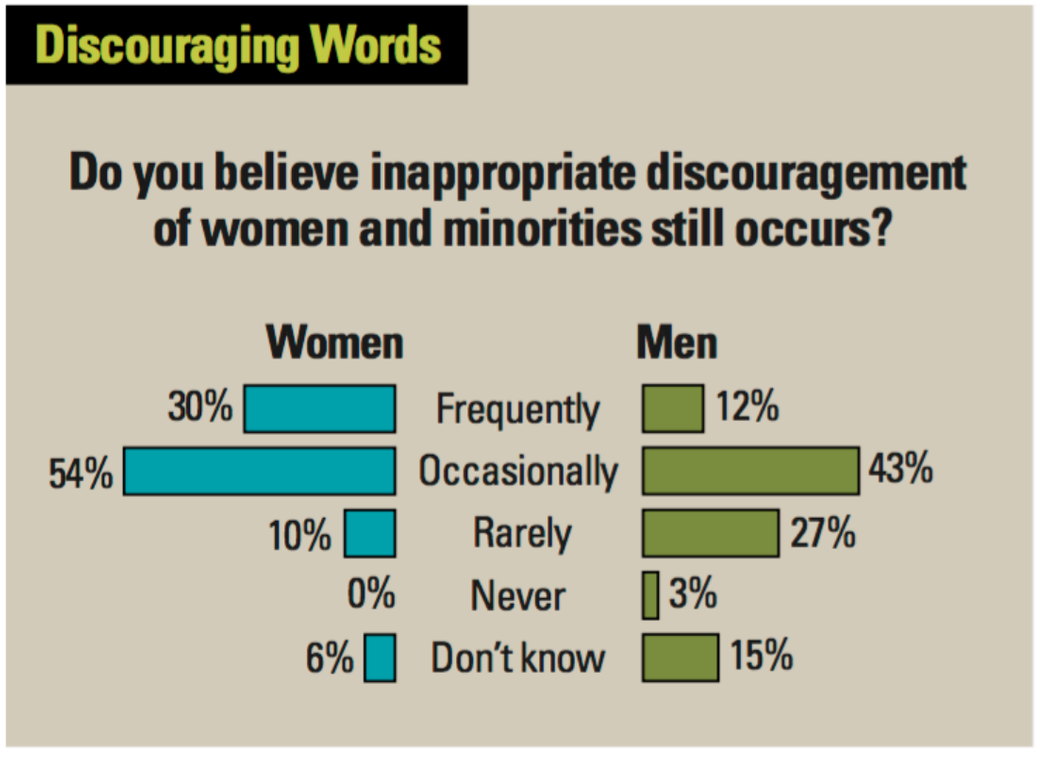
The time-honored practice of using introductory courses to weed out students seeking degrees in science and engineering hinders efforts to attract more women and minorities into those fields, say the chairs of science departments at U.S. universities. But the professors see no need to change their approach to teaching.

That contradiction appears in a survey by the Bayer Foundation, the 15th in its annual series on science education. More than 400 chairs from the top-200 research universities and from minority-serving institutions responded to a series of questions on their attitudes toward underrepresented minorities (African-Americans, Hispanics, and Native Americans) and women. It's a follow-up to last year's survey asking those students about the obstacles they face in pursuing STEM (science, technology, engineering, and mathematics) degrees.

The new survey (bayerus.com) found that 84% of the science chairs say broadening participation in STEM fields is important to their

courses are taught. There's a growing literature indicating that student collaborations, interactive technology, and hands-on learning are more effective in reaching students than are the traditional lecture and by-the-numbers lab sessions. But Hrabowski says many faculty members don't have the time, resources, or institutional support to try anything new.

Some faculty members also assume erroneously that they must lower the bar in order to have more students succeed. Their con-



Two perspectives. Most women chairs reported that women and minority students face roadblocks on their way to a STEM degree.

Mervis 2011.
Science **334**: 1333.

I asked chatGPT: Are there follow ups to Mervis 2011. Science 334: 1333.

I asked chatGPT: Are there follow ups to Mervis 2011. Science 334: 1333.

Happy to see

Curricular Design and Student Engagement: Michael Klymkowsky's 2023 article in *Frontiers in Genetics* discusses the challenges of designing effective biology curricula. He emphasizes the importance of aligning course content with students' backgrounds and experiences to enhance engagement and understanding. This approach aligns with efforts to move away from traditional "weed-out" courses towards more supportive learning environments .

Affordances:
Socratic, patient personalized tutors

Affordances:
Socratic, patient personalized tutors

- students have (24/7) access with knowledgable, patient socratic tutors who ...

Affordances:
Socratic, patient personalized tutors

- engage in formative activities (practice for summative assessment)

Affordances:
Socratic, patient personalized tutors

- engage in <u>formative</u> activities (practice for secure summative assessment)

Affordances:
Socratic, patient personalized tutors

secure
^

- encourage self-reflection through conversation

Affordances:
Socratic, patient personalized tutors

secure
^

- interactions that reflect where the student is at

Affordances:
Socratic, patient personalized tutors

secure
^

- lets the student know when they appear to “get it”

Affordances:
Socratic, patient personalized tutors

secure
^

- **Q:** Where do people fit into this process?

Threatening prospects: What if?

Threatening prospects: What if?

- socratic AI tutors are better teachers / teaching assistants

Threatening prospects: What if?

- impacts on grad. student support / educational jobs

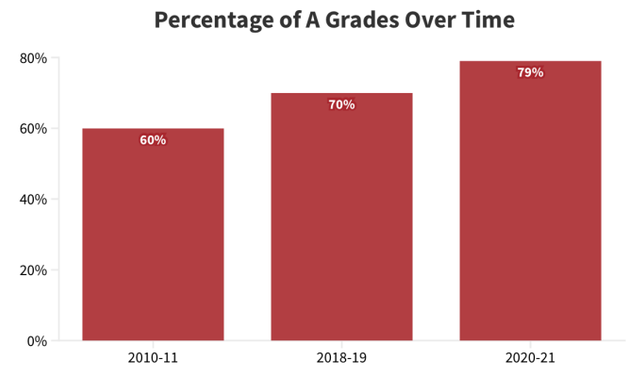
Threatening prospects: What if?

- will courses and degree programs: judged on disciplinary proficiency, rather than institutional status?

Threatening prospects: What if?

- will courses and degree programs: judged on disciplinary proficiency, rather than institutional status?

Harvard Report Shows 79% A-Range Grades Awarded in 2020-21, Sparking Faculty Discussion

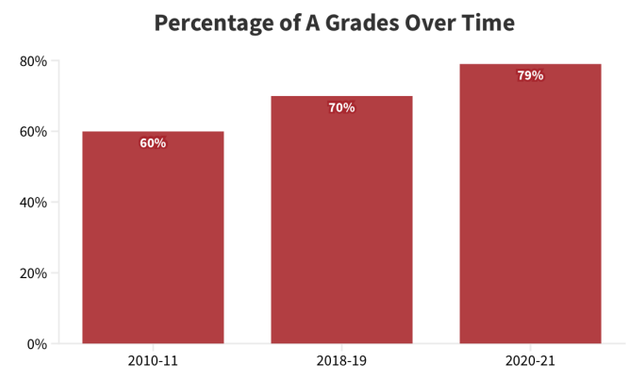


Source: Report on Grading at Harvard College • By Elias J. Schisgall

Threatening prospects: What if?

- so much to ponder and explore …

Harvard Report Shows 79% A-Range Grades Awarded in 2020-21, Sparking Faculty Discussion

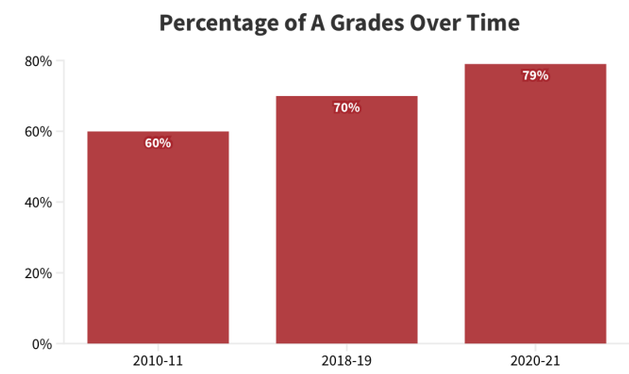


Source: Report on Grading at Harvard College • By Elias J. Schisgall

Threatening prospects: What if?

- so much to ponder and explore …

Harvard Report Shows 79% A-Range Grades Awarded in 2020-21, Sparking Faculty Discussion



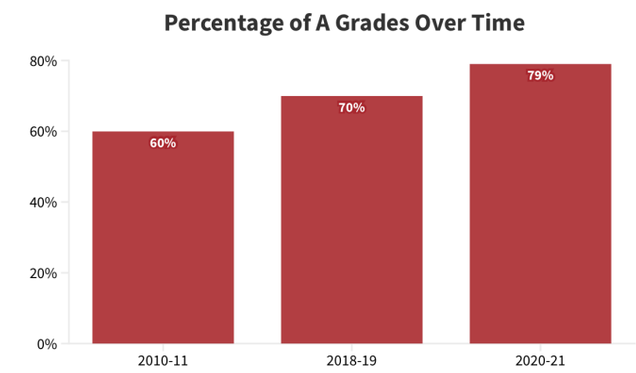
Montaigne concludes, like Socrates, that ignorance aware of itself is the only true knowledge - Roger Shattuck

Threatening prospects: What if?

- so much to ponder and explore …

comments / questions?

Harvard Report Shows 79% A-Range Grades Awarded in 2020-21, Sparking Faculty Discussion



Montaigne concludes, like Socrates, that ignorance aware of itself is the only true knowledge - Roger Shattuck

Papers cited:

Bryfczynski, S., Pargas, R. P., Cooper, M. M., & Klymkowsky, M. W. (2012). BeSocratic: Graphically-assessing student knowledge. In *Proceedings of the IADIS international conference on mobile learning* (pp. 3-10).

Cooper, M., & Klymkowsky, M. (2013). Chemistry, life, the universe, and everything: A new approach to general chemistry, and a model for curriculum reform. *Journal of Chemical Education*, *90*(9), 1116-1122.

Cooper, M. M., & Klymkowsky, M. W. (2022). Aligning Assessment Goals with the Current and Future Technologies Needed to Achieve Them. In *Technologies in Biomedical and Life Sciences Education: Approaches and Evidence of Efficacy for Learning* (pp. 241-257). Cham: Springer International Publishing.

Cooper, M. M., & Klymkowsky, M. W. (2024). Let us not squander the affordances of LLMS for the sake of expedience: using retrieval augmented generative AI chatbots to support and evaluate student reasoning. *Journal of Chemical Education*, *101*(11), 4847-4856.

Klymkowsky, M. W., & Garvin-Doxas, K. (2008). Recognizing student misconceptions through Ed's Tools and the Biology Concept Inventory. *PLoS Biology*, *6*(1), e3.

Klymkowsky, M. W., & Cooper, M. M. (2012). Now for the hard part: The path to coherent curricular design. *Biochemistry and Molecular Biology Education*, *40*(4), 271.

Klymkowsky, M. W., Rentsch, J. D., Begovic, E., & Cooper, M. M. (2016). The design and transformation of Biofundamentals: A nonsurvey introductory evolutionary and molecular biology course. *CBE—Life Sciences Education*, *15*(4), ar70.

Klymkowsky, M., & Cooper, M. M. (2024). The end of multiple choice tests: using AI to enhance assessment. *arXiv preprint arXiv:2406.07481*
.
Matz, R. L., Fata-Hartley, C. L., Posey, L. A., Laverty, J. T., Underwood, S. M., Carmel, J. H., ... & Cooper, M. M. (2018). Evaluating the extent of a large-scale transformation in gateway science courses. *Science advances*, *4*(10), eaau0554.