

A FACULTY WEBINAR

"Show Your Work" — Then AI Called My Bluff

Rethinking what our assessments actually measure

Dr. Rachel Tatro-Duarte, Porterville College

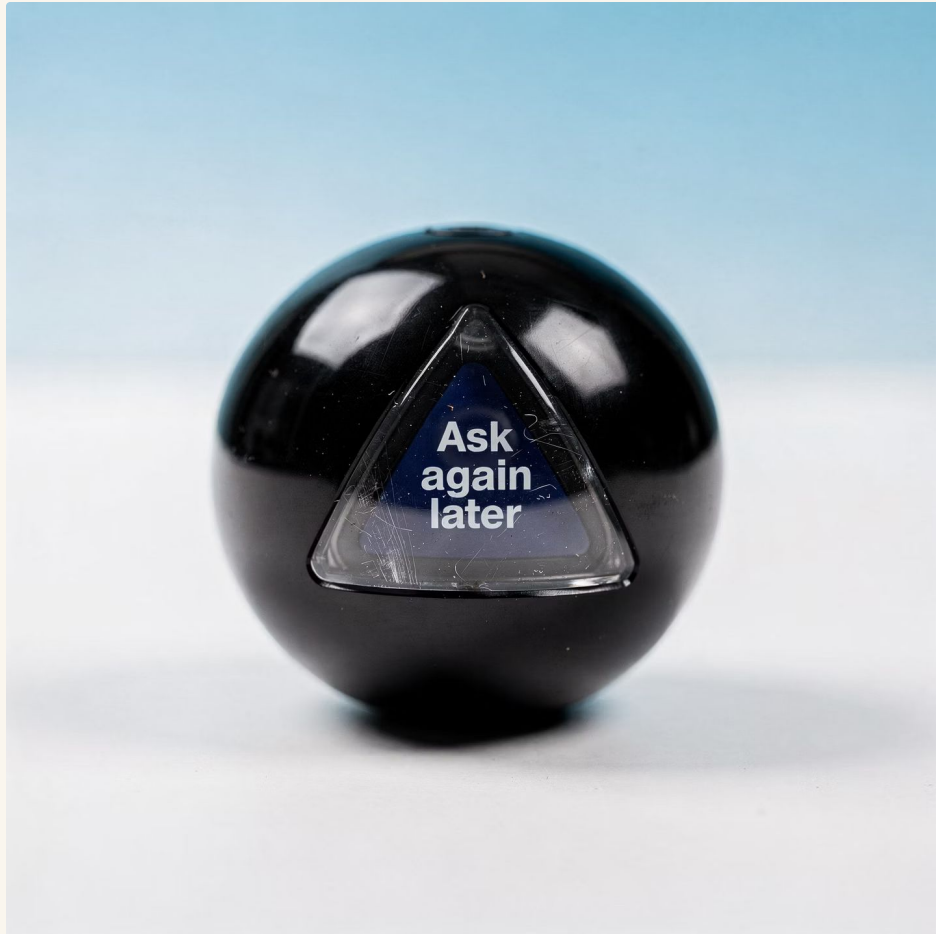
California Community Colleges Chancellor's Office AI Fellow

Show your work. Write the paper.

Now AI could show the work too.

WHERE I STARTED

The Magic Assessment Ball:



Did learning occur?

Shake it, turn it over, and read the little window.

Where did our assessment habits
come from?

A century of evidence: performance is not learning

Nicholas Soderstrom and Robert Bjork, cognitive psychologists at UCLA, reviewed research on learning reaching back nearly a century.

- **Learning** is lasting change. **Performance** is what we see during instruction, and it is often a poor sign of learning.

Four Key Findings

- Learning can happen with no visible change, and strong performance can happen with no real learning happening
- What appears as the scattered, often resulted in learning.
- It was the act of remembering, not the score, that builds learning.
- When something feels easy, we assume we have learned it.

Soderstrom, Nicholas C., and Robert A. Bjork. "Learning versus Performance: An Integrative Review." *Perspectives on Psychological Science*, vol. 10, no. 2, 2015, pp. 176–99, <https://doi.org/10.1177/1745691615569000>.

The Tolman & Honzik Maze Study (1930)

1. Latent Learning – learning that happens with no visible performance

Learning can occur without any obvious reinforcement or noticeable behavior change. Soderstrom and Bjork illustrate this with a bus ride example: A person in a new city takes the bus daily, absorbing the route through observation alone. The learning is invisible – until the day they have to drive themselves to work.

Three groups of rats. One complex maze. 17 days. The results changed how we understand learning.

Group 1: No Reward

Never rewarded. Simply removed from the maze each time.

Result: Most errors throughout all 17 days.

Group 2: Always Rewarded

Received food every time they reached the goal box.

Result: Fewest errors. Consistent improvement from the start.

Group 3: Rewarded from Day 12

No reward for the first 11 days. The rats wandered, and bumped into walls. Then rewarded from day 12 onward.

Result: Started like Group 1 – then caught up to Group 2 almost immediately.

"The learning was happening all along. It just had no reason to show itself – until it did."

Spacing teaching out and providing various activities between learning made performance look worse – but created real learning

The "Distribution of Practice" section in Soderstrom and Bjork (2015) makes one central argument.

Distributed Practice

- Spreading practice out over time, or mixing different skills and activities together
- Performance looks **worse** during training (swimming example)
 - Lead to **stronger and longer-lasting learning**
- Not something that can be captured during a short-term evaluation

Blocked Practice-Crammed teaching

- Practicing the same thing repeatedly—drill learning
- Performance looks **good in the moment**
- Great for capturing performance especially on test or quiz
 - Learning **fades faster-does not last**
 - Semester terms: concepts to be mastered in a week or two leads to block practice.

Desirable Difficulties

Ways of practicing that make learning feel harder and look worse in the moment — but help it last longer.

What this looks like:

- Frustratingly hard
- Produce more mistakes
- Bad results—constant wrong answers
- Looks like failing and confusion
- May not look like learning when we need it to

Why aren't these used as part of the assessment?

- The struggle looks like failure — to students and teachers alike
- Short-term performance dips, so they feel like the assignment isn't working
- Appears as though learning is not happening
- Not great data for our assessment

These Desirable difficulties are exactly what makes the learning stick—though we give very little room for this as our final assessment

Six things the research gave me

We confuse short-term performance with long-term learning.

Learning can happen without visible performance, and often settles after we measure it.

Strong flawless performance can happen without resulting in learning

Grading punishes the struggle that builds learning.

Testing helps when it is small, low-stakes, and for the learner vs. High-stakes tests made to serve outside viewers.

There is little evidence that outcomes-based assessments improve learning.

Plato, *Phaedrus*, c. 370 BCE

- In Plato's *Phaedrus*, Socrates tells a story about the invention of writing.
- The Egyptian god Theuth offers writing to King Thamus as a remedy for forgetting.
- Thamus refuses. Writing will weaken memory and give people the show of wisdom without the reality.
- Socrates adds that writing is like a painting: "if you ask them a question they preserve a solemn silence."
- What he trusted instead was living conversation, knowledge that can answer back and defend itself.

We took the technology Plato distrusted, and made it our test of knowledge.

Plato. *Phaedrus*. Translated by Benjamin Jowett, Project Gutenberg, 1 Feb. 1999, www.gutenberg.org/ebooks/1636.

Fast forward to the viva voce: examination by the living voice

- For centuries, students were not graded. They were examined in a **disputation**, a formal debate held aloud.
- One student, the respondent, put forward a position and defended it. Opponents argued against it.
- A moderator presided, pressed where the argument was thin, and could keep it going until satisfied.
- The judgment formed as the argument unfolded, in the room, in front of witnesses.
- It could not be faked for long. A position you did not hold collapsed under questioning.

The written exam grew out of mathematics, not out of volume

- By the early 1700s the disputation was a formality, and Newton's mathematics had spread through Cambridge.
- Around 1725, a voluntary exam began to rank the better students. At first it was oral.
- Then came written answers, and by about 1790, printed papers: the Mathematical Tripos.
- Students were ranked in order of merit. The top were still called wranglers, from the old arguing.
- Writing proved a more exact test for mathematics. Enrollment fell, so it was not about volume.

"A History of Mathematics in Cambridge." *Faculty of Mathematics*, University of Cambridge, www.maths.cam.ac.uk/about/history. Accessed 24 Sept. 2026.

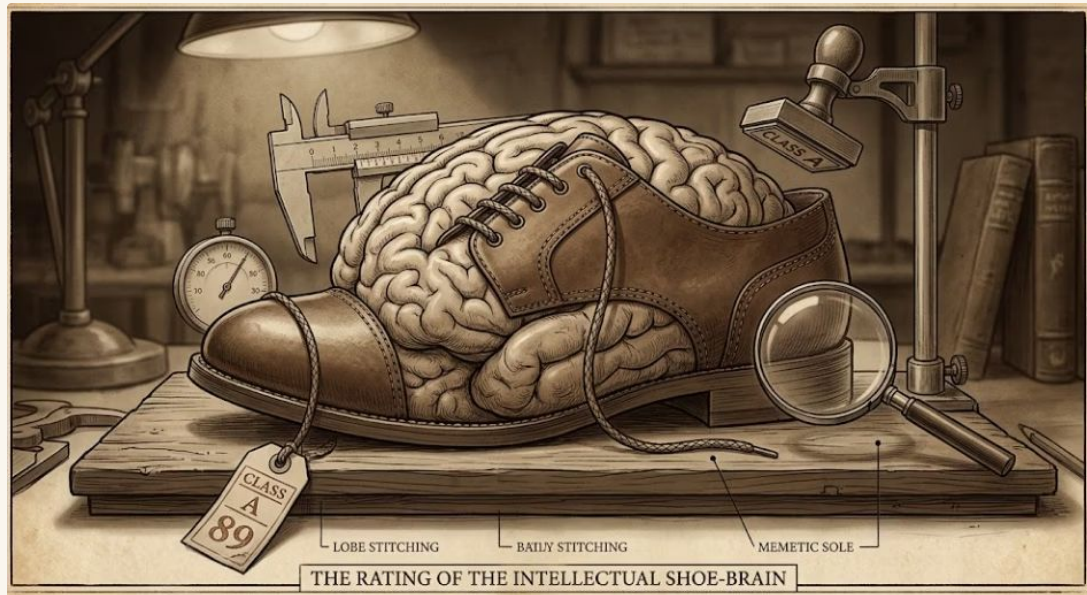
"The University of Cambridge: The Age of Newton and Bentley (1660–1800)." *A History of the County of Cambridge and the Isle of Ely*, edited by J. P. C. Roach, vol. 3, Victoria County History, 1959, pp. 210–35. *British History Online*, www.british-history.ac.uk/vch/cambs/vol3/pp210-235.

Cowan, David. "Pittite Triumph and Whig Failure in the Cambridge University Constituency, 1780–96." *Parliamentary History*, vol. 43, no. 1, 2024, pp. 91–111, <https://doi.org/10.1111/1750-0206.12724>.

1792

Then there was the beginning of the mark

Neil Postman traces the first grading of student papers to Cambridge in 1792, at the suggestion of a tutor named William Farish. Postman adds that almost nothing is known about Farish. Historians dispute whether he was first. Yale was ranking students by 1785.



Educational folklore says Farris borrowed a shoe factory's quality scale. The story may not be true. But the allegory is: a mind, graded like a product.

Postman, Neil. *Technopoly: The Surrender of Culture to Technology*. Knopf, 1992.

Durm, Mark W. "An A Is Not an A Is Not an A: A History of Grading." *The Educational Forum*, vol. 57, no. 3, 1993, pp. 294–97.

The Beginning of the Distrust: Horace Mann and the fight over who controlled the schools

- Secretary of the Massachusetts Board of Education from 1837 to 1848, and a leader of the common school movement.
- In 1843 he toured European schools, especially in Prussia, and held them up as better than Boston's.
- In 1844, thirty-one Boston schoolmasters published a 144-page attack on his report. They ran their schools with real independence.
- In 1845, Mann and his ally Samuel Gridley Howe answered back with a surprise written exam.
 - This becomes a full on pamphlet war and grab for power—who runs the educational system—educators or administration
- The doesn't create the exam—instead he uses it as a move in a political fight over who controlled the schools. Mann wins.

Association of Masters of the Boston Public Schools. *Remarks on the Seventh Annual Report of the Hon. Horace Mann, Secretary of the Massachusetts Board of Education*. C. C. Little and J. Brown, 1844.

Reese, William J. *Testing Wars in the Public Schools: A Forgotten History*. Harvard UP, 2013.

Kaestle, Carl F. *Testing Policy in the United States: A Historical Perspective*. Educational Testing Service, 2013, www.ets.org/Media/Research/pdf/kaestle_testing_policy_us_historical_perspective.pdf.

HORACE MANN, BOSTON, 1845

Mann published the test results, claims that the written test made the classroom visible to everyone outside for general inspection.

"A transcript, a sort of Daguerreotype likeness, as it were, of the state and condition of the pupils' minds, is taken and carried away, for general inspection. Instead of being confined to committees and visitors, it is open to all."

"History of Standardized Testing." *College of Education*, Lehigh University, 2013, ed.lehigh.edu/news/2013/history-standardized-testing.

Reese, William J. *Testing Wars in the Public Schools: A Forgotten History*. Harvard UP, 2013.

Kaestle, Carl F. *Testing Policy in the United States: A Historical Perspective*. Educational Testing Service, 2013, www.ets.org/Media/Research/pdf/kaestle_testing_policy_us_historical_perspective.pdf.

The first standardized test in America was not about the student-it was used as a weapon to take control.

- Mann and Samuel Gridley Howe planned the exam in secret, to show that Boston's schools were not as good as the city believed.
- Examiners tested 530 of the city's best pupils below high school. Most flunked.
- The results were published to shame the schoolmasters, fire teachers and take control.

We have that now

Our assessments still judge educators, not only students.

K-12 learned to teach to the test. Higher education is not that different

Higher education institutions use assessments **to measure student learning, improve teaching quality, and prove institutional accountability-maintain accreditation and funding**

Reese, William J. *Testing Wars in the Public Schools: A Forgotten History*. Harvard UP, 2013.

Reese, William J. "First Race to the Top." *Larry Cuban on School Reform and Classroom Practice*, 17 Sept. 2013, larrycuban.wordpress.com/2013/09/17/first-race-to-the-top-william-reese/.

"History of Standardized Testing." *College of Education*, Lehigh University, 2013, ed.lehigh.edu/news/2013/history-standardized-testing.

Mann's exam moved control out of the classroom and educators in the room lost the power to decide what learning was.

- Before, the schoolmaster decided what was taught and judged his own students.
- Reformers wanted test results used to decide which teachers were rehired.
- Superintendents used tests to centralize their own authority. Chicago's compared them to the gauges of industry.
- Graded classrooms, graded textbooks, and promotion tests set the sequence from outside the room, and publishers found a market.

The test became a tool of political control, and the educators in the room lost the power to decide what learning was.

Kaestle, Carl F. *Testing Policy in the United States: A Historical Perspective*. Educational Testing Service, 2013, www.ets.org/Media/Research/pdf/kaestle_testing_policy_us_historical_perspective.pdf.

Reese, William J. *Testing Wars in the Public Schools: A Forgotten History*. Harvard UP, 2013.

Specify the product, and the school will manufacture it

- He wanted schools run like a plant, at one hundred percent efficiency, and in 1918 applied that logic to curriculum. Rotating classes.
- If business named the product it wanted, he argued, schools could produce it.
- The student becomes a shapeable product. The school becomes the plant that makes the employee.

"Education is a shaping process as much as the manufacture of steel rails."

Once the student is a product, assessment becomes a matter of checking whether the product meets its specification, which is far easier than asking whether a person has grown.

Very little room is left for human creativity. And AI is extremely good at producing to specification.

Bobbitt, John Franklin. "The Elimination of Waste in Education." *The Elementary School Teacher*, vol. 12, no. 6, 1912, pp. 259–71.

Kaestle, Carl F. *Testing Policy in the United States: A Historical Perspective*. Educational Testing Service, 2013, www.ets.org/Media/Research/pdf/kaestle_testing_policy_us_historical_perspective.pdf.

Bobbitt, John Franklin. *Some General Principles of Management Applied to the Problems of City-School Systems*. The Twelfth Yearbook of the National Society for the Study of Education, Part I, University of Chicago Press, 1913, p. 10. *Internet Archive*, <https://archive.org/details/somegeneralprinc00bobbbrich/page/10/mode/2up>.

Tyler gave us learning outcomes

- He asked four questions: what should students learn, what experiences get them there, how are those organized, and how will we know they learned it?
- Start from the objective, and design backward from it.
- This is where our learning outcomes come from.
- It assumes the objective shows up in an assignment that we can inspect, and that the assignment proves the learning.

Outcomes turned learning into something that could be reported upward.

Scriven named formative and summative

- Formative evaluation happens during the process, to improve it.
- Summative evaluation happens at the end, final submission to judge it. No changes can be made.
 - The difference between is largely is whether there is still time to change something.

The assessment mostly uses summative: the finished product, judged at the end rather than formative which allows for room to fix and improve.

Scriven, Michael. "The Methodology of Evaluation." *Perspectives of Curriculum Evaluation*, edited by Ralph W. Tyler et al., Rand McNally, 1967, pp. 39–83.

AND LATER — JAMES NOTTINGHAM, 2017

Nottingham said learning runs through confusion

- The Learning Pit: you meet a real problem, get stuck, build your way out, and reflect on how.
- It draws on Socratic aporia, productive puzzlement
- Confusion is not failure. It is the precondition of deep understanding.

Our grading already punished the pit: students are graded on what they do not know yet, within one semester.

Nottingham, James. *The Learning Challenge: How to Guide Your Students through the Learning Pit to Achieve Deeper Understanding*. Corwin, 2017.

THE ASSESSMENT MOVEMENT, 1983 TO 1985

Born from outside pressure, and accountability won

1983

A Nation at Risk

1984

Involvement in Learning

1985

Integrity in the College Curriculum, and the first national assessment conference

National Commission on Excellence in Education. *A Nation at Risk: The Imperative for Educational Reform*. U.S. Department of Education, 1983.

Study Group on the Conditions of Excellence in American Higher Education. *Involvement in Learning: Realizing the Potential of American Higher Education*. National Institute of Education, 1984.

Association of American Colleges. *Integrity in the College Curriculum: A Report to the Academic Community*. Association of American Colleges, 1985.

The founders of the field of assessment could not show that it worked-Closing the Loop

6%

of 146 profiles of good practice
showed that learning improved

- Asked for examples of assessment improving learning, they found only a handful.
- Their own standard: students actually learn more, and programs actually change.
- Newer work agrees: no good evidence it improves learning (Ontario, 2015), and few programs check whether their changes worked (CUNY, 2025).
- So an outcome is an instruction about what we are asked to measure, not a description of learning.

Banta, Trudy W., and Charles Blaich. "Closing the Assessment Loop." *Change: The Magazine of Higher Learning*, vol. 43, no. 1, 2011, pp. 22–27, <https://doi.org/10.1080/00091383.2011.538642>.

Goff, Lori, et al. *Learning Outcomes Assessment: A Practitioner's Handbook*. Higher Education Quality Council of Ontario, 2015, heqco.ca/wp-content/uploads/2020/03/heqco.LOAhandbook_Eng_2015.pdf.

"Assessment Cycle in Higher Education: A Critical Appraisal." *Assessment at CUNY*, CUNY Academic Commons, Oct. 2025, assessatcuny.commons.gc.cuny.edu/2025/10/assessment-cycle-in-higher-education-a-critical-appraisal/.

Some saw the problem long before AI

Grant Wiggins, 1989

- Tests should ask students to do the real work of a discipline, not recall isolated facts.
- Called it authentic assessment, and questioned the essay and exam as measures of thinking.
- Later built *Understanding by Design* with Jay McTighe.

Lorrie Shepard, 2000

- Traced our testing to behaviorism, "scientific measurement," and the social efficiency movement.
- Assessment belongs inside learning, as feedback while there is still time.
- Warned in the 1980s that high-stakes tests would lead to drilling.

Wiggins, Grant. "A True Test: Toward More Authentic and Equitable Assessment." *Phi Delta Kappan*, vol. 70, no. 9, 1989, pp. 703–13.

Wiggins, Grant, and Jay McTighe. *Understanding by Design*. Association for Supervision and Curriculum Development, 1998.

Shepard, Lorrie A. "The Role of Assessment in a Learning Culture." *Educational Researcher*, vol. 29, no. 7, 2000, pp. 4–14, <https://doi.org/10.3102/0013189X029007004>.

Kaestle, Carl F. *Testing Policy in the United States: A Historical Perspective*. Educational Testing Service, 2013, www.ets.org/Media/Research/pdf/kaestle_testing_policy_us_historical_perspective.pdf.

Even the people who built it tried to turn it back

Carl Brigham

Built the SAT in 1926 from the Army intelligence test, after ranking races by intelligence in 1923.

He later recanted, and tried to slow the SAT's spread. But the machinery was already in motion.

Diane Ravitch

Helped lead the national push for standards and test-based accountability.

In 2010 she publicly reversed herself and argued it had failed.

Their reversals never traveled as far as their tests.

Kaestle, Carl F. *Testing Policy in the United States: A Historical Perspective*. Educational Testing Service, 2013, www.ets.org/Media/Research/pdf/kaestle_testing_policy_us_historical_perspective.pdf.

Ravitch, Diane. *The Death and Life of the Great American School System: How Testing and Choice Are Undermining Education*. Basic Books, 2010.

WHERE THAT LEAVES US

There is a lot of weight being put on AI for this assessment crisis, but it really did not start with AI

We are being asked to guard a measure that so far, has been proven to largely serve people outside the classroom, not the learning inside it

WHERE THAT LEAVES US

Perhaps less is more

- Every time a test was criticized, the answer was more testing.
- The one thing this history never tried is less testing, less assessment
- But less only works if we trust the people who teach, and the students who learn.

Essay Outline

I. Introduction

- A. Hook
- B. Background
- C. Thesis

II. Body Paragraph 1

- A. Topic Sentence
- B. Evidence

C. Analysis

III. Body Paragraph 2

- A. Topic Sentence
- B. Evidence
- C. Analysis

IV. Counterargument

- A. Opposing View
- B. Refutation

V. Conclusion

- A. Restate Thesis
- B. Summary
- C. Final Thought

"The unexamined life
is not worth living."

- Socrates

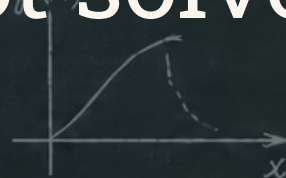
$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}$$

$$\frac{\partial^2 u}{\partial t^2} = c^2 \nabla^2 u$$

$$E = mc^2$$

"Education is not
the learning of facts,
but the training of
the mind to think."

- Einstein



repeatedly do. Excellence,
then, is not an act,
but a habit."

$$R_{ij}^n - \frac{1}{2} R g_{ij} + A g_{ij} = \frac{8\pi G}{c^4} T_{ij} \quad - \text{Aristotle}$$



```
def forward(x):  
    x = linear1(x)  
    x = relu(x)  
    x = linear2(x)  
    return x
```

```
class Net(nn.Module):  
    def __init__(self):  
        super().__init__()  
        self.layer = nn.Sequential(  
            nn.Linear(128, 64),  
            nn.ReLU(),  
            nn.Linear(64, 10)  
        )  
    def forward(self, x):  
        return self.layer(x)
```

```
import torch  
import torch.nn as nn  
import torch.nn.functional as F
```

```
loss = criterion(output, target)  
loss.backward()  
optimizer.step()
```

```
# learning never stops  
for epoch in range(epochs):  
    for batch in loader:  
        output = model(batch)  
        loss = criterion(output, target)  
        loss.backward()  
        optimizer.step()
```

So, I did not solve the AI Assessment problem

ἀπολογία assessment

التعليم والتربية

PART TWO

المعالم 學習 learning

What the research showed me: It's not AI, it's not our students, it's not our assignments.

التقييم

examination

التقييم

grading

WHAT THE HISTORY DID FOR ME

Realizing this helped me set aside an inherited, flawed tool and willingly lean into trust.

Trust

- Myself, and the judgment I bring to my classroom.
- My colleagues using AI, and trying to figure it out.
- My colleagues who have decided not to use AI in the classroom
- My students, who want to learn, who have been measured far more than they have been trusted, and who need our guidance.

WHERE I LANDED

Three small ways I apply my research in teaching:

Mixed activities-SLOs

The assessment work for the outsiders. The less exciting stuff.

The viva voce

Brought back in whatever small ways I can, especially online-with and without AI

Fostering autonomy in adult learners

I open the door, and they choose how to walk through it. Putting the ownership of learning back into the students' hands, allowing and trusting them to make decision about AI and their learning

For the SLOs, I use a two-step method: Examine and Surface

1. Examine

Your outcome may hold several learning goals in one sentence. Take one.

Example: Collect, **evaluate**, analyze, and **synthesize** primary and secondary sources in order to write a focused, well-organized 2,000-word academic argumentative research paper that demonstrates awareness of audience, purpose, and voice.

- The learning goal, in a word or two
- Evaluate-Students will evaluate sources
 - This is the learning

2. Surface

What is the best way to see this learning happen in a varied practice? (Soderstrom & Bjork)

I typically use the research paper to evaluate everything. Part of evaluation is already visible in the sources they chose to use in their essay.

Varied practice:

Create a list of sources they will use and sources they will reject—This can be submitted a viva voce where they submit the list along with an brief audio talk through telling me about they why.

The Mending Apparatus -Playlab Bots

What changes

- The essay as it is—stays the same.
- What it adds is a varied assignment: that mixes up the review of the sources.
- Isolating learning how to evaluate sources
- Learning I can't always see: The evaluation of articles they rejected.

THE VIVA VOCE, IN SMALL WAYS

Finding more ways for the living voice to enter the learning

THE VIVA VOCE, IN SMALL WAYS

Online

- Quizzes with an audio follow-up, where students explain one or more answer in their own voice.
- A short audio note with each essay submission, where students talk through what still confuses them, what they doubt, or what they are not sure they got right despite submitting the essay.

Face-to-face

- Randomized viva voce questions, so there is nothing to rehearse and the response comes out natural.
- Mini-low stress, fun and playful disputations in class with peers and myself as the moderator
 - groups, pairs, individual--offer extra credit

Instead of assessing only what students already know, give credit for showing what they do not know yet and how they work through it.

Grade the confusion note. The short audio note on what still confuses them becomes a graded part of the submission. You grade how specific it is, not whether the student is right. "I don't get it" earns little, and "I'm not sure whether this source counts as biased or just one-sided" earns full credit.

Explain the wrong answer. Students take something they got wrong and explain why it seemed right at the time. Then they find the exact point where the mistake happened. The grade goes to how well they understand their own thinking, not to whether they got it right the first time.

Earn back missed quiz questions. On the mastery quiz, a student can earn back a missed question by recording a short audio explanation of why they missed it and what they see now. The mistake opens a viva voce moment.

That is not a flaw in the answer. That is the answer.

Let them choose: Empowering adult learners

Mastery quiz

Timed, but retake it as many times as you want, for the score you want.—low stakes, learning for the student kinda quiz (Sonderstrom & Bjork)

Active learning path

A hands-on assignment. Do the thing, photograph it, send it to me. Audio follow-up or video work. These vary.(Varied pieces)

Both paths are worth the same points. I tell them to mix it up. Busy week, with kids or another class or a job? Take the one that gives you time back. Excited about an active assignment? Go deeper.

We remember what connects to us and has meaning. Choice is how I try to make room for that.

Bibliography

- Anderson, Lorin W., and David R. Krathwohl, editors. *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Longman, 2001.
- "Assessment Cycle in Higher Education: A Critical Appraisal." *Assessment at CUNY*, CUNY Academic Commons, Oct. 2025, assessatcuny.commons.gc.cuny.edu/2025/10/assessment-cycle-in-higher-education-a-critical-appraisal/.
- Association of American Colleges. *Integrity in the College Curriculum: A Report to the Academic Community*. Association of American Colleges, 1985.
- Association of Masters of the Boston Public Schools. *Remarks on the Seventh Annual Report of the Hon. Horace Mann, Secretary of the Massachusetts Board of Education*. C. C. Little and J. Brown, 1844.
- Banta, Trudy W., and Charles Blaich. "Closing the Assessment Loop." *Change: The Magazine of Higher Learning*, vol. 43, no. 1, 2011, pp. 22–27, <https://doi.org/10.1080/00091383.2011.538642>.
- Banta, Trudy W., et al. *Designing Effective Assessment: Principles and Profiles of Good Practice*. Jossey-Bass, 2009.
- Bloom, Benjamin S., editor. *Taxonomy of Educational Objectives, Handbook I: Cognitive Domain*. David McKay, 1956.
-
- Bobbitt, John Franklin. *The Curriculum*. Houghton Mifflin, 1918.
- Bobbitt, John Franklin. "The Elimination of Waste in Education." *The Elementary School Teacher*, vol. 12, no. 6, 1912, pp. 259–71.
- Bobbitt, John Franklin. *Some General Principles of Management Applied to the Problems of City-School Systems*. The Twelfth Yearbook of the National Society for the Study of Education, Part I, University of Chicago Press, 1913, p. 10. *Internet Archive*, <https://archive.org/details/somegeneralprinc00bobbbrich/page/10/mode/2up.z>
- Cowan, David. "Pittite Triumph and Whig Failure in the Cambridge University Constituency, 1780–96." *Parliamentary History*, vol. 43, no. 1, 2024, pp. 91–111, <https://doi.org/10.1111/1750-0206.12724>.
- Durm, Mark W. "An A Is Not an A Is Not an A: A History of Grading." *The Educational Forum*, vol. 57, no. 3, 1993, pp. 294–97.

SOURCES, 2 OF 3

- Ewell, Peter T. "An Emerging Scholarship: A Brief History of Assessment." *Building a Scholarship of Assessment*, edited by Trudy W. Banta and Associates, Jossey-Bass, 2002, pp. 3–25.
- Goff, Lori, et al. *Learning Outcomes Assessment: A Practitioner's Handbook*. Higher Education Quality Council of Ontario, 2015, heqco.ca/wp-content/uploads/2020/03/heqco.LOAhandbook_Eng_2015.pdf.
- "A History of Mathematics in Cambridge." *Faculty of Mathematics*, University of Cambridge, www.maths.cam.ac.uk/about/history. Accessed 24 Sept. 2026.
- "History of Standardized Testing." *College of Education*, Lehigh University, 2013, ed.lehigh.edu/news/2013/history-standardized-testing.
- Kaestle, Carl F. *Testing Policy in the United States: A Historical Perspective*. Educational Testing Service, 2013, www.ets.org/Media/Research/pdf/kaestle_testing_policy_us_historical_perspective.pdf.
- National Commission on Excellence in Education. *A Nation at Risk: The Imperative for Educational Reform*. U.S. Department of Education, 1983.
- Nottingham, James. *The Learning Challenge: How to Guide Your Students through the Learning Pit to Achieve Deeper Understanding*. Corwin, 2017.
- Plato. *Phaedrus*. Translated by Benjamin Jowett, Project Gutenberg, 1 Feb. 1999, www.gutenberg.org/ebooks/1636.
- Postman, Neil. *Technopoly: The Surrender of Culture to Technology*. Knopf, 1992.
- Ravitch, Diane. *The Death and Life of the Great American School System: How Testing and Choice Are Undermining Education*. Basic Books, 2010.
- Reese, William J. "First Race to the Top." *Larry Cuban on School Reform and Classroom Practice*, 17 Sept. 2013, larrycuban.wordpress.com/2013/09/17/first-race-to-the-top-william-reese/.

SOURCES, 3 OF 3

- Reese, William J. *Testing Wars in the Public Schools: A Forgotten History*. Harvard UP, 2013.
- Scriven, Michael. "The Methodology of Evaluation." *Perspectives of Curriculum Evaluation*, edited by Ralph W. Tyler et al., Rand McNally, 1967, pp. 39–83.
- Shepard, Lorrie A. "The Role of Assessment in a Learning Culture." *Educational Researcher*, vol. 29, no. 7, 2000, pp. 4–14, <https://doi.org/10.3102/0013189X029007004>.
- Soderstrom, Nicholas C., and Robert A. Bjork. "Learning versus Performance: An Integrative Review." *Perspectives on Psychological Science*, vol. 10, no. 2, 2015, pp. 176–99, <https://doi.org/10.1177/1745691615569000>.
- Sousa, David A. *How the Brain Learns*. 6th ed., Corwin, 2022.
- Study Group on the Conditions of Excellence in American Higher Education. *Involvement in Learning: Realizing the Potential of American Higher Education*. National Institute of Education, 1984.
- Tyler, Ralph W. *Basic Principles of Curriculum and Instruction*. U of Chicago P, 1949.
- "The University of Cambridge: The Age of Newton and Bentley (1660–1800)." *A History of the County of Cambridge and the Isle of Ely*, edited by J. P. C. Roach, vol. 3, Victoria County History, 1959, pp. 210–35. *British History Online*, www.british-history.ac.uk/vch/cambs/vol3/pp210-235.
- Warwick, Andrew. *Masters of Theory: Cambridge and the Rise of Mathematical Physics*. U of Chicago P, 2003.
- Wiggins, Grant, and Jay McTighe. *Understanding by Design*. Association for Supervision and Curriculum Development, 1998.
- Wiggins, Grant. "A True Test: Toward More Authentic and Equitable Assessment." *Phi Delta Kappan*, vol. 70, no. 9, 1989, pp. 703–13.