



Daniel Bernstein, UNIVERSITY OF KANSAS, djb@ku.edu

Andrea Follmer Greenhoot, UNIVERSITY OF KANSAS, agreenhoot@ku.edu

Team-Designed Improvement of Writing and Critical Thinking in Large Undergraduate Courses

ABSTRACT

Helping students achieve advanced critical thinking and writing skills in large undergraduate classes is a challenge faced by many university faculty members. We addressed this challenge in a three-year project using team course design, built around a cognitive apprenticeship model, to enhance undergraduates' writing, critical thinking, and research skills in courses ranging in size from 70 to over 400 students. Faculty members partnered with specialists from the university library, writing center, and teaching center, and with graduate student fellows who received supplemental training in those units. Together they designed progressive learning activities and written assignments based on meaningful, situated critical thinking scenarios. Instruction teams also developed rubrics for tracking students' progress on each step, and they used this information to inform the next wave of course enhancements and generate continual and iterative improvement. Assessments developed by the instruction teams showed that students in the team-designed courses improved in their critical thinking and writing skills from the beginning to the end of the semester. Furthermore, an evaluation of student work from the team-designed courses using the AAC&U Value rubrics showed that these students displayed more advanced critical thinking and writing skills than students in roughly comparable but conventionally designed courses. Our results demonstrate that team design involving specialists and graduate students can be a feasible and worthwhile strategy for engaging faculty members in developing advanced instructional and assessment designs that enhance high-end learning in a large university setting.

KEYWORDS

team-design, cognitive apprenticeship, critical thinking, writing, assessment

Large public universities face a major challenge in educating diversely prepared undergraduate students, particularly in courses that enroll hundreds of students in a single section. Although faculty members often seek to generate high-end learning goals that involve critical thinking and writing, it can be quite challenging to help students achieve these goals in large-class environments. At the University of Kansas, we undertook a three-year project to test the efficacy of team-designed courses using a cognitive apprenticeship model to enhance undergraduate writing, critical thinking, and research skills in

courses ranging in size from 70 to over 400 students. The overarching goal was to maximize the effectiveness of each course for the wide range of students who attend a large public university.

TEAM DESIGN

We approached this challenge using design teams made up of faculty members working in collaboration with specialists from the KU Libraries, the Writing Center, and the Center for Teaching Excellence. Collaboration on teaching among librarians, writing specialists, and faculty members is promoted for improving information literacy (e.g., Miller & Pellen, 2005; Raspa & Ward, 2000; Rockman, 2004) and for engaging writing across the curriculum (e.g., McLeod, 1988; McLeod & Soven, 1992). Our project used a version of collaboration whose primary purpose is achieving the instructor's course goals by improving student writing and critical thinking within the discipline (Ianuzzi, 1998). Excellent critical reasoning/research skills and the quality of writing are interdependent, and the design teams developed them in parallel. Faculty members first identified complex, expert-like tasks and then broke them down into a series of smaller, component skills that combine with knowledge to enable performance of the complex task. Our colleagues from the libraries and writing center are extremely skilled in guiding the design of staged assignments that give students an opportunity to practice each of those component skills and receive feedback. The partners engaged in collaborative, proactive instructional design instead of calling upon those colleagues for assistance only after specific skill deficits emerged.

The design team also developed rubrics to evaluate student performance and to provide feedback to students about their skill development. The staged assignment approach, along with rubrics for tracking students' progress on each step, fit very well with a cognitive apprenticeship teaching framework (e.g., Collins, Brown, & Hollum, 1991; Lave & Wenger, 1990) that guides students toward thinking more like experts in a field. Having access to evidence from the staged assignments also enabled the faculty member to identify missing skills or misconceptions early in the process and intervene as needed.

Since the potential for adoption of this approach was constrained by the amount of professional time available from library and writing center staff, we made this innovation scalable by creating graduate student fellowships as a supplement to a regular teaching assistant or research assistant position. These Graduate Student Fellows (GSFs) received supplemental training from the KU Libraries and Writing Center and then used those skills to assist with course and assignment design and to support undergraduates' work in the target course. The fellowships make the team design model sustainable in a large research university environment, while simultaneously creating new professional development opportunities for our graduate students.

TEACHING AS INQUIRY

We built the team-designed collaborations around a model of teaching as intellectual inquiry, in which each faculty member poses a question about the skills of her/his students (Hutchings & Shulman, 1999). For example, a psychology faculty member identified the critical reading of empirical research articles and synthesis of diverse research findings as essential skills for students in her class. After examining how well students showed those skills by examining students' work in previous course offerings, she then



adjusted course procedures to raise the demonstrated level of skill. Individual instructors collaborated with specialists and their GSFs to consider the strengths and weaknesses of student understanding, and then designed enhanced instruction targeting the specific weaknesses identified. The ongoing instruction team then examined student work in the next offering of the course to see if more students demonstrate the higher forms of learning. The analysis resulted in a second wave of enhancements as part of a cycle of continuous improvement in instructional design. Thus, iterative, student-learning based course enhancement was a core feature of the course redesign model.

Our primary evaluation of these team design collaborations focused on two courses: an upper-level psychology course on cognitive development and an introductory political science course on international relations. Professors in both courses engaged in iterative course redesign across three consecutive years of the project, focusing on continual improvement of students' writing, critical reasoning, and research skills. We measured growth in students' writing and critical thinking skills across each semester offering. We also evaluated the relative value added by our model by comparing student performance in these courses to performance in courses not using collaborative, cognitive apprenticeship design. Finally, we examined the generalizability of our approach beyond the two primary courses by extending the model to several courses from diverse disciplines during Years 2 and 3 of the project. Our goal with these "extension courses" was to see if the feasibility and utility of the model was applicable to a new set of courses and instructors.

Faculty members in this project shared several interrelated research questions: Can specialists in content and instructional design collaborate successfully to help students write clearer and better argued assignments that demonstrate a more sophisticated understanding of what the research in the discipline showed? Does collaborative design of staged assignments and use of shared rubric measures of writing and thinking skills generalize to a wider cohort of instructors? How well does team design work to break intellectual complexity into component skills to be taught in stages (which we identify as a form of cognitive apprenticeship)? Is it useful and sustainable to enhance the skills of graduate teaching assistants through apprenticeship in library instruction and/or a writing center? The study also allowed faculty and academic leaders to observe how insights from systematic student learning assessments can be used to produce continual, iterative improvement in student learning outcomes over time. The entire enterprise provided a test of concept, asking how well instructional teams of specialists (with specifically prepared graduate students) can function in a research university environment.

METHOD

Participants

The participants were 389 undergraduate students at a large, research extensive state university who were enrolled in the team-designed ($n=301$) and comparison ($n=78$) classes. At the beginning of each selected course, we provided information to the students about the project. We asked students for their consent to have their regular coursework included in analyses that went beyond the usual evaluation for grading for the purposes of this project. Students were told that their choice would not affect their grade in any way. All available records indicate that the majority of students enrolled in these courses provided consent (61%); most of those who did not provide consent simply did not respond

to an online request for consent. Because we could not analyze all student assignments from these courses within a reasonable time frame, we selected a subset of students who consented to the Value rubric analyses by randomly sampling within each student year cohort (e.g., first year, second year), so that the distribution by year in the sample from each class was similar to that of the whole class. The number of assignments analyzed from each class depended in part on how much student work was available. In addition, for internal purposes we were interested in comparing our Value rubric-based assessment with performance on a standardized test, so the samples included those students who also completed the standardized assessment (the number of students recruited for the standardized test varied across courses for reasons unrelated to the goals of this project).

Project overview

Figure 1 presents a schematic diagram of the activities in the project and their sequence.

Two Primary courses were iteratively redesigned and assessed each year of the project: Psychology 430, an upper-level undergraduate course on cognitive development that enrolls between 60 and 100 students, and Political Science 170, an introductory-level course on international relations that enrolls between 100 and 200 students. The faculty instructors offered these two courses once each year of the three-year project, enhanced by team design and a cognitive apprenticeship model. Each year, the instructors systematically evaluated the writing and critical thinking skills of their students, and used the evidence of student learning to inform additional course design elements in the subsequent offering. In Years 2 and 3, a different instructor took over the team-designed Political Science 170 course when the first instructor left the university. The same faculty member taught Psychology 430 all three years.

We also identified conventionally designed political science and psychology classes to provide comparison data for each Primary course. For political science, the Comparison courses were sections of the same course on international relations, taught by a different instructor without the team design approach. Psychology 430 was only offered by the primary instructor; therefore, the Comparison psychology courses were different courses taught by different instructors without team design, but were of similar size and at comparable levels of the curriculum (i.e., enrolling mostly third- and fourth-year students), and required at least one written assignment that involved critical thinking. We present a list of all courses included in this project and enrollment data for each course in Table 1. We did not include comparison classes in Year 3 for several reasons. In the Year 3 conventionally-designed political science section, copies of student assignments were not systematically retained. In psychology, we were unable to identify an appropriate comparison course in Year 3 because previous Comparison course instructors had begun to incorporate similar scaffolding methods into their own courses. Because we have no reason to believe that any of the Comparison courses would vary from year to year, we aggregated across the Comparison sections and contrast with the successive offerings of the Primary courses.

During Years 1 and 2, new cohorts of additional faculty members from various disciplines participated in a planning seminar to adopt the same methods in their courses (the Extension Courses) and implemented their redesigned courses the subsequent year (i.e., Years 2 and 3). Like the faculty instructors of the Primary courses, these faculty



Table 1. Team-Designed and Comparison Courses Included in the Project

COURSE NUMBER AND SEMESTER	COURSE NAME	COURSE TYPE	PROJECT YEAR	# ENROLLED	# ASSIGNMENTS ANALYZED
Political Science 170 Fall 2008	International Relations	Primary	1	227	18
Political Science 170 Fall 2009	International Relations	Primary	2	159	36
Political Science 170 Fall 2010	International Relations	Primary	3	180	32
Psychology 430 Spring 2009	Cognitive Development	Primary	1	100	22
Psychology 430 Fall 2009	Cognitive Development	Primary	2	68	35
Psychology 430 Fall 2010	Cognitive Development	Primary	3	67	36
			Subtotal	801	179
Political Science 170 Spring 2009	International Relations	Comparison	1	135	13
Political Science 170 Spring 2010	International Relations	Comparison	2	97	16
Psychology 605 Spring 2009	Health Psychology	Comparison	1	77	29
Psychology 370 Fall 2010	Brain and Behavior	Comparison	2	104	22
			Subtotal	413	80
American Studies 110 Fall 2009	American Identities	Extension	2	280	7
History 129 Spring 2010	US History after the Civil War	Extension	2	173	25
Astronomy 191 Spring 2011	Contemporary Astronomy	Extension	3	225	25
Biology 570 Fall 2010	Biostatistics	Extension	3	104	25
Communications 356 Spring 2011	Research Methods	Extension	3	114	20
Psychology 598 Spring 2011	Positive Psychology	Extension	3	91	20
			Subtotal	987	122
			Total	2201	391

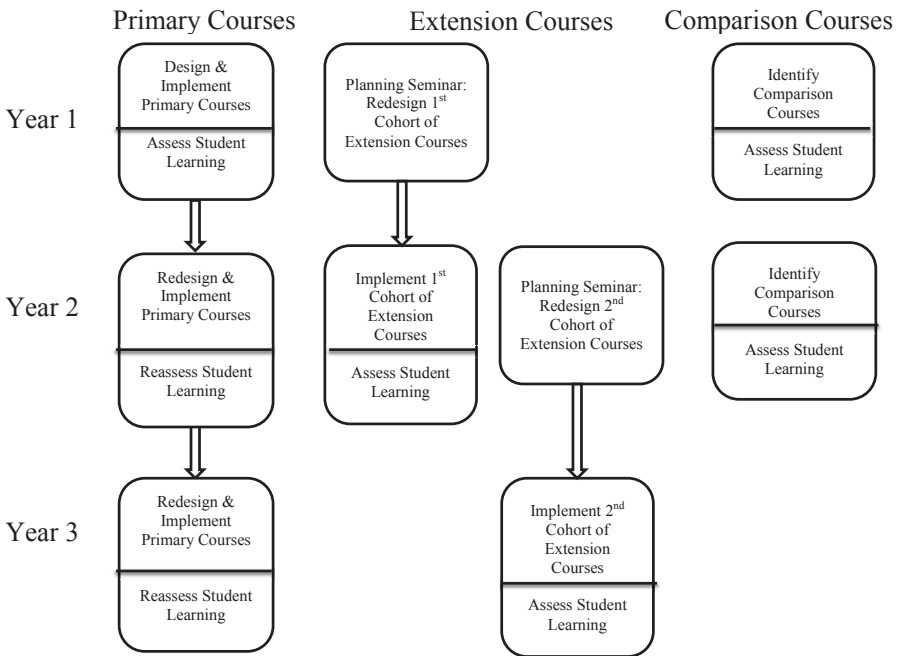


Figure 1. Project Timeline

members partnered with specialists from the library and writing center as well as one or more graduate student fellows. A list of the six Extension courses is also shown in Table 1.

Pedagogical model

The team-designed courses enhanced students’ skills of analytical and argumentative reasoning through incorporation of intensive and progressive analytical and writing assignments. Prior to or early in the semester in which the course was offered, the instructional team, including the GSFs, worked together to design the analytical and writing assignments as well as grading rubrics. Course design followed a cognitive-apprenticeship framework: faculty members began by identifying real-world critical thinking scenarios that might be encountered by a well-informed person in the field and then designed a written argument assignment situated in that meaningful context. They then delineated the mental steps involved in successful completion of this task and, working with their instruction team, created a series of staged, scaffolded assignments to help students progressively develop the skills needed for each step. Thus, the instruction team used a backwards design strategy to develop assignments and activities to enhance critical thinking and writing skill (Wiggins & McTighe, 2005). The final assignment or assignment stage represented an accumulation and integration of the progressively developed skills.

During the semester, the GSFs also consulted with students on assignments as needed. The GSFs received a small stipend to support their work on the courses. The model for recruiting and distributing the GSF positions varied across courses. Some of the GSFs were the regularly assigned GTAs for the course and received the stipend for extra work; whereas, others held research assistantships and viewed the GSF position as an oppor-



tunity to build their teaching experience. Furthermore, some faculty chose to divide the work and the stipend among multiple graduate students.

Planning seminar

In Years 1 and 2, our university teaching center convened a bi-monthly seminar that brought together a cohort of faculty members and specialists from the libraries and writing center to redesign their courses with team-supported, scaffolded writing assignments. The root metaphor for our center is shared inquiry (e.g., Shulman, 2004), so each seminar series was facilitated by a faculty member who had already implemented the pedagogical model in at least one of his/her own courses. The seminar series followed the same backwards design strategy described above, in which the faculty member designed a written assignment based on meaningful critical thinking scenarios in the field, identified the mental steps involved in successful completion of the assignment, and designed a series of assignment stages based on those mental steps. Faculty participants prepared for each seminar meeting by reading two or three articles/chapters on relevant topics (e.g., cognitive apprenticeship, teaching and evaluating critical thinking) and generating a product based on the readings (e.g., draft assignment stages, rubric draft). Thus, one very important role of the seminar was to provide a series of occasions for faculty to think through and collaborate on their course design. Faculty and other seminar participants shared their ideas and exchanged feedback in seminar meetings, refining their redesign plans across the semester.

Examples of enhancements to primary courses

In Year 1, the political science instruction team developed a six-stage writing assignment culminating in an opinion editorial (Op-Ed), in which students developed an argument about a contemporary issue in world politics and supported it with appropriate and persuasive evidence. The early assignment stages guided students in recognizing and using high quality data sources and also provided iterative writing practice. The instruction team also designed rubrics to evaluate and provide feedback on student work, and students had access to these rubrics before completing the assignments. Finally, the GSFs were available to consult with students on their writing. The project in Year 2 was also a staged assignment but culminated in an international news portfolio on one of several pre-identified topics. Students also completed a “pretest” news portfolio early in the semester, which the professor used to identify students’ weak spots and design additional scaffolding. This iteration also included peer review of other students’ portfolios. In Year 3, the professor and her team returned to the Op-Ed assignment because it required a broader range of valuable skills than the news portfolio, and they employed similar scaffolding and feedback strategies as in the previous semesters. They also encouraged students to submit their second Op-Ed in parts (voluntarily) and receive feedback on each submission.

In the psychology course, the major assignment was for students to write a mock advice column for parents based on their critical reading and synthesis of relevant empirical research papers. All three years of the project, students worked on the assignment through several stages that had been designed by the instruction team. The professor had assigned the advice column in previous course offerings, but in the Year 1 offering for this project she added a literature search lab session co-led and co-designed by a librarian and the GSFs, a peer workshop and analysis session, and additional in-class assignments modeling

the critical reading of empirical journal articles. Students also had access to the GSFs for one-on-one writing consultations. Based on evidence of student learning in Year 1 (and later Year 2), the instruction team designed additional revisions for the Year 2 and Year 3 offerings. For instance, they modified the timing and procedures involved in the writing consulting process, and developed assignments involving application of the paper rubric to sample papers. In Year 3, the students first wrote an academic-style literature review paper and then produced a much briefer advice column based on their literature review.

Assessment of critical thinking and writing skills

Development of disciplinary skill within the course. To measure the development of their students' cognitive skills across the semester, the Primary course faculty created assignments that required critical thinking and writing skills situated in course content. The instructors evaluated these skills with rubrics designating graduated levels of development on each of the major components of the assignments, such as use of research, synthesis of research, and writing mechanics. This assessment was done before and after the scaffolded teaching to see how much these skills were enhanced even within the course of a single semester.

In Year 1 of the political science class, students completed a critical review assignment early and towards the end of the semester. This assignment required them to analyze and critically evaluate an article in the field, skills which were central to the integrative Op-Ed assignment. In Years 2 and 3, the professor used the final integrative assignment (the news portfolio in Year 2 and the Op-Ed in Year 3) as the "posttest" and asked students to complete a similar assignment shortly before mid-semester as a baseline.

In the psychology class, students completed short writing assignments that were similar to the advice column assignment but did not require them to find and read original research. Rather, students were given descriptions of three empirical studies on a topic and were asked to use the research to write a letter to the editor or a blog post in response to a real-world question. Students completed the first essay as a baseline in the third week of the semester, and in Years 1 and 2 they completed a second essay at the end of the semester, after completing the advice column assignment. In Year 3, the advice column itself served this "posttest" function, but because this assignment also required students to select, read and summarize their own original sources, and four rather than three sources, it was more rigorous than the baseline assessment.

Measures of general skills through AAC&U Value rubrics. We also evaluated course assignments collected in all of the courses using the AAC&U Value rubrics for Written Communication and Critical Thinking skills (AACU, 2013). Each rubric identifies five dimensions of the overall skill (e.g., for Critical Thinking: Explanation of Issues, Conclusions), and describes performance at four graduated levels of skill development (1-Benchmark, 2- Milestone 1, 3- Milestone 2, and 4- Capstone) for each one. Students who do not meet even Benchmark levels of the skill can receive a 0 (Not Met). The details of these rubrics are displayed on the previously cited AACU website. Conversations within our project team, and with faculty across our campus for other assessment projects, have identified the Milestone 1 level of these rubrics as representing the minimum level of basic competence in a skill (for a university student), whereas Milestone 2 performance describes the level of skill mastery that most faculty would like to promote in their students.



We selected samples of 15-35 students from each Primary, Comparison, and Extension course (see Table 1), and evaluated their course-embedded writing assignments using these rubrics. For the team-designed classes, we analyzed the final integrative or capstone assignment that was designed to call upon the critical analysis and writing skills progressively developed across the semester. These assignments varied considerably across disciplines, but they all asked students to develop a written argument about an issue in the field and to support that argument using high quality sources and evidence. All were completed in the last third of the semester. In the Comparison classes, we selected assignments that also met these criteria but were not supported by team design or progressive and iterative assignments.

Beginning in Year 2, we trained a group of graduate students to score the written work using the Value rubrics. We began by convening a three-hour session of four graduate students (the raters), two faculty members, and two specialists from the libraries and writing center to practice and become reliable in using the rubrics. Participants in this session iteratively rated and discussed 5 or 6 assignments, making minor adjustments to the rubric language until they came to a shared understanding of the rubric categories and criteria. Each assignment was then independently scored on both the Written Communication and Critical Thinking rubrics by two different raters, with each rater paired with the three others on a comparable number of assignments. The raters then met again to discuss scoring disagreements and were permitted, but not compelled, to change their ratings following the discussion. In Year 2 we used this process to score the Year 1 assignments. At the conclusion of Year 3, we repeated this entire process, replacing two of the graduate student raters (who had graduated) to score the Year 2 and 3 assignments. Within all pairings, the raters were quite reliable with each other, providing scores that were identical or one category apart at least 90% of the time.

RESULTS

We conducted four analyses of students' writing and critical thinking skills. As a preliminary step, we first established that students in the Primary courses improved in their disciplinary skills over a single semester. Next, we tested the value added by the team design model by looking at whether the skill levels achieved by students in the Primary courses were higher than those achieved in comparable but conventionally designed courses. We then examined a year-by-year breakdown of the Value rubric data from one course, to illustrate the iterative course redesign process applied to the Primary courses. Finally, we used the Value rubric data from the Extension courses to assess whether the benefits of team design generalized beyond the Primary courses.

Development within a course

By comparing performance before and after the newly designed staged assignments, the instructors of the primary courses observed growth in students' writing and critical thinking skills from the beginning to the end of each semester. Table 2 shows the overall scores on the pretest and posttest assessments, converted to percentages, for each course and each project year. For both courses and in each year, students had improved scores from the early assessment to the end-of-semester assessment, but showed greater growth in some dimensions of these assignments than in others. For example, in the psychology course in all three years, students showed particularly notable improvements in *synthe-*

Table 2. Mean Scores on Early and End-of-Semester Assessments in the Primary Courses

	POLITICAL SCIENCE 170	PSYCHOLOGY 430
Year 1		
Early-Semester	.59	.72
End-of-Semester	.75	.86
Year 2		
Early-Semester	.83	.83
End-of-Semester	.86	.89
Year 3		
Early-Semester	.85	.85
End-of-Semester	.86	.89

sis, which involved integrating and drawing conclusions from diverse research findings. The overall improvements in psychology were statistically significant in Year 1, $t(80) = 8.28, (p = .0001)$, and Year 2, $t(53) = 3.29, (p = .0018)$, and marginally significant in Year 3, $t(49) = 1.87, (p = .067)$. Recall, however, that the Year 3 posttest was a more difficult assignment than the Year 3 pretest (and more difficult than the Year 1 and Year 2 posttests). The changes in subcomponents were also statistically significant, with $t_s \geq 2.99$ ($p_s \leq .0042$).

The political science instructors also noted greater growth in some areas than others. For instance, in Year 2, there were particularly strong improvements in the skills of *critical analysis* and *report organization*, and in Year 3 students showed notable growth in *writing skill*. The overall improvement was statistically significant in both Year 1, $t(207) = 12.01 (p = .0001)$, and Year 2, $t(153) = 2.51, (p = .013)$, but not in Year 3. Writing skill, however, did show statistically significant improvement in the Year 3 offering, $t(152) = 2.67, (p = .008)$.

Value rubric analyses of primary and comparison courses

To compare the quality of students' critical thinking and writing in the team-designed and traditional courses, we looked at the proportions of scores at each skill level (from Not Met to Capstone) on the Written Communication and Critical Thinking rubrics. We first looked at the scores on each rubric aggregated across project year and aggregated across the five categories of the rubric, which were highly correlated. Thus, these distributions represent the skills of the 259 students in these courses whose written work was evaluated with the Value rubrics (179 in the Primary courses and 80 in the Comparison courses).

The top panel of Figure 2 shows two separate distributions of Written Communication scores—one for the Comparison courses and one for the Primary courses. The Primary course distribution is shifted to the right, indicating that students in the team-designed courses produced higher level work in terms of writing skill. The Primary courses had fewer scores below minimum standards (i.e., at the Not Met or Benchmark levels), and more scores at the Milestone 1 and 2 levels, than the Comparison courses. The bottom panel of Figure 2 shows the two distributions of Critical Thinking scores. These scores were lower than Written Communication scores for both course types, but the Primary courses again had fewer scores than the Comparison courses at the bottom two levels,

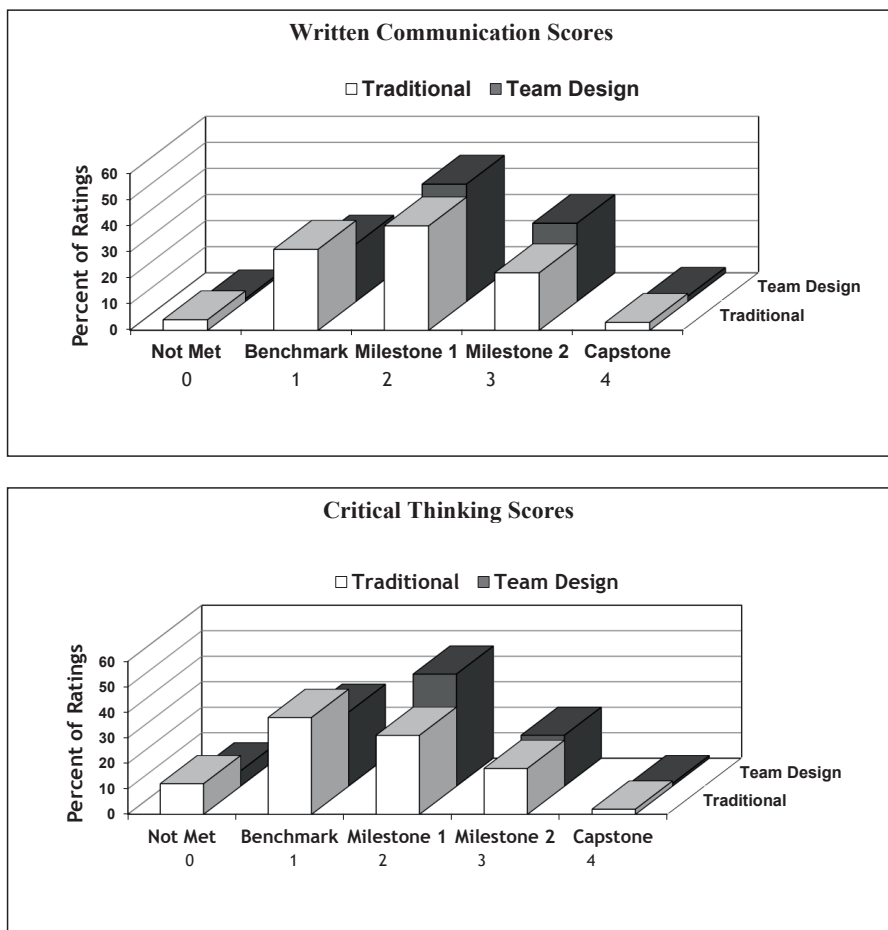


Figure 2. Written Communication and Critical Thinking Scores for Primary and Comparison Courses, Aggregated across Year

and had more scores at Milestone 1. The Primary courses substantially increased the rate of work that met or exceeded basic competence levels in both Written Communication (from 65% for Comparison courses to 77% for Primary courses) and Critical Thinking (from 51% for Comparison courses to 65% for Primary courses). Chi-square tests showed that the differences between the Primary and Comparison course distributions were statistically significant for both Written Communication, $N = 259$, $X^2(4) = 50.1$ ($p < .0001$) and Critical Thinking, $X^2(4) = 58.9$ ($p < .0001$).

To take a closer look at the specific kinds of skills the team-designed courses promoted, we analyzed the distributions of scores on the individual dimensions of writing and critical thinking skill specified by the rubrics. To illustrate, in Figure 3, we show the distributions for the Comparison courses and Primary courses in two skill categories. Figure 3a—in the top panel—displays scores on Selection of Sources and Evidence. In the Comparison classes, 35% of students used sources that were not credible or relevant, or used no sources at all. The Primary courses cut that percentage almost in half, and practi-

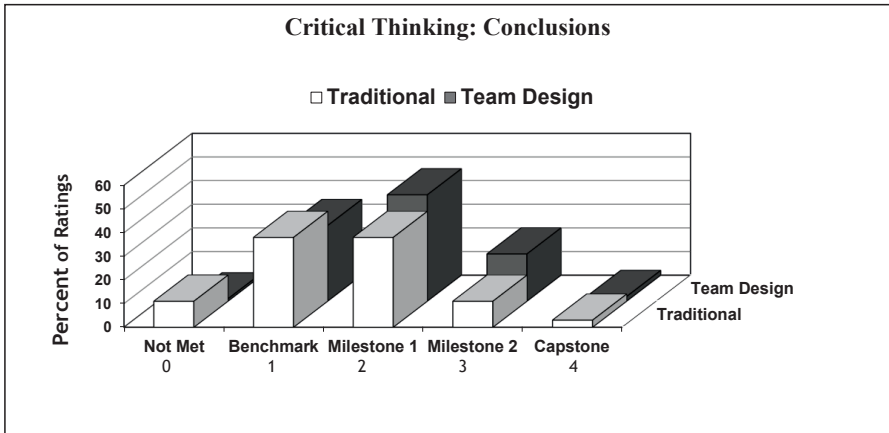
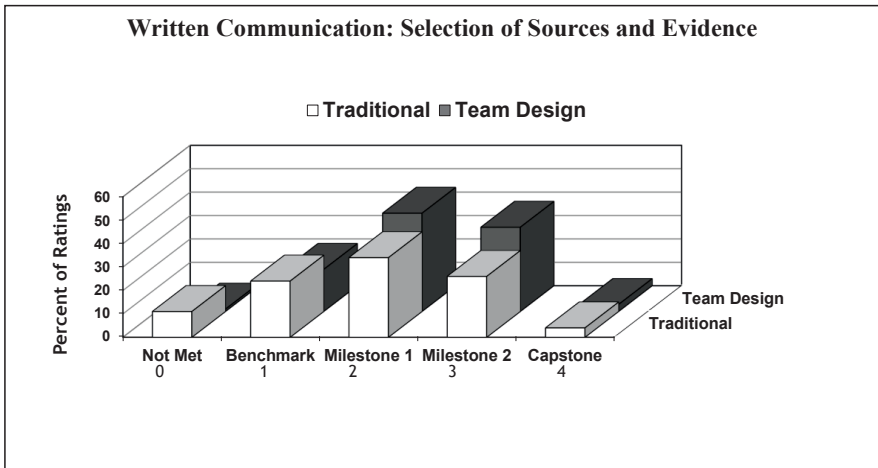


Figure 3. Value Rubric Scores on Individual Rubric Dimensions for Primary and Comparison Courses, Aggregated across Year

cally no students wrote papers without sources. Further, almost 40% of the Primary course students showed consistent use of high quality sources, whereas less than a third of the Comparison students did so. Thus, our team design model seemed to be quite effective in helping students learn to identify, access, and use high quality, relevant sources for their course assignments. Figure 3b—in the bottom panel—displays Conclusions scores. In the Comparison courses, about half of the students state conclusions that are not well tied to the evidence (38%) or provide no conclusions at all (11%). That percentage is reduced by 16% in the Primary courses, and almost all students in these courses provided some conclusions. Relative to the Comparison courses, more students in the Primary courses reached Milestone 1, drawing conclusions that were logically tied to the evidence, and twice as many students met or exceeded Milestone 2, indicating that they drew conclusions based on a wide range of information, including opposing viewpoints.

The overall pattern of upgraded student performance in the Primary courses was evi-

dent in seven of the individual skill categories; there were clear and statistically significant Primary course advantages in these categories ($X^2s(4) \geq 12.4$ [$ps \leq .015$]), but the advantage was especially dramatic for Genre and Disciplinary conventions and Selection of Sources and Evidence (from the Written Communication rubric) and for Explanation of Issues, Students' Position, and Conclusions (all from the Critical Thinking rubric). Thus, the team design model was particularly effective in enhancing students' abilities to find and use appropriate and relevant sources, follow writing conventions that are appropriate for the task and discipline, introduce a written argument and articulate a clear position within it, and draw conclusions that are based on the available evidence. In three categories (Content Development and Syntax/Mechanics from the Written Communication rubric or Evaluation of Sources and Evidence from the Critical Thinking rubric) there were no differences between the Primary and Comparison courses.

The results presented thus far reveal that, across the three years of the project, students in the Primary courses improved their critical thinking and writing skills across the semester, and by the end of the semester displayed more advanced skills than students in comparable but traditionally designed courses. Given that iterative course modification was a key feature of the team design work on the Primary courses, we next illustrate how evidence of student learning (i.e., performance on the pretest and posttest and on the final integrative assignment) in the Primary courses was used to produce continuous and iterative course improvement across semesters.

Iterative changes over time

Each year, the Primary course instructors examined measures of student learning to identify strengths and weaknesses in student performance. In collaboration with their instruction teams, they then designed additional assignment stages, learning activities, and/or opportunities for feedback that targeted weak skill areas. To illustrate this iterative process, we present rubric scores on a few individual rubric dimensions for the team-designed Psychology 430 class by project year alongside the Comparison psychology course data, aggregated across year, on the same dimensions.¹

Some skills showed consistent and steady improvement over time, commensurate with the team-designed course enhancements. To illustrate, students in the Psychology 430 showed more skilled use of Genre and Disciplinary Conventions by the Year 1 offering (see Figure 4) than students in Comparison courses. Over one third of students in the Comparison courses did not adhere to basic expectations for writing for the discipline or task, but this percentage was reduced by at least half in all three team-designed semesters. Each year of the project, moreover, the professor added new course elements to promote even more advanced writing skills. For instance, in Year 1, the students rarely took advantage of the optional writing consultations with the GSFs, so in Year 2 these sessions were scheduled to take place shortly after students received feedback on a draft, rather than as a preliminary step. The instruction team also developed assignments to reduce excessive quoting, focusing instead on how to explain empirical research in one's own words, and an assignment in which students applied the paper rubric to sample papers, to increase their understanding of high quality written work. Consistent with these revisions, use of important writing conventions was even stronger in Years 2 and 3, with the proportion of students who consistently adhered to conventions three to four times as high as in the Comparison classes.

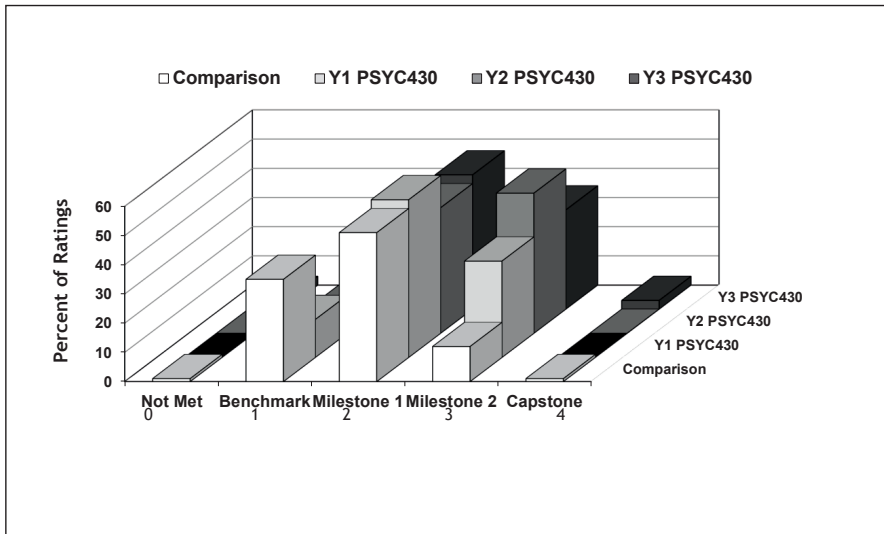


Figure 4. Change over time in Genre and Disciplinary Convention Scores for the team-design psychology course (PSYC 430), relative to Comparison psychology course

Other skill categories showed little enhancement in Year 1 but for this reason were targeted in course revisions in Years 2 and 3. For example, although students in the team-designed class showed large upgrades in their *selection* of sources and evidence, in Year 1 they engaged in less *critical evaluation* of sources and evidence than Comparison course students (see Figure 5a). Therefore, in Year 2 the Psychology 430 instruction team designed several new learning activities to provide guidance and practice in evaluating and drawing conclusions from empirical research. They also required students to add evaluative comments to the summaries of their sources that they produced for an in-class peer review session. Figure 5a shows a clear boost in evaluation of Sources and Evidence in Year 2, with 98% showing some interpretation or evaluation of sources and evidence (compared to 75% in the Comparison classes) and 59% questioning the viewpoints of experts. Nonetheless, Year 3 scores shifted back to levels similar to the Comparison courses. This pattern may reflect the reduced length of the advice column in Year 3; many students struggled to gauge the appropriate level of detail when writing about empirical evidence in this concise format for a general audience. This issue will likely guide future modifications of this course.

As shown in Figure 5b, there was no team design advantage in Conclusions scores in Year 1. Similar to the Comparison courses, half of the scores for the team-designed course were in the Not Met and Benchmark categories. Thus, even with the initial team-designed enhancements, many students in Year 1 drew conclusions that were not clearly evidence-based. Another third showed basic competence in evidence-based conclusions (Milestone 1), but only 14% considered the full range of information such as opposing viewpoints (the Milestone 2). This skill area was addressed by some of the Year 2 course enhancements described above: the new assignments on evaluating and drawing conclusions from research, and the rubric application activity, which emphasized identifying

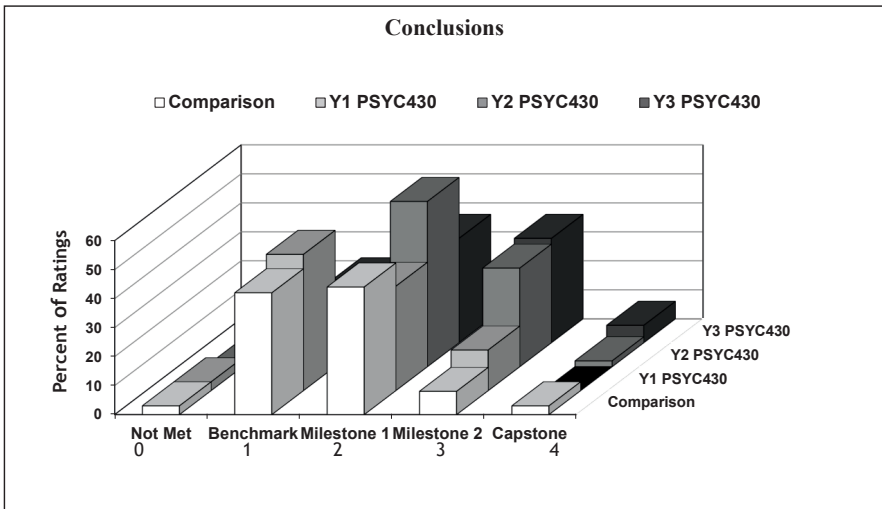
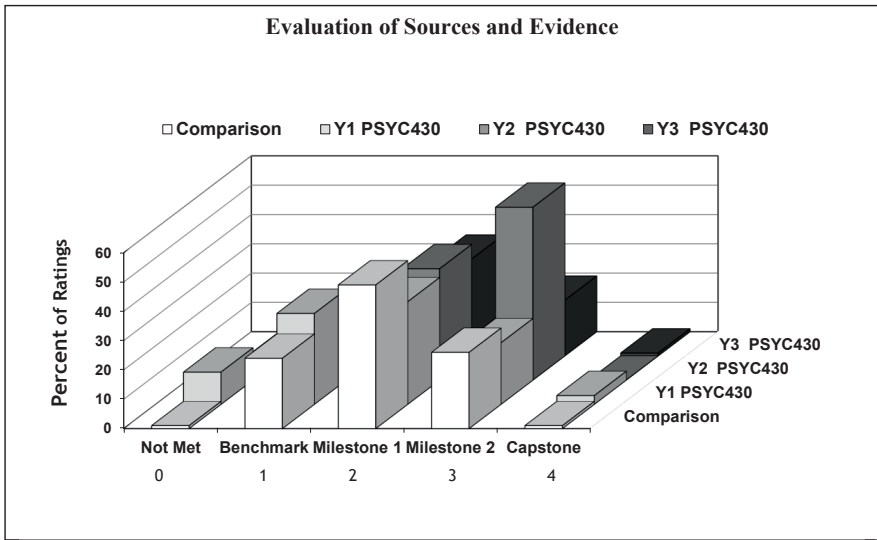


Figure 5. Change over time in two dimensions of Critical Thinking scores for the team-design psychology course (PSYC 430), relative to Comparison psychology course

high-end examples of the synthesis of multiple research findings. As shown in Figure 5b, almost all students drew evidence-based conclusions in Year 2, and the rate of advanced level Conclusions also increased, with over one-third of the students considering the full range of evidence. In Year 3, the professor made further modifications to support high quality research synthesis and conclusion: requiring students to write an academic-style literature review paper before producing a much briefer advice column, integrating more empirical journal articles into the regular course reading, and shifting in-class time from information delivery to the analysis and synthesis of information delivered via the read-

ing. Following these changes, 42% produced advanced level (i.e., Milestone 2 or Capstone) conclusions.

Student learning in extension courses

We also scored the assignments of 122 students in the Extension classes with the Value rubrics, to determine whether the upgraded skills observed in the Primary team-designed courses generalized to our extension courses (taught by participants in the faculty planning seminar). In Figure 6 we show the scores of the Extension courses, in comparison to those from the Primary courses and from the Comparison courses. Overall, students in the Extension courses displayed critical thinking and writing skills that were at least as advanced as those observed in the Primary courses, and that were more advanced than those in the Comparison courses. Written Communication scores in the Extension courses were almost identical to those in the Primary courses, and were significantly better than those in the Comparison courses, $X^2(4) = 38.37$ ($p < .0001$). Critical Thinking scores in the Extension courses actually exceeded those in the Primary courses, $X^2(4) = 23.85$ ($p < .0001$), as well as scores in the Comparison courses, $X^2(4) = 106.79$, ($p < .0001$). Thus, these results suggest that our planning seminar was quite effective in helping new cohorts apply the pedagogical model to their courses.

DISCUSSION

Limitations and advantages of the study methodology

The present work was done within eight courses that were regular offerings in the course schedule. There was no random assignment of students to classes, nor was the instructor or type of instruction randomly assigned to sections or semesters. The courses that were team designed all had innovative teaching and scaffolded assignments as well, while the comparison courses had regular assignments designed by the faculty member alone. A three-way factorial design would be required to separate the effects of team design from those of innovative assignment design or instructor quality *per se*, but that is not possible in the flow of course delivery in an institution (Neter, Wasserman, & Kutner, 1990). Further, it was not possible to obtain broad demographic information to match comparison and target courses; rather, they were matched roughly by department and level.

For these reasons, we cannot identify the effects of team design *per se*, independent of innovative teaching methods. We do note, however, that very few instructors create scaffolded assignments or cognitive apprenticeship activities on their own; the point of creating the teams was to enhance courses by adding the resources of design partners. Our findings simply make the claim that teams are a good way to create forms of instruction that may be beyond the time or expertise of a typical faculty member; we make no claim that individual faculty members cannot create innovative course designs. Our project demonstrates that team design yielded methods and results that were unlikely to occur in courses operated in typical fashion at our campus. If such innovative teaching is not present on a campus, creating design teams organized in this way could increase the likelihood of developing enhanced teaching on that campus.

Also, while a random sample of student work was taken from each class for evaluation and the raters did not know the treatment condition of the course, the raters did know

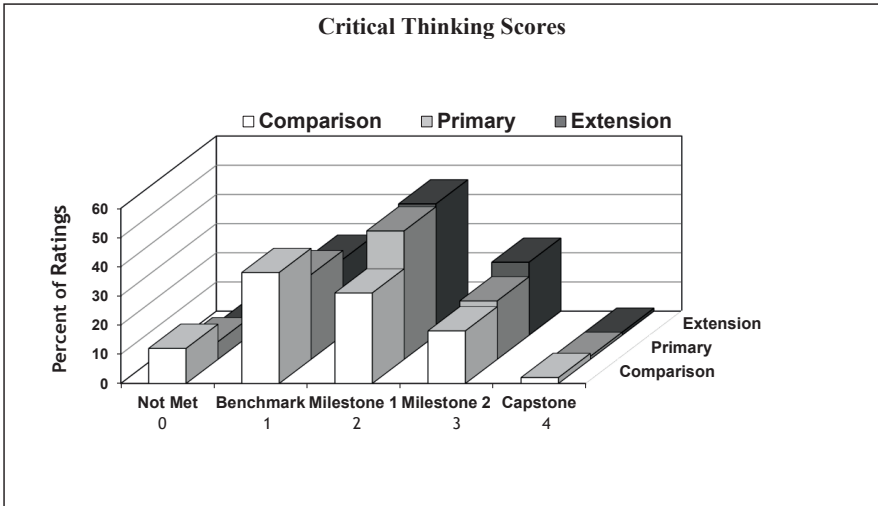
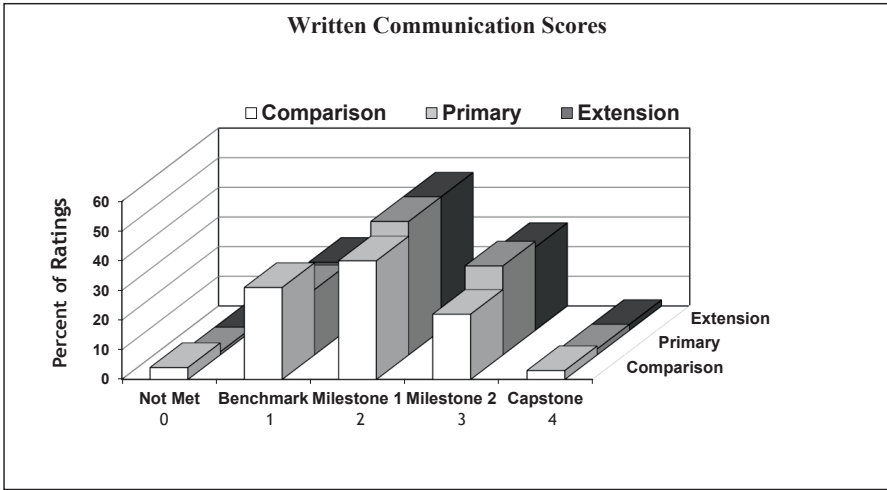


Figure 6. Written Communication and Critical Thinking Scores for Comparison, Primary, and Extension Courses, Aggregated across Year and Rubric Dimensions

the identity of the instructor of each course. Therefore we cannot rule out the possibility that the raters’ scores were influenced by pre-existing knowledge about the instructors. However, given that most of the graduate student raters were from different departments than the pool of instructors, we think this type of bias is unlikely.

We also note that our project has some advantages over well-controlled laboratory studies of learning. The participants were regular students taking a course under typical motivational conditions, and the program was applied to large classes operating within their normal resource constraints. The faculty members and teaching assistants involved had many other responsibilities besides this one course, and we can say that whatever

impact the program had would be sustainable under typical working conditions for instructors. The generalizability of this project may compensate for the imprecision of identification of whether team-designed instruction was the major responsible factor.

Value of instructional team design

Our assessments of student learning as well as participating faculty members' reflections suggest that team-designed courses using a cognitive apprenticeship approach can be an effective and efficient way of supporting the development of undergraduate students' critical thinking and writing skills, even in very large courses. It was very valuable to have staged assignments carefully designed to scaffold the desired skills, as actual implementation of this form of cognitive apprenticeship lived up to its conceptual billing. The intellectual resources the graduate students acquired through their fellowships with the writing center and library were critical to the design phase of the courses. Their fellowship preparation also made the graduate assistants even more effective in their direct work with students in their usual teaching assistant roles.

Figure 7 shows a schematic diagram of the different relationships among instructional partners in a conventional course and a team-designed course. Many faculty members use the version shown in the top diagram, in which specialists in writing and library research are visibly available to students to improve their performance after feedback, possibly on drafts. In a more progressive version, the specialists are invited to give a lecture to the class on their services or lead a one-session workshop before students begin their writing. In the fully realized version in the bottom of the figure, the specialists help design a sequence of assignments that allow students to develop the component skills of a complex task created by the instructor, receiving feedback and assistance at each step.

Our data suggest that this model works very well to enhance student performance even when classes are very large and have a wide range of students, including non-majors. Gathering the skills of many people is a very good way (and maybe the only way) to meet the challenge of large enrollment lower division courses, especially when faculty members have many responsibilities. Whether the time pressure comes from doing research, teaching four or five classes per semester, or substantial service and advising, many faculty members cannot engage in advanced teaching methods on their own. While seminars and some specialized upper division courses may not need as much input, all instruction will benefit from a larger vision of the skill set of instructional design.

We also note that this team design was focused explicitly on the disciplinary skills and goals of the instructors' complex assignments (Ianuzzi, 1998), rather than primarily aimed at producing information literacy or writing skill *per se*. This aspect of the project fits well with the strategies articulated in John Bean's *Engaging Ideas* (2011). Bean's book offers strategies for helping students achieve disciplinary goals, often through writing. That form of collaboration may be welcomed more by faculty members than when specialists approach faculty members asking for assistance in achieving the library or writing center's goals. Either way, the collaboration also helps develop general intellectual skills through the enhancement of disciplinary skill building.

Of course, when faculty members believe that a course and its content are a unique product resulting from a personal vision, they may find it hard to benefit from the advantages of multiple course designers. Therefore, we scaffolded the collaborative design work in the extension courses by creating faculty cohorts that participated in planning seminars.

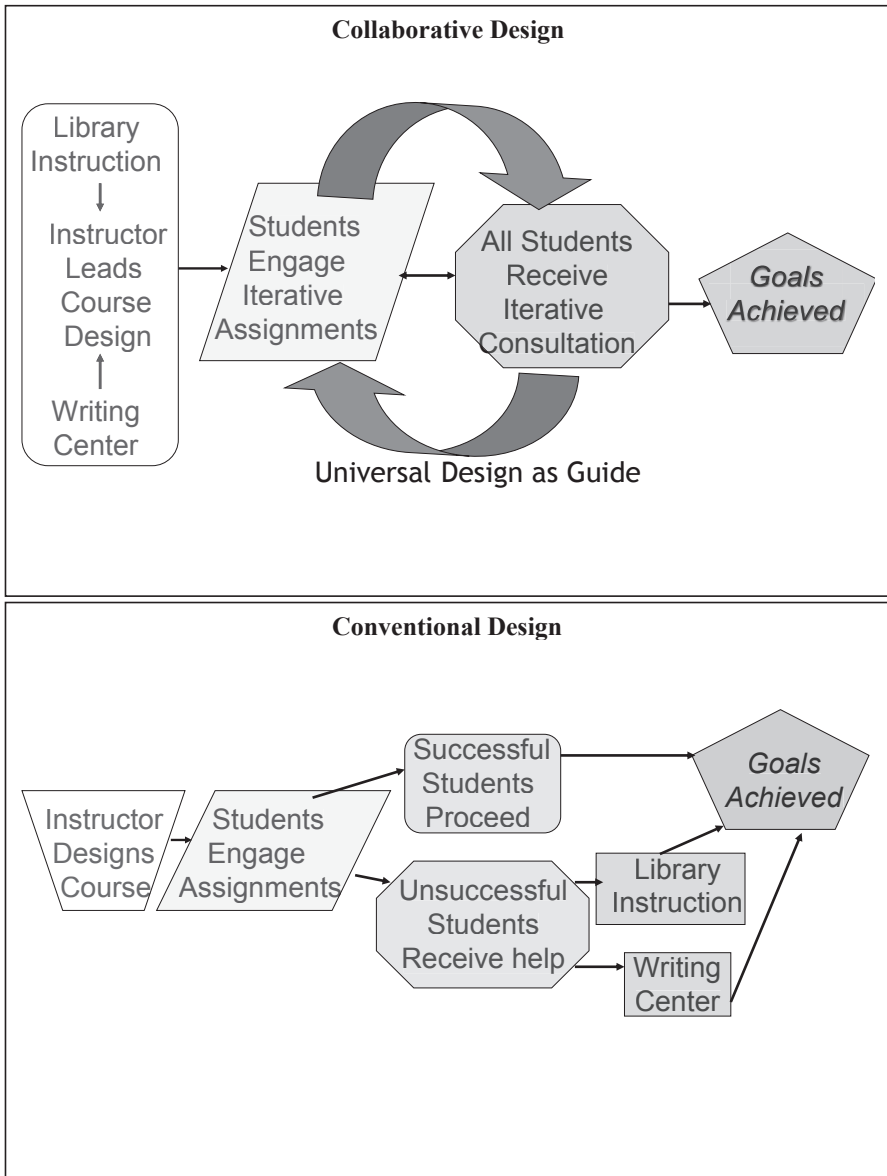


Figure 7. Traditional and collaborative course design models

These activities provided a view of collaborative design that helped faculty members move beyond their existing practices of solo course design and authorship.

The utilization of graduate students as course design partners and student consultants not only enhanced the course designs and support for students, but it also provided valuable experience that enhanced the graduate students' teaching repertoires. In addition to increasing their skills in assignment design and feedback, those GSFs were learning how

to collaborate on course design. These graduate students will begin their careers with direct experience of the benefits of team design and documentation of student learning. The add-on fellowship model we developed worked very well, and it is popular with both faculty mentors and the fellows themselves. Thus, engaging graduate students as colleagues in course design should be considered an effective strategy for both improving learning and preparing future faculty members. There is a delicate dance between the competing interests of dissertation research and teaching innovation, but we will definitely move toward offering more options on the teaching side. Given the likely career paths of most doctoral students, having strong preparation in course design will be a professional asset.

Value of identifying steps in promoting complex intellectual skill

Many educational researchers and theorists have focused on breaking complex tasks into manageable smaller steps, allowing students to learn more successfully. As documented in their comprehensive review of cognitive research on learning, Bransford, Brown, and Cocking (1999) note the advantages of cognitive apprenticeship approaches that build upon component skills. While this idea is not new, it can be very challenging to transform an entire course into a series of staged, sequenced assignments, and then even more difficult to implement such a plan in a large course. While not a perfect answer, having instructional partners who can provide the intellectual resources to accomplish much of that design is a great asset. The team-designed course is more likely to achieve in reality what many have been calling for in theory, and doing so in large classes is an even more critical need in contemporary higher education.

When there is sufficient energy and background in a design team, other advantages also accrue to the resulting teaching. With more specialist time available, it is easier to design assessment through authentically generated course assignments. While the goals of a course are often targeted at meta-cognitive skills like critical thinking, the particular assignments need to be situated in the context of the field being taught. A single instructor may have neither the time nor the experience needed to construct complex activities that give students a chance to demonstrate understanding in ways that go beyond simply answering questions. Another advantage is that instructors who become familiar with the use of rubrics (as in the AAC&U VALUE project) that describe development of component skills will be able to provide meaningful feedback to students, along with a conceptual roadmap of what they need to accomplish. Having specialists combine their varying expertise in course design provides an opportunity for exceptional teaching.

Institutional impact and faculty development

Producing change in teaching and learning practices within an institution requires sustained participation, collaboration, and support for participating faculty members. Drive-by faculty development is a good start (or a loss-leader for a teaching center), but we believe it is better to invest more resources and time in a smaller and more focused group of people. They will generate good examples that will bring others along. The initial attempt to spread the impact of this model at our university was the inclusion of six additional instructors and courses beyond the initial psychology and political science initiative. That provided a wider range of examples across disciplines, and it provided a broader test of the graduate student fellow model of making the work scalable to a larger



number of participants. Our data indicate that there were meaningful enhancements in student learning among those faculty members who had some modest but formal contact with the primary course instructors and their methods through the planning seminars. Adding the graduate student fellowship, while minor in cost, provided a benefit from participation that helped faculty colleagues justify the changes in their teaching.

The collection and rubric-based analysis of student work also had impact on our campus policies and practices in assessment. Largely based on our work (Greenhoot & Bernstein, 2011), the AAC&U scoring rubrics for writing and critical thinking that we used in the project were adopted as models for a university-wide assessment initiative. Both Profs. Greenhoot and Bernstein have been asked to take leadership roles in campus development of course redesign for enhanced completion and retention and of systems of measurement for writing, critical thinking, and recently for effects of the campus global awareness program. Further, the concept of team designed courses helped lay the groundwork for a new Center for Online and Distance Education and the development of a post-doctoral fellows program in support of large courses in science and mathematics. Based on the success of our project in collaborative design with libraries and the writing center, the idea was expanded to include instructional designers and e-learning specialists as partners with faculty members in developing online content and learning activities. We take this new and intense interest in our skills and strategies to be a good indicator that our work on team design of cognitive apprenticeship teaching was visible on campus and valuable to our colleagues.

The project affirmed our expectation that faculty colleagues respond well to highly-engaged and innovative teaching programs. When we combined a high quality program with excellent occasions for discussion with colleagues, good things happened, both in the immediate experience for participants and in the documented outcomes of the resulting instruction. We also saw that providing intellectual resources (a graduate student fellow as a real teaching partner) and logistical support (labor for handling and analyzing evidence) make a big difference in the willingness of colleagues to participate and in the quality of the resulting intellectual products. Neither of these is terribly surprising in retrospect, but as we move forward we have these ideas firmly in mind for planning purposes.

The next steps

We believe that teams are central to high quality course redesign in general. Our data cannot support a claim that team design is necessary to create innovative teaching and superior learning, but we can state that faculty members who embraced collaboration with instructional specialists demonstrated better learning results with typical undergraduates in large classes than was found by comparable faculty members who worked alone. With intense demands for both higher level learning and greater retention and completion of a wide range of students (e.g. AAC&U, 2002), faculty members are expected to adopt the most effective teaching methods possible, even in large classes. Attaining the forms of instruction found, for example, in universal design for learning requires both intellectual re-tooling and substantial time for revision and production of course materials. Given the work we have pioneered and documented through this project, our campus is moving toward team-designed courses as more of an expectation than an anomaly. It will be a challenge to sustain the educational benefits of redesigned courses while scaling up

from a handful of courses taught by volunteers to the significant number of foundational courses taught to large numbers of students each semester. We believe our experience and our data suggest that such investment is worthwhile and achievable.

ACKNOWLEDGEMENTS

This study was supported by grants from the Spencer Foundation and the Teagle Foundation. We are grateful to our colleagues Kate Weaver and Mariya Omelicheva for their work in developing the political science course as one of our primary targets. The authors of the Extension courses also contributed greatly during planning seminars. Jennifer Church-Duran and Erin Ellis of the KU Libraries and Terese Thonus, Director of the KU Writing Center contributed enormously to the development and delivery of this program. We also want to acknowledge the support of Bob Thompson, Duke University, whose guidance of the overall Spencer and Teagle projects was instrumental in our having the opportunity to lead this project.

Daniel Bernstein is a Professor of Psychology and the Director of the Center for Teaching Excellence at the University of Kansas.

Andrea Follmer Greenhoot is an Associate Professor of Psychology at the University of Kansas.

NOTE

1. We had no reason to believe that these courses would change systematically across project years because the comparison data were drawn from courses taught by different instructors.

REFERENCES

- Association of American Colleges and Universities (AAC&U) (2002). *Greater Expectations: A new vision for learning as a nation goes to college. National Panel Report available at <http://greaterexpectations.org/> (retrieved 17 July 2013).*
- Association of American Colleges and Universities (AAC&U) (2013). VALUE: Valid assessment of learning in undergraduate education. <http://www.aacu.org/value/index.cfm> (retrieved 30 June 2013).
- Bean, J. (2011). *Engaging ideas: The professor's guide to integrating writing, critical thinking and active learning in the classroom*. San Francisco: Jossey-Bass.
- Bransford, J.D., Brown, A.L., & Cocking, R.R. (1999). (Eds.) *How people learn: Mind, brain, experience and school (Expanded ed.)*. Washington, DC: National Academy Press.
- Collins, A., Brown, J.S., & Hollum, A. (1991). Cognitive apprenticeship: Making thinking visible. *American Educator*, 6(11), 38-46.
- Greenhoot, A.F., & Bernstein, D. (2011/2012). Using the VALUE rubric as a tool in evaluating a teaching innovation. *Peer Review*, Fall 2011/Winter 2012, 22-26.
- Hutchings, P. & Shulman, L.S (1999). The Scholarship of Teaching and Learning. *Change*, 31 (5), 11-15.
- Iannuzzi, P. (1998). Faculty development and information literacy: Establishing campus partnerships. *Reference Services Review*, 26, 97-102.

- Lave, J. & Wenger, E. (1990). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.
- McLeod, S.H. (Ed.). (1988). *Strengthening programs for writing across the curriculum*. San Francisco, CA: Jossey-Bass.
- McLeod, S.H. & Soven, M. (Eds.). (1992). *Writing across the curriculum: A guide to developing programs*. Thousand Oaks, CA: Sage.
- Miller, W. & Pellen, R.M. (Eds.). (2005). *Libraries beyond their institutions: Partnerships that work*. Binghamton, NY: Haworth Press.
- Neter, J., Wasserman, W., & Kutner, M. H. (1990). *Applied Linear Statistical Models*. Boston, MA: Irwin.
- Raspa, R. & Ward, D. (Eds.). (2000). *The collaborative imperative: Librarians and faculty working together in the information universe*. Chicago, IL: American Library Association.
- Rockman, I.F. (Ed.). (2004). *Integrating information literacy into the higher education curriculum: Practical models for transformation*. New York, NY: Wiley.
- Shulman, L.S. (2004). *Teaching as community property: Essays on higher education*. P. Hutchings (Ed.). San Francisco: Jossey-Bass.
- Wiggins, G., & McTighe, J. (2005). *Understanding by design (2nd Edition)*. Alexandria, VA: Association for Supervision and Curriculum Development.