How to teach critical thinking (and writing) without assigning essays: a novel approach

The purpose of this essay is to present a novel approach for preparing undergraduate students to use evidence wisely, particularly in the social sciences (though these tools can be amended for virtually any field). This approach is informed by science education researchers who have begun to catalog common mistakes in the ways students think or visualize scientific concepts to help instructors create pedagogies to help their students confront and avoid those mistakes. Teaching students to avoid a systematic set of logical errors before assigning grades for written work is likely more equitable to students from disadvantaged backgrounds, who may be less prepared than some of their peers through no fault of their own. Importantly, this approach can be as easy on instructors as multiple-choice exams without the need to rely on fact-based exams. Happily, my method is also AI-proof.

Keywords: critical thinking; causality; equity and ethics in teaching undergraduates, assessments

The purpose of this essay is to reflect on some problems in our current way of teaching undergraduates and to present some preliminary tools I created that I hope are not only more effective and equitable but also easier on instructors than what we are currently doing. In short, over time, I have created a teaching methodology I call "Teaching to identify logical errors systematically" (TILES) in which I created a systematic list of common mistakes in the way students think and use evidence for three kinds of arguments my students make: ethical normative arguments, legal and other forms of text analysis, and the use of empirical observations and statistics for descriptive and causal inference. Then, through two decades of trial and error, I created (and would like to share) lecture materials, in-class exercises, and assessment tools that help students confront and correct those logical mistakes. I am also in the process of creating software that can be used to facilitate a more efficient two-way communication between instructors and students than graded essays with comments, based on a similar teaching philosophy.

In this essay, I share some of the common logical errors I have noticed as well as the tools I created to help students learn to avoid those mistakes. My hope is that these strategies can help students learn to write before they are graded for their written work. I am hoping that the big selling point of these techniques is that they are easier on instructors than assigning and grading essays. This makes them easy to implement in large introductory courses without the need to fall back on testing students solely for their skills at memorizing lecture and reading materials. I then briefly summarize the scholarship from science education congruent with this approach and provide examples of common logical mistakes I have noticed. The rest of the essay (including a lengthy appendix) provides examples of exams and quizzes that can also be used as in-class practice exercises so that instructors can amend what I have done for their own courses.

Part of my motivation has been informed by the disadvantages that some students face from their poorly funded K-12 schools that punish students for what they did not learn before

they came to my class.[1] Furthermore, the ways that most universities structure their curriculum requirements can unwittingly exacerbate both the inequities and our ability as instructors to correct mistaken ways of thinking systematically. While we expect students to evaluate alternative arguments in upper division courses, the emphasis on factual knowledge in lower division courses may not be preparing them to do so.

Indeed, some preliminary evidence suggests that the focus on knowledge could be making the problem worse. Stein, et al. (2007) created a measure for critical thinking skills, loosely defined as the ability to identify that an alternative possibility exists given one explanation, with higher level skills measured as the ability to create those alternatives. They report that students who recalled having memorized knowledge in their undergraduate courses perform (much) worse on their critical thinking assessment exam. If this is true, then introductory courses that require students to memorize are literally making students worse at critical thinking. If they were prepared by a high-quality K-12 education, they do well. If they do not, we may be again, unwittingly, reinforcing those inequities (and charging them a ton in tuition for the opportunity).

But without new assessment strategies that help students understand what makes the use of a particular kind of evidence valid, we are ill-equipped to ameliorate those inequities. First, we do not have the resources to have students hand in drafts, give them comments, and reread multiple revisions for multiple kinds of essays in large introductory courses. But much more importantly, students do not have the opportunity to make all the kinds of mistakes in their written work. Even if they read the comments (they do not), they cannot learn to avoid all these mistakes with a few essays.

Moreover, the gulf between the scholarly consensus about effective and equitable ways of teaching and what we are currently doing is widening. In the words of Nobel Prize winner,

---

[1] See Reardon, Kalogrides and Shores (2019) for a disturbing account of the lack of equity in K-12 education.

physicist Carl Wieman, what we are doing in our large lecture courses is "the pedagogical equivalent of bloodletting" (2014). But this does not have to be difficult. The good news about my strategy is that once instructors have implemented these strategies in their courses, all they do is show up to class.

Three kinds of assessments, in order of difficulty

Categorizing kinds of statements

Students do not seem to understand what kind of evidence is valid for which kinds of arguments. Indeed, my inference from a multiple-choice quiz I give every student I have taught since 2007 is that many students at all levels cannot discern a fact from an argument. Even after a week of lecture with in-class exercises, many cannot easily distinguish text-based arguments, from causal arguments, from ethical arguments. What I have noticed as I have been teaching this, is that some are not straightforward: is it a fact that smoking "causes" cancer, or is this a causal argument? And given the scientific consensus, isn't this also a question of what we should and should not do? If it is unconstitutional to pray in school, isn't it also true that we should not do it? And moreover, sometimes the best of us can confuse a fact with an inference (such as the WHO tweet: "FACT: COVID-19 is not airborne!) And moreover, probabilistic causality obeys different logical rules such that exceptions do not necessarily undermine general rules. What further complicates this is that legal causality is often considered a "fact" of law even if it is not true at all and moreover, normative ethical arguments can be self-contained, but they can also depend on causality and legality.

Not surprisingly, given the complexity, the average grade generally ranges close to a B-. With Canvas, I am able to test my items with item response analysis to ensure that the answers to each item is correlated with doing well on other items on the quiz. I have also allowed students to argue for how they read the question, which allows me to improve the quality of the items, but

also ensures that it is fair. (Some example statements from such quizzes are attached to this paper in Appendix A).

"Essay" exams

I started including essays that I write as an open note, open book, open internet exam, with a list of what I call "potential critiques." Students must indicate whether these critiques are valid. The essays have grammatical errors (not ones I did not tell them about), factual problems from the readings of the class, the usual logical fallacies, and they use the incorrect kind of evidence for an argument. The list of potential problems also have problems: they accuse the essay of factual errors that are not errors, have their own logical fallacies, etc. This teaches them the principles of writing but does not grade them for knowing anything that I did not teach them explicitly, which is much more equitable. It forces them to learn logical principles rather than allowing them to turn in lazy essays. (Some example essays with their list of potential criticisms are in Appendix C).

Critical thinking multiple choice

The other new assessment technique I have been using are what I call "cruel and unusual" multiple-choice exams. These are extraordinarily difficult to write, but they are easy to grade, and very easy to revise such that they cannot easily share the exam or cheat. The questions are usually 1-5 sentences with a long list of potential answers with a set of A and C, all but G type answers. They might ask something like: which of the following pairs of reading are the most inconsistent logically? Or: put the following conspiracy theories in order of them being possible, given what we have learned about collective action theories. With the constitutional law class, these easily work as hypotheticals where the answers are the most likely outcome in court. With my statistical methodology courses, I can present some regression tables with a researcher's interpretation and then ask them a list of questions about whether those interpretations make sense. I can include problems related to the logical fallacies in the mix as well.

I was surprised that it generally takes students an hour to answer five multiple choice questions! The exam can be open note, open book, and open internet. I test for knowledge in ways that are not easy to look up during the exam, making it impossible for a student who has not been coming to class to do well. They have time to look up a fact that they might have missed in the reading, but they don't have time to comb the reading for the facts. This helps solve problems of the students not knowing a word I hadn't anticipated, or other problems I associate with equity. Further, they have to know the material but do not need to memorize. (Some example exams can be found in Appendix B; one of them has detailed commentary on the purpose of the question along with an explanation for why the answers are correct or incorrect. I have also included some examples of exams where I present mistakes in statistical research that the students are supposed to identify from my course in methodology at the graduate level in Appendix F).

A note about common logical errors in three types of courses

Teaching causal inference and statistics: concepts and units

In my experience, it takes a full six weeks for students to understand concepts and units, whether they are freshman or PhD students. And they literally can make no progress in a methods course until they get it. Causal thinking requires an empirical connection between two concepts that vary across a set of units. First of all, at their most basic, concepts must be unidimensional, with a qualitative or quantitative way to distribute units along that dimension, with the poles on either extreme defined and (mostly), and with the quantitative "high" or large values implied by the name. In this way, ideology and party identification are not concepts, but conservatism and Republican identification are.

Some political phenomena are not concepts, by this definition. So, the first quiz I give is a list where the students are asked to evaluate whether they are concepts. "Rules," "local governments" and "bilingual education" are examples of political phenomena that are not

concepts. They are also asked to evaluate which units could be attached to the concepts. Murder rate, as an example, is not an individual level concept. Urbanization cannot be used as a concept for cities. In twenty introductory methods courses, where I ask what units can be attached to urbanization, I have never had a student say countries; they always guess cities. It is not an easy skill to develop and yet we expect them to just figure that out on their own. The tendency is to teach the math or probability and the rest will follow.

PhD students have trouble with unit concept matching in the introductory statistics course. But they have a need to get more sophisticated more quickly. I had a student who had a theory that the price of opium in Afghanistan leads to more terrorist attacks, which on one level makes perfect sense because he believed the Taliban sells opiates for money that they then use to orchestrate terrorist attacks. The problem is that the units were regions. It took him an entire semester to be convinced that the units should have been years; price doesn't vary sufficiently by region and the terrorist attacks don't happen in the same regions they generate the money from opium. And it can even get much more difficult than that. In my own field of Supreme Court scholarship, we have to consider whether the proper units are issues, case citation, docket number (there are often multiple docket numbers per case citation, only reversals, justices, justice*year or policy area*year, etc.) This illustrates how prepared they need to be on this subject but we don't give them enough examples when we focus too much on the math or the statistical software. We need examples where they can learn this because depending on them just to be smart enough to figure it out causes retention issues in a graduate program. (An example of these exam questions assessing the skill of concept-unit matching can be found in Appendix E).

Teaching text analysis: the building blocks of interpreting caselaw

The objective of text analysis is to argue about the appropriate meaning of the text, given alternatives. This requires some very difficult steps. First, students must be able to identify phrases or places from the text that would give rise to alternative plausible interpretations. Then, they have to provide two interpretations. The difficult part about this is that they must be

logically distinct: they must lead to different interpretations that are sufficiently different from one another such that it makes a difference which one we would believe. In case law interpretation, phrases have to be identified that would lead to a different conclusion about what the law is, given a hypothetical. Is a described police search constitutional? Is some hypothetical funding of religious purposes allowable?

To teach this, I ask them to find a phrase or set of text to interpret. Then I ask them to provide two logically distinct but plausible interpretations given a particular hypothetical legal question. An example I use is whether the underlying logic in the case that created the right to privacy necessitated the logic where we got the right to sexual privacy in general. The text is very short, and there are lots of places that are ripe for multiple interpretations. This is a building block step because their job, if they were to write a paper would be to make an argument about what the words in the case mean, given a counterargument. The counterargument would be a strawman if it is not plausible. And it is not a counter if it is not logically distinct; the alternative interpretation would have to lead to a different conclusion from the main argument to be considered a counterargument. That students find this very difficult suggests that the building block step is an important one to teach, before they learn to write about what the law means (this applies to religious texts, or works of philosophy as well). (Examples of this assignment can be found in Appendix D).

Teaching the building blocks of normative arguments

Teaching how to write careful normative arguments may be the most difficult task, probably because making these arguments is the most difficult. This is also fraught with the complication that students' identities are most likely to be connected to their normative positions than their positions on causal relationships or interpretations of text. These are also problematic because they are the most likely to be susceptible to students' perceptions that instructors' own normative positions get in the way of assessments. For this reason, this skill is likely to be the

most important to teach in terms of the building blocks rather than having students write an actual normative essay.

The best resource I have found here is the first chapter of Tom Regan's *Matters of Life and Death*, where he outlines how to make a normative argument from a first principle. He explains that what you do is argue a specific proposition carefully from a first principle. He also outlines what to do and most importantly, what not to do. The idea is to make the argument as intersubjective as possible, making sure to figure out which first principle is likely to appeal to those who would be likely to disagree with the specific proposition (in Regan's case, he is making the argument for veganism – in 1983, so he knows how to argue something that most of the planet disagrees with). Then, he outlines the logical fallacies that might be tempting. I like this essay because it makes it clear to the students that rants are not acceptable.

The assignment I give is this: write a specific proposition, such as "accepting money for a political campaign is unethical." Then, defend the first principle that would be most likely to appeal to your most serious opponents. So, then, students are told not to make the case, but rather they are to argue how a case could be made from a first principle. This way, their identities are not as threatened when they are criticized for their argument. And they become aware of how difficult it is to do, which has the happy outcome that they are more tuned into logical flaws made in the case of rants.

Courselets: Software that makes two-way communication between instructor and students more efficient

I will be using these assessments in a large introductory course, combined with practice exams such that the grade for the practice is based on completion. These practice assignments are due two days before the exam in question. I use Courselets software (free for instructors of all levels at courselets.org) created by Professor Christopher J. Lee (see Lee et al. 2018) to facilitate two-way communication between myself and students. When students confront a difficult logical problem, they can provide an answer along with their commentary about how difficult they find

the problem in question. If they answer incorrectly, they ascertain that immediately by reading what I have written about the wrong answer. They can also explain (from a given set of explanations) what they think led them astray. Students are hearing from me about the underlying logic of a particular skill, in real time, just before they get the next problem, and will have the opportunity to revisit the foundational skill necessary to get the question correct. I am hearing from them how often they think a particular problem is.

An LMS can tell me how often students got an answer incorrect, but the students themselves tell me how hard the question was for them and what more foundational skill they needed to improve to help them do better in the next exam. This information can then be revisited in lecture if enough students were confused or can be explained in more detail before the exam. Using this technology should prepare all students to get an A on the exam, leading to higher retention rates. Indeed, Professor Lee's retention rates skyrocketed after implementing this technology. My hope is that this will also lead to higher retention rates in upper division courses, as they will have had a chance to practice the logic related to a variety of kinds of arguments that they will be required to make as they progress through their career.

Conclusion

I have outlined some problems I see in how we teach undergraduates and come up with a plan for countering what I believe to be their most common pitfalls, which has to do with various logical fallacies, and in particular, the fallacies associated with identifying which kind of evidence is required for which kinds of arguments. I have tried to identify the building blocks to the logical steps in writing essays and communicated some examples of assignments and assessment strategies that help the students begin to learn to think before they are tasked with writing an effective essay. I am also hoping this produces a more equitable situation, in particular for underpaid untenured faculty.

What I envision is a new kind of textbook, new kinds of assignments and assessments, and a new system of sharing resources that we could share in a password protected internet

venue for faculty so that the task of developing B- essays, or critical thinking multiple choice exams, and other resources can be created more easily. The K-12 environment has this down where the curriculum is mostly written for teachers, but we as higher education teachers have only begun to share resources at all.[2] The way to make this work for instructors is to create a system where questions can be shared on a larger scale. I am hoping to be a part of the beginning of this process.

[2] There are already some resources: APSA, and poliscidata have some websites. And there are teaching with technology resources as well, but these are mostly syllabi. Sharing exam resources would have to be a password protected space, perhaps provided for by a textbook company.

References

Bok, Derek. 2006. *Our Underachieving Colleges: A Candid Look at How Much Students Learn and Why They Should Be Learning More*. Princeton, NJ: Princeton University Press

Brown, Scott W., and Frederick B. King. 2000. "Constructivist Pedagogy and How We Learn: Educational Psychology Meets International Studies." *International Studies Perspectives* 1 (3): 245–54

Burbules, Nicholas C., and Rupert Berk. 1999. "Critical Thinking and Critical Pedagogy: Relations, Differences, and Limits." In Critical Theories in Education: Changing Terrains of Knowledge and Politics, ed. Thomas S. Popkewitz and Lynn Fendler, 45–66. New York: Routledge

Facione, Peter A. 2000. "The Disposition toward Critical Thinking: Its Character, Measurement, and Relationship to Critical Thinking Skill." *Informal Logic* 20 (1): 61–84

Fitzgerald, Jennifer, and Vanessa Baird. 2011. "Taking a Step Back: Teaching Critical Thinking by Distinguishing Appropriate Types of Evidence". *PS: Political Science & Politics* 44(3) 619-624.

Grant, Michael, and Marshall Smith. 2018. "Quantifying Assessment of Undergraduate Critical Thinking." *Journal of College Teaching & Learning* 15, no. 1 (2018): 27-38.

Lee, Christopher J, Brit Toven-Lindsey, Casey Shapiro, Michael Soh, Sepideh Mazrouee, Marc Levis-Fitzgerald, and Erin R Sanders. 2018. "Error-Discovery Learning Boosts Student Engagement and Performance, While Reducing Student Attrition in a Bioinformatics Course." CBE—Life Sciences Education 17 (3): ar40.

Monk, Martin, and Jonathan Osborne. 1997. "Placing the History and Philosophy of Science on the Curriculum: A Model for the Development of Pedagogy." *Science Education* 81 (4): 405–24.

Noel, Hans. 2010. "Ten things political scientists know that you don't." In *The Forum*, vol. 8, no. 3. De Gruyter.

Pearl, Judea, Madelyn Glymour, and Nicholas P. Jewell. 2016. *Causal inference in statistics: A primer*. John Wiley & Sons.

Reardon, Sean F., Demetra Kalogrides, and Kenneth Shores. "The geography of racial/ethnic test score gaps." *American Journal of Sociology* 124, no. 4 (2019): 1164-1221.

Regan, Tom. 1980. "Introduction." In *Matters of Life and Death*, ed. Tom Regan, 3–34. New York: Random House.

Smith, Daryl G. 1977. "College Classroom Interactions and Critical Thinking." *Journal of Educational Psychology* 69 (2): 180–90.

Stein, Barry, Ada Haynes, and Michael Redding. 2007. "Assessing Critical Thinking in STEM and Beyond." In *Innovations in E-Learning, Instruction Technology, Assessment, and Engineering Education*, ed. Magued Iskander, 79–82. Dordrecht, The Netherlands: Springer.

Stein, Barry, Ada Haynes, Michael Redding, Kevin Harris, Mark Tylka, and Elizabeth Lisic. "Faculty driven assessment of critical thinking: National dissemination of the CAT instrument." In *Technological Developments in Networking, Education and Automation*, pp. 55-58. Springer, Dordrecht, 2010.

Stephens, Nicole M., MarYam G. Hamedani, and Mesmin Destin. 2014. "Closing the social-class achievement gap: A difference-education intervention improves first-generation students' academic performance and all students' college transition." *Psychological Science* 25, no. 4 (2014): 943-953.

Terenzini, Patrick T., Leonard Springer, Ernst T. Pascarella, and Alberto Nora. 1995. "Influence Affecting the Development of Students' Critical Thinking Skills." *Research in Higher Education* 36: 23–39.

Tsui, Lisa. 2002. "Fostering Critical Thinking through Effective Pedagogy: Evidence from Four Institutional Case Studies." *Journal of Higher Education* 73 (6): 740–63.

Tsui, Lisa. 1999. "Courses and Instruction Affecting Critical Thinking." *Research in Higher Education* 40 (2): 185–200.

Wieman, Carl E. 2014. "Large-Scale Comparison of Science Teaching Methods Sends Clear Message." Proceedings of the National Academy of Sciences 111 (23): 8319–20.