

Fostering Citizenship Skills with Kritik in the Undergraduate Classroom

by

Karen M. McCurdy and Jacek Lubecki
Political Science and International Studies Department
Georgia Southern University
kmccurdy@georgiasouthern.edu
jlubecki@georgiasouthern.edu

Paper Presented at the Tools for Civic Engagement and Digital Literacy Panel in the
Civic Engagement Track

American Political Science Association Teaching & Learning Conference
February 10-12, 2023
Sheraton Inner Harbor Hotel, Baltimore, MD

Abstract

Citizen Engagement is a skill that people learn from other human beings. The role of the undergraduate classroom in political science fostering citizen engagement has not been examined in detail. This paper focuses on the use of a new pedagogical software product that can be used to increase the ability of political science students building arguments in support of or opposed to public policy positions. Kritik was used by two instructors in a variety of undergraduate course offerings. This paper will provide initial findings from three semesters of use of the Kritik software.

The question of what tools can be used in the classroom to foster the development of political discourse is urgent. The propensity for American political discussion to devolve into shouting matches has been on public display for about a decade and is worsening. It is reasonable to believe that political scientists should worry about this problem, and explore various solutions.

The authors implemented a new pedagogical tool, Kritik, in several of our courses during academic year 2021-2022, and in fall 2022. We are reporting preliminary results of our experience with this software in Face to Face, software supported courses. We used Kritik in two small sections of Introduction to American Government (65 and 45 students); and in several different upper division political science courses, Russian Politics (12 students), Politics of the Middle East (16 students), State and Local Government (15 students), and Environmental Politics (6 students).

The first course is required of all undergraduates at our university, while the upper division courses are electives in the Bachelor of Arts degree program in our department. We implemented Kritik in three different contexts, and are curious whether the software helped our students better articulate political concepts. Our ultimate question is whether this software will help improve political discourse. While the data we intend to report here is necessary but not necessarily the sufficient evidence needed to answer that query, we are convinced that Kritik is a useful tool in most Face to Face political science classroom environments.

The Societal and Pedagogical Problem

Educational approaches come and go, and each of us has been the beneficiary or not of pedagogical techniques that were popular in their day. Each of us in preparing courses for the college level students in our classroom sifts through the things that worked and those that didn't in our own educational experiences. This is traditional and common sensical.

But how should we consider technological innovations that influence our approaches to teaching? The authors of this paper are old enough to have not experienced multiple choice exams in our primary educations, and one of us did not encounter them in a large state university education, lo these many years ago. Each of us however has made difficult decisions about how to handle increasing class sizes. Is it reasonable to grade essays in classes of 250, 100, or 50 students? How many essays should one assign to students in large classes? What is the cost-benefit for the faculty member in terms of additional student benefit?

These are questions that are beyond the scope of this paper. What we present here is the preliminary data from a new software platform designed to give undergraduate students increased practice with the essay format. Kritik is the software that we used, although there are now several companies offering similar products. We had a good experience with Kritik, enough so that we participated in a one hour web [workshop](#) last summer with Valentina Kozlova from the University of Alberta who uses Kritik for group projects.

There are broader societal consequences from at least two educational innovations of the past generation that are worth considering here. One is the introduction of optical mark recognition (OMR) technology in universities and the other is the results based national educational reform approach signed into law by President George W. Bush on January 8, 2002, when the No Child Left Behind Act became [PL 107-110](#). NCLB was amended in 2015, but continues the process of annual standardized testing in grades 3-12 in the United States public schools.

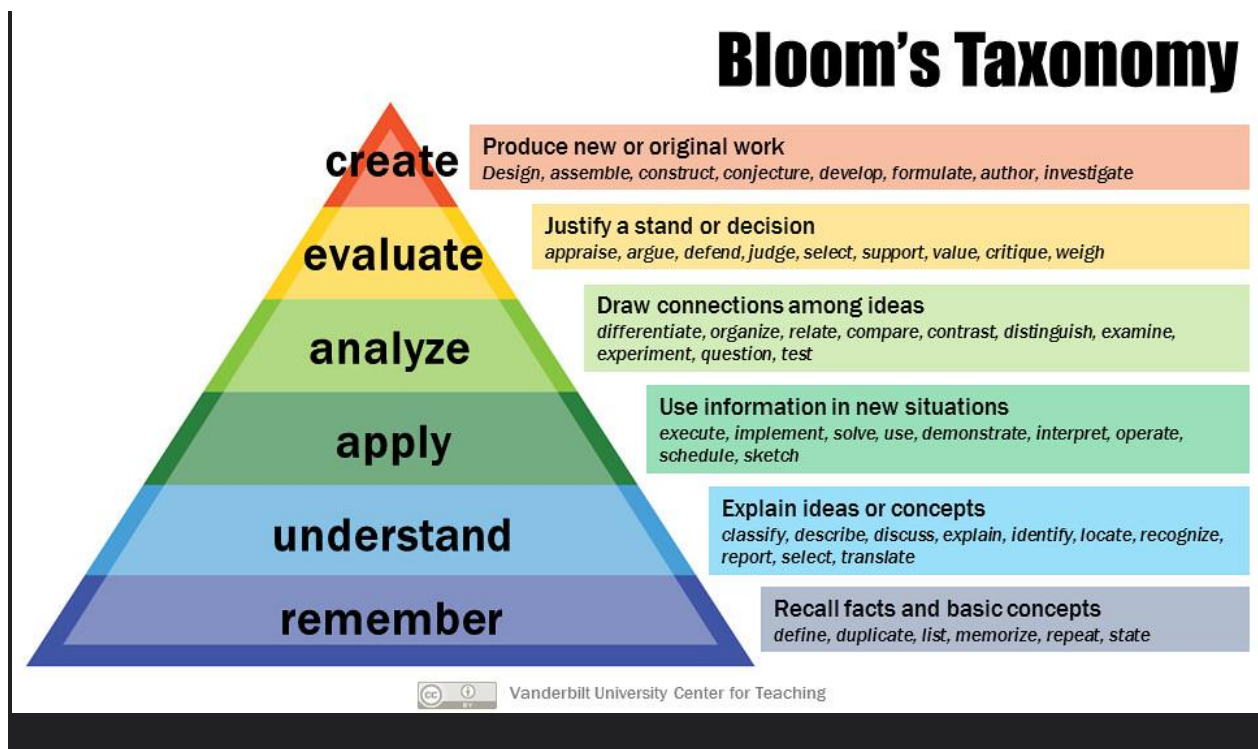
A brief history of Optical Mark Recognition Technology

Optical Mark Recognition (OMR) was developed in the early twentieth century. Several successful approaches that could be used for standardized educational tests were patented in the late 1950s and early 1960s, becoming commercially viable by the mid-1960s. These approaches were dependent on the IBM punch card, having developed out of the computer industry. The Scantron company deployed a new approach in 1972, when inexpensive readers were marketed to schools, with the company making its profits by selling scantron forms. This is the version of OMR that most of us are familiar with, either from surveys, student evaluations, examinations. The PSAT, SAT and GRE examinations may be the earliest encounters that the oldest of us in the profession had with OMR. The younger members of the professorate probably engaged with OMR in their primary schools.

What is wrong with using OMR technology in the undergraduate Political Science classroom? Nothing and everything. There are facts that students should learn,

examinations designed to be scored by OMR technology do a good job of assessing those facts. Yes, as we all know from our assessment workshops and annual preparation of assessment documents, when we rely on standardized tests, no matter how sophisticated the multiple choice question, we are assessing our students only at the recall and understanding categories of Bloom's Taxonomy (see Figure 1-1). How can we move undergraduate students into the apply/analyze/evaluate levels of the taxonomy, without resorting to unrealistic amounts of time reading student essays? Perhaps there is a technological solution that can help political science faculty help students in large format classrooms apply, analyze, and evaluate political information.

Figure 1-1: Bloom's Taxonomy



Source: Vanderbilt University Center for Teaching

A brief overview of Results Based Educational Reform

Communities and their elected officials have engaged intensely about the education of children, at least since the mid-nineteenth century in Great Britain, including a consideration of educational quality, the controversial notion of Payment by Results which would be familiar to twenty-first century parents, educators, and elected officials (Loveless p 15). School accountability has been a difficult policy arena from the beginning, with frequent alternating groups of traditional and progressive reformers changing what occurs in textbooks and classrooms. Complaints about annual examinations and demoralized teachers complaining that their instruction had been reduced to drill and rote memorization. Teaching to the test then is not a new phenomenon, although it comes with a twenty-first century new label. In the United States, education reform also began in the nineteenth century with the Committee of Ten created by the National Education Association in 1892. This committee was charged with aligning high school education and college, was comprised of a group of college presidents, high school principals, and was co-led by the U.S. Commissioner of Education, William T. Harris and the president of Harvard University, Charles W. Eliot. The committee recommended that all students receive the same liberal education at the high school level which would prepare them for life whether that included a college education or not based on four core subject tracks. The progressive education movement, heavily influenced by scientific management principles, immediately responded critically to the report. By 1918, the published critique appeared as *The Cardinal Principles of Secondary Education*, which proposed seven objectives of

secondary education. The two camps have been active and vocal in U.S. educational policy since these initial debates and published reports.

The most recent iteration of this philosophical pedagogical debate is where the professoriate and our students find themselves today. The authors of this paper interject the additional concern of educating students in civic engagement in our political science classrooms. Necessarily the discussion of civic learning is constrained by the older and more general pedagogical debate. Changes in state university classroom size precipitated by the dramatic expansion in the number of students following World War II brought the need for adjustments to traditional teaching techniques to the fore. Those of us employed at public institutions in Georgia that recently experienced rapid growth are well acquainted with the dilemma of how to retain essay assignments when class size grows beyond 50. And then there is technological innovation.

No Child Left Behind and the Common Core State Standards are simply the most recent iteration of a very long societal concern about whether our children are being adequately prepared by the educational infrastructure. An ancillary aspect of this concern is a cost-benefit concern. It is important for our discipline to place our pedagogical choices within this larger social and political context.

Our disciplinary responsibilities to civil society have not been discussed much during my career. Students move through Introduction to American Government core courses, and yet there is little professional discussion about how what happens in our classrooms influences our country's political discourse. This paper is a call to change the focus of our professional attentions to building civil capacity for political discourse during the wide variety of locations where that discourse has degenerated to angry

words even to the point of violence. Students need to learn to build a political argument, and to defend their position. Underlying all of political philosophy is the truth that there is no single correct answer to how to help a polity flourish. We should be helping students understand this not only in the upper division political science courses, but most importantly in the Introductory course required of most Georgia students. The remainder of today's paper provides our preliminary impressions of a software program that can facilitate teaching this important skill.

Kritik software as a pedagogical tool to enhance civil society

Kritik is a peer-grading platform that distributes fair and accurate assessments by harnessing collective intelligence to simplify workflows and reduce turnaround time on feedback. The "Kritik" peer grading platform app was used in our respective courses, POLS 4490 Politics of Russia by Jacek Lubecki, and POLS 3330 State and Local Politics and POLS 1101 Introduction to American Government by Karen McCurdy in the fall 2021 semester. This paper will present "Kritik" as a teaching and learning tool, and summarize the experience of its use in McCurdy's and Lubecki's courses. Data will be provided to answer the question whether the tool fulfilled its purposes, which include an improvement in students' quality of writing, and engagement with each other. Briefly, evidence for the former is ambiguous, while the latter has been obviously achieved. Given the limited and preliminary nature of the data, more experience with the tool is needed to reach firmer conclusions.

Kritik is advertised as following by its creators:

Kritik is a peer-grading platform that distributes fair and accurate assessments by harnessing collective intelligence to simplify workflows and

reduce turnaround time on feedback (Kritik website, <https://us.kritik.io/welcome>, accessed Feb. 15, 2022).

The tool does precisely what it promises to do. It allows the faculty to set up a substantive written assignment (typically, an essay or a research paper) together with an anonymous peer grading and comprehensive evaluation process.

The instructor creates the assignment prompt, together with a set of detailed instructions for the assignment and a rubric that student peer evaluators can use.

The students are then guided on how to use the tool – for instance, through a calibration exercise. Students are asked to evaluate each other's work, to give meaningful feedback to each other, and evaluate the feedback that they each received from a peer. Normally, all three elements of the work are given credit: creation of the substantive essays, the peer evaluation, and feedback or the student response to each peer's evaluation. The instructor can set the credit value of each evaluative element as desired. In our case, for instance, Jacek Lubecki used gave 70% of credit to substantive essays (so, based on the evaluations that students gave to each other), 20% of credit was based on quality of evaluations, and 10% of credit was based on students' evaluations of each other's feedback. McCurdy was more cautious in using only the Creative score in calculating fall semester final grades in both courses. With the positive instructor's experience, the Kritik component of the final grade was increased for spring semester.

Overall, and on face value, the system appears useful, for several reasons:

- 1) For instructors who have long wanted to embed peer-evaluations of substantive written work into their classes, the challenge had often been technical and workload-related: the challenges of distribution of the papers, their anonymity,

setting up a system of grading, and feedback were also laborious processes. The system takes care of this entire burden.

- 2) Likewise, for teachers who want students to write a number of short essays – pedagogically the best technique of engaging students in complex learning activities – the workload of grading all these papers, especially their formal qualities (grammar, etc.) is often impossible. The system takes the burden from the teacher, who can focus on the substantive ideas expressed in the essays, instead of their formal qualities.
- 3) Most importantly, students learn how to evaluate each other's substantive ideas and formal qualities of writing, thus developing higher-level skills of critical perspective and critical evaluation of their own and each other's work.

Thus, the case for adopting “Kritik” in a class that features a number of essay assignments seems to be compelling. However, how did it work in practice? Did the use of the tool actually improve the quality of student writing in the course of the courses taught by us? Did “Kritik” improve students’ critical perspectives and ability to give feedback to each other? While the evidence is limited (for instance, we did not conduct our own evaluations of all the student essays separate from the students’ own “Kritik” evaluations) it is worth presenting the preliminary impressions of the use of the tool over the course of fall 2021. Although McCurdy did evaluate the final essays in both of her courses, and was anxious to see whether the quality of the final essays had suffered from peer review in the earlier weeks of the semester.

Jacek Lubecki used Kritik in his POLS 4490B Politics of Russia course, which enrolled 13 students, who were asked to write and evaluate 9 weekly essays on various topics. Here the summary of average scores for various essays in his course:

Table 1: Average Kritik Scores in POLS 4490 Politics of Russia course

#	Essay questions	Overall	Create	Evaluate	Feedback
1	How necessary and "good" was Putin for Russia?	74%	78%	75%	50%
2	The nature of Putin's regime in a context.	85%	86%	85%	85%
3	Did democracy have to fail in Russia?	82%	81%	81%	92%
4	Kritik discussion of patrimonial capitalism and semi-authoritarianism	82%	76%	91%	72%
5	Kritik discussion - Russia as a post-modern dictatorship. A warning to the world?	90%	91%	89%	92%
6	Forms of delirium. The West.	80%	81%	77%	92%
7	The meaning of communism and its fall	83%	82%	83%	92%
8	Essays on Russian and Soviet military and imperial experience	80%	79%	84%	69%
9	Understanding Russian foreign policy	89%	91%	85%	92%

Notice that scores on all dimensions improved dramatically with the second essay, but, then, the progress was uneven and more elusive, with the one exception of the feedback quality, which, from the 4th essay on stayed consistently high with the exception of essay 8. The data actually gives the impression that overall quality of the student performance peaked with essay 5 – in many ways, the most intellectually challenging and creative of essays.

A more granular analysis of the data by the creation rubric categories gives the following metrics to us:

Table 2: Student scores by rubric categories on seven individual essays

Rubric categories	Essay 1	Essay 2	Essay 3	Essay 4	Essay 5 (post-modern dictatorship)	Essay 6	Essay 8	average of categories
Concept Understanding	91%	88%	91%	83%	94%	86%	95%	90%
Clarity of Thoughts	89%	87%	86%	84%	93%	87%	92%	88%
Accuracy	87%	89%	92%	84%	95%	92%	97%	91%
Grammar	79%	87%	83%	83%	90%	90%	90%	86%

Conclusion	88%	83%	91%	79%	88%	85%	91%	86%
Creative and critical thinking	88%	87%	89%	84%	90%	86%	95%	88%
Source and evidence	69%	78%	86%	79%	90%	87%	95%	83%

Please, notice that only 7 out of 9 essays are covered by the tables, as two of the essays had slightly different rubric categories (a category: “applicability to real life” was added to “evaluation of communism” and “Russian foreign policy essays”). What stands out from the analysis is consistently high performance on “concept understanding,” “clarity of thought” and “accuracy” categories, as well as a strong progress that students made on “sources of evidence” categories. The very high score of essay 5 (Russia as post-modern dictatorship) is also in evidence. What made essay 5 so outstanding in terms of students’ scores is a bit of a puzzle. The essay prompt was as follows:

Kritik discussion - Russia as a post-modern dictatorship. A warning to the world?

Objectives:

To develop a critical perspective on Russia's realities and Putin's role in Russia.

To learn about each other's perspectives on Russia.

To get familiar with Kritik as a tool.

Instructions

What, according to Pomerantsev, makes Russia into a perfect example and perhaps a trend-setter for a "post-modern dictatorship"? What is the essence of the concept? What danger to democracy does this model pose, and is it a warning for the United States and Western democracies? What are the remedies?

The topic was very conceptual, and asked these upper division students to bring in knowledge of not just Russia, but also the United State, a topic in which most of these students specialize. This perhaps explains than their better than average performance – in each other’s eyes. In this respect, given that the topic was potentially contentious and controversial, the remarkable high scores that students gave to each other might be evidence of “Kritik’s” ability to foster not just good writing, but also good public discourse on controversial issues.

McCurdy used Kritik in two different types of courses in fall semester 2021, the Introduction to American Government course with 64 students, and the State and

Local Politics upper division elective with 15 students. The assignments used in peer evaluation were also quite different in the two courses. In both courses, the pedagogical goal was to give students greater interaction with the work of their classmates to facilitate peer learning. Students had direct knowledge of the work of nine of their peers in POLS 1101, and the written work of six of their peers in POLS 3330.

The immediate improvement for students using Kritik is that their evaluations were available to them within fourteen days (mostly because of the instructors inexperience with the software). With greater faculty experience, this delay in receiving evaluations is reduced to seven days in Spring 2022 (the five days of the evaluation phase followed by a two day feedback period). The peer evaluation scores for Essay One, an average of 57, are slightly higher than the historical pattern of instructor evaluated submissions which have ranged from 35 to 60 in various semesters.

Fall 2021 was the first semester returning to the normal face to face classroom. Everyone was a little rusty, and the lack of stamina of the lower division students was particularly notable. The average score for essays two and three generally rises slightly in a typical semester. Somewhat surprisingly last fall the average scores fell to 45 for Essay Three from an average of 57 for both Essays One and Two. The final exam essay average rebounded though with an average score of 93. Given the instructor learning curve, I was encouraged by these results enough to continue using Kritik in Spring 2022. The results for the initial essay in Spring 2022 are encouraging, with an average creation score of 63.

Table 3: Comparison of Scores Across Assignments POLS 1101 Fall 2021

	Topics	Overall Score	Creation Score	Evaluation Score	Feedback Score
Essay One	Political Demography and the Founding Generation (Quant & Qual data)	60	57	67	47
Essay Two	Civil Rights in Comparative Perspective: 2021 to 1832 and 1941 (Qual data)	57	57	55	62
Essay Three	Policy Letter to a Member of Congress; and Subcommittee Analysis (Quant & Qual data)	45	45	43	56
Average Kritik Essay Scores		54	53	55	55
Final Essay (Scored by Instructor)	How are artists using their form of popular culture to influence public policy outcomes? (data choice open)		93 Range 82.5 to 100		

Source: Kritik Course Statistics compiled by instructor

The way that McCurdy uses Kritik in the upper division is to support a multi-part guided research project, with three opportunities for feedback and revision. A proposal of five to seven pages is reviewed in week six of the semester, a first

draft of twelve to fifteen pages is reviewed in week 10 of the semester, with a public presentation in weeks twelve to fourteen. The final revision of the research report is due on the last day of class. I have been looking for ways of allowing students to see the high quality of research projects that are generated by their peers for a decade or more. It seems unfortunate that I am the only one who gets to read all of these fantastic papers. I was very hopeful that Kritik would provide the opportunity for expanding peer learning into the written work in my upper division courses.

Table Four: Comparison of Scores Across Assignments POLS 3330 Fall 2021

	Combined Score	Proposal	First Draft
Overall Score	67	68	66
Creation Score	73	67	78
Evaluation Score	59	70	47
Feedback Score	72	63	80

Note: Evaluation and Feedback scores were not included in the final grade calculation, and the drop in student participation in the evaluating drafts was noticeable.

Source: Kritik Course Statistics compiled by instructor

Upper division students are better at following directions from the outset of the semester, as indicated by the ten point increase in the creation scores. Although there are always students who are taking their first upper division course who are not adept at including their title page and bibliographic information formatted in the Chicago Style used in Political Science publications.

From discussion in the class following the Kritik evaluation period closing for proposals and the first drafts, student's comments supported my belief that they would learn from each other. This instructor writing that typos and grammatical

oversights pull your grade down tend to roll off the backs of students. Seeing how typos and minor grammatical errors interfere with ideas being conveyed was much easier for students to grasp while reading the work of their peers.

In depth analysis of McCurdy's POLS 1101 Essay Two.

The Essays used in POLS 1101 are designed to expose students to a range of data types that are used by political scientists, and the punditry who consider themselves students of politics. For each essay students are given the materials from which they are to craft an essay. For the first essay, they are directed to data from the Census Bureau providing statistics from the colonial period, the first federal census in 1790, as well as the most recent census data; brief biographical essays of delegates to the Constitutional Convention at the National Archive and Records Administration, as well as two classic essays about the motivations of the Founding Generation. While there is a lot in this first essay, the essay is easy for graduating seniors to juggle, and quite difficult for first semester students. Part of the point of the exercise is to learn to ask for clarification when you are confused by the instructions, or having trouble negotiating the electronic resources. Honors students and graduating seniors have no difficulty asking for clarification, while students in general education sections seem to what to do essays by themselves. They then seem willing to voice their frustration after the assignment is completed, but interestingly before they receive their scores.

Essay Two does not include numerical data which makes the assignment slightly easier for the typical Introduction to American Government student. This essay however asks students to draw comparisons directly between two separate

historical instances of civil rights being abridged by government, an instance of Connecticut passing Black Codes in 1832, and the Executive Order to intern Japanese American citizens from 1942-1945, and an example of civil rights being abridged in the present. The student must select their own contemporary example of civil rights being pushed aside. This essay involves very sensitive issues, yet students are able to put aside their own ideological positions, and evaluate whether their peers crafted a good argument in their essay creation. Interestingly, the highest feedback score during the semester was associated with Essay Two. While additional analysis is needed, it does appear that students were engaged to see what their peers thought of their essays (the average feedback score is low when evaluations are not read, and not commented upon).

Appendix A: Description of the Six Political Science Courses

This paper is a progress report on a new pedagogical technique for which we have results from Fall semester 2021, and the first few assignments of Spring semester 2022 which is in its seventh of sixteen weeks of instruction at the time of the conference at which this paper was presented. Lubecki used Kritik in two upper division comparative politics electives, the Politics of Russia, and Middle Eastern Politics, while McCurdy used Kritik in two sections of the core curriculum required course in American Government along with two upper division electives in American politics, State and Local Government and Environmental Politics.

Lubecki used Kritik for weekly essays in his upper division courses. McCurdy used Kritik for a proposal and first draft in her upper division courses, and three essays occurring at approximately three week intervals in the introductory course.

Fall Semester 2021

POLS 4490, Politics of Russia, 13 students.

In Fall 2021, students were assigned nine peer-evaluated essays. These occurred at weekly intervals.

POLS 1101, Introduction to American Politics, 64 students.

In Fall 2021, students were assigned three peer-evaluation essays. These essays have been used in at least ten previous semesters. The limitation on the number of essays that students can be asked to write has been dependent on the grading turn-around time for the instructor. No grading assistance is available from the department until the course size reaches 200.

POLS 3330, State and Local Politics, 15 students.

In Fall 2021, students engaged in peer-evaluation of the first two stages of a multi-part research project. This multi-part format has been used since 1990 in all upper division and graduate courses that McCurdy has taught. The research proposal and first draft have been evaluated by the instructor, with peer evaluation occurring in the public presentation phase of the project, returning again to instructor evaluation of the final draft. With Kritik, the proposal and first drafts of the research project were added to what had been the peer learning stage of the project in previous semesters.

Spring Semester 2022

POLS 3134, Middle Eastern Politics, 9 students.

In Fall 2021, students were assigned nine peer-evaluated essays. These occurred at weekly intervals.

POLS 1101, Introduction to American Politics, 36 students.

Based on the results from the initial use of the Kritik peer evaluation platform in the previous semester, five essays were assigned for Spring 2022. This was the number of essays that had been assigned before adjustments to the course for COVID-19 were made in the previous three semesters. The first essay round Kritik creation, evaluation, and feedback had been completed before the current draft of this paper was prepared.

POLS 3231, Environmental Politics, 8 students.

Given the feedback from the initial use of Kritik for the multi-stage research project that have historically been assigned in all upper division courses taught by McCurdy, the peer evaluation platform is being used this term to provide a learning by teaching component to the earliest written elements of the research project. Students had completed the Kritik creation, evaluation, and feedback phases for their research project proposals before the current draft of this paper was prepared.

POLS 1101, Introduction to American Politics, 64 students.

In Fall 2021, students were assigned three peer-evaluation essays. These essays have been used in at least ten previous semesters. The limitation on the number of essays that students can be asked to write has been dependent on the grading turn-around time for the instructor. No grading assistance has been available from the department.

POLS 3330, State and Local Politics, 15 students.

In Fall 2021, students engaged in peer-evaluation of the first two stages of a multi-part research project. This multi-part format has been used since 1990 in all upper division and graduate courses that McCurdy has taught. The research proposal and first draft have been evaluated by the instructor, with peer evaluation occurring in the public presentation phase of the project, returning again to instructor evaluation of the final draft. With Kritik, the proposal and first drafts of the research project were added to what had been the peer learning stage of the project in previous semesters.

References

Lorin W. Anderson and David R. Krathwohl [eds.] 2001, *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*, New York: Longman.

Benjamin S. Bloom, M. D. Englehart, E.J. Furst, W.H. Hill, D.R. Krathwohl, 1956, "Taxonomy of Educational Objectives: The Classification of Educational Goals", *Vol. Handbook I: Cognitive Domain*, New York: David McKay Company.

Kathleen F. Gabriel, 2008, *Teaching Unprepared Students: Strategies for Promotion Success and Retention in Higher Education*, Stirling, Virginia: Stylus Publishing, LLC.

Charles E. Glassick, Mary Taylor Huber, Gene I. Maeroff, 1997, *Scholarship Assessed: Evaluation of the Professoriate*, San Francisco, CA: John Wiley and Sons.

Kritik. <https://www.kritik.io>

Karen M. McCurdy, *Representation and Policy Change: A Lab Manual for American Government* (2019) TopHatMonacle Inc. publisher, ISBN:978-1-77412-069-9.

Tom Loveless, 2021, *Between the State and the Schoolhouse: Understanding the Failure of Common Core*, Cambridge: Harvard Education Press.

TopHat. <https://tophat.com>