

# Teaching an International Distributed Discussion-Based Course

Jeff Offutt<sup>1</sup>, Birgitta Lindström<sup>2</sup>, and Kesina Baral<sup>1</sup>

<sup>1</sup>Software Engineering, George Mason University, Fairfax, VA, USA

<sup>2</sup>School of Informatics, University of Skövde, Skövde, Sweden

**Abstract**—*Small discussion-based courses pose several challenges. Low enrollments make the course difficult to justify and can restrict active discussions. Impromptu discussions are hard to encourage. Students come to class tired, not well prepared, reluctant to speak out, or not able to verbalize abstract thoughts fast enough to fully engage. Sometimes a few students dominate the discussions while other students stay silent. This paper describes a novel teaching model that was created to allow one professor to teach the same course at multiple universities. As the course design emerged, the asynchronous online distributed nature of the course turned out to not only solve the initial problem, but also other challenges of discussion-based courses. Instructors and students found this model led to more engagement, increased learning, and higher performance.*

**Keywords:** Asynchronous online learning; Distance learning; Innovative uses of technology in the classroom; Collaborative learning

## 1. Introduction

The phrase “the web changes everything” has been widely used since the early days of the world wide web [1], [6], [13]. Change is exciting, full of potential for great good and great harm, and also hard to manage. Many domains and fields have been disrupted and are still trying to come to terms with the changes and whether they should be stopped, directed, managed, or accepted. Education is no different.

The printing press caused similar disruption to education, allowing a few teachers to disseminate knowledge to thousands of students [5]. Many teachers viewed printing as a threat, but over time recognized that books enabled them to teach more effectively. The web can easily be viewed as a threat, but like the printing press, can give effective teachers more tools to improve student learning.

Many types of courses have been taught online using different strategies, tools, and techniques [12], [2], [4]. Not surprisingly, results have been mixed. Some have reported unqualified success, some failure, but most attempts have resulted in mixed outcomes.

Means et al. studied more than a thousand empirical studies of online learning, finding that online students can perform just as well, if not better than, those taught in classrooms [9]. They also found that effectiveness varied with content and students [9]. This indicates that solutions

probably vary with the type of course, its contents, and the students.

This paper reports on multiple offerings of an innovative course taught in a primarily online asynchronous mode. The professors were not trying to push an online teaching agenda. In fact, their prior experience made them dubious of online education. Rather, they were trying to satisfy a specific need by starting with a fairly innocuous question: “*How can one professor teach the same course at three different universities?*”

The resulting course surprised us in several ways. Our general solution was to teach asynchronously online—out of necessity. Our goal was to maintain quality, or as medical doctors say, “do no harm.” To our surprise, the resulting course not only maintained quality, but was better.

Although our course is specific in topic, this strategy can be applied to other small discussion-based courses to solve several problems. Universities are reluctant to invest in low enrollment courses. Enrollment can be leveraged by distributing enrollment across multiple universities. Educators agree that peer-learning is effective [12], [11] by allowing students to bring different perspectives, knowledge-bases, and solution approaches. A distributed course increases diversity by integrating students from different educational environments. Online discussions can be more comfortable for students who hesitate to speak in groups or need more time to think before speaking.

The rest of this paper describes our novel solution to our initial educational problem (section 2), and our course in detail (section 3). We then present assessments in section 4, and close by discussing how this model of online education could be applied to other courses in section 5.

## 2. A Novel Solution

Offutt taught SWE 763, Software Engineering Experimentation, several times as a traditional, face-to-face (*F2F*) course that met once a week for 2.5 hours. The general goal was to prepare PhD students to carry out the empirical portion of their research. The course had three components: (1) three weeks of lectures on designing, conducting, and reporting on experiments; (2) nine weeks of in-class discussions of case study research papers; and (3) a semester-long research project that culminated in a paper and a conference-style presentation at the end of the semester. The course was successful in terms of student qualitative comments as well as quantitative outcomes such as published student papers.

Table 1: Stats on experimentation courses taught in a face-to-face classroom and online asynchronously

| Style  | Year | Enrollment |        | Paper publ. | PhD starts |
|--------|------|------------|--------|-------------|------------|
|        |      | US         | Sweden |             |            |
| F2F    | 1994 | 12         | -      | 4           | 3          |
|        | 1996 | 15         | -      | 3           | 1          |
|        | 2001 | 9          | -      | 3           | 1          |
|        | 2005 | 11         | -      | 3           | 3          |
|        | 2008 | 15         | -      | 4           | 3          |
| Online | 2012 | 9          | 6      | 10          | 3          |
|        | 2015 | 14         | 6      | 8           | 3          |
|        | 2019 | 5          | 9      | -           | -          |

Table 1 summarizes the enrollment and selected outcomes. Some student papers were submission-ready at the end of the semester, whereas others were revised or expanded before publication. Table 1 includes papers that were revised before being published. Several students used their class projects to start their PhD dissertation research.

Then in 2011, a research collaborator (Lindström) asked Offutt to teach his experimental course at Skövde University in Sweden. Although not practical, the request led to the above question, and eventually, to a new way to teach this course.

Teaching a class at multiple, international, universities presents practical problems. Even if delivered online, classes could not be synchronous because of time zones. Would the students be enrolled at GMU or their home institutions? USA students pay tuition directly to universities but tuition is paid by taxes in Sweden. What grading system should we use? Lastly, whose learning management system (LMS) should we use? It is difficult to enroll students into another university’s LMS, and the discussion board available at Mason would not support the kind of dynamic interaction needed for a discussion-based course.

The solution came in several parts. Students enroll at their home universities, we use the local grading systems, and we found a tool from a young startup company, piazza.com. The tool is a free discussion board with a modern, social networking style user interface. It is a neutral system where instructors can register students with an email address; thus it did not matter what university they attend. This tool allowed the students to be virtually merged into one class—exactly what was needed.

The first distributed version of the experimental course in Spring 2012 kept the previous structure, while being online and asynchronous. Lectures were recorded and all discussions were online. The only F2F meetings were the first meeting to set expectations and the final student presentations.

### 3. An Asynchronous Course

The first offering of the course in a distributed, online, asynchronous format (simply *distributed* in the rest of the

paper) was in 2012. Table 1 also provides numbers on the distributed versions of the course.

The learning objectives of the course remained the same from the F2F to the distributed version. As taken from the syllabus<sup>1</sup>, they are:

- 1) Understanding of the scientific process; particularly using the experimental method.
- 2) Knowledge of how empirical studies are carried out in computing and software engineering.
- 3) Understanding when experimentation is required in computing, and what kinds of questions can be answered using experimentation.
- 4) Knowledge of how to control variables and eliminate bias in experimentation.
- 5) Knowledge of how to design and carry out experiments in ways appropriate for a given problem, and develop skills in analyzing and presenting experimental data.

The overall course organization is the same in the distributed version. However, in the distributed version, the professor recorded and posted the lectures. To check engagement, students post answers to simple questions as responses to the lectures.

The first part of the course also includes papers about experimentation (six in 2019) that loosely correspond to the lectures. Students are required to post short summaries of half of those papers. We use these initial summaries to encourage the students to think deeply and apply critical thinking.

The case study section is the most different in the distributed course. In F2F courses, one student was assigned to present each paper in the classroom. Students presented 10-20 minute lectures. Other students listened, took notes, skimmed the paper, surfed the web, or slept. After presentations, other students engaged with questions, insightful comments, or otherwise tried to feign polite interest. These discussions would sometimes be dynamic and last 30 minutes; more often they would tail off after 5 or 10.

In the distributed version, two students for each paper write a critical summary and post to piazza. One student for each paper writes a “dissent.” The dissent focuses on limitations and problems with the paper, and (polite) disagreements with the initial summaries. Summaries are due on Monday (6:00 PM EST at George Mason and midnight in Sweden). Dissents are due Tuesday. Other students join the discussion through the remainder of the week with a nominal deadline of Friday.

We also intersperse project weeks throughout the semester. The project is to design, carry out, and write up an experiment. The topic and form (controlled experiment, qualitative human-based study, industrial field study, etc.) are chosen by students. As a fallback opportunity, we suggest

<sup>1</sup>Online: <https://cs.gmu.edu/~offutt/classes/763/> (some details have changed over time, this paper is based on the most recent offering in Spring 2019).

several projects. Early in the term students post their topic selection. Professors and students then ask questions, suggest refinements and alternatives, and otherwise help the students get started. A few weeks later, students post an initial experimental design. Again, the entire class evaluates the design looking for accidental biases, potential problems and risk factors, and possible additions and refinements. In other weeks we focus on writing and professional presentations, with recorded lectures by the professors.

Late in the semester, students post paper drafts for feedback and reviews. Each student is assigned to review two classmate's papers, just as in a research conference. Our goals are not just to help students with their own topic selections, experimental designs, and experiment, but to enable students to learn from their peers.

Our first surprise was the very first discussion of a published experimental paper. The initial summaries were more thoughtful and the discussion was longer and deeper than in the F2F format. More students joined the discussion. Although unexpected, we could immediately see why. Instead of coming to the classroom tired at the end of a day, and possibly not prepared, the asynchronous format allows the students to prepare at their convenience, consider other student's comments, and add to the discussion anytime. In the F2F format, a student may have a thought after class that would be lost forever. In the distributed format, that thought can be posted to enrich the entire class's learning. In short, the asynchronous format freed us from something we never knew restricted us—the tyranny of the clock!

Advantages were also clear during the week of project proposals. In the F2F format, each student spent five minutes presenting their experimental designs, and the professor gave feedback. In the distributed version, students posted 10 to 15 minutes worth of material online, and received up to four hours' worth of feedback. Some discussions continued for several days. These advantages, and others, are explored in more detail in the following sections.

## 4. Assessing Quality

Assessing the quality of online courses is notoriously difficult [3] and we were not able to find any agreed upon standards, checklists, or methods for evaluating course quality. Alley and Jansak suggest that it's not practical to develop a standard way to measure learning quality [3]. Thus a significant challenge was to identify quality factors that could be assessed in both environments.

We start with specific research questions, then investigate separate qualitative and quantitative measures.

**RQ1.** Do students report a better learning experience in a distributed or F2F course?

**RQ2.** Do professors report better results in a distributed or F2F course?

**RQ3.** Do students perform better in a distributed or F2F course?

**RQ4.** Do students give higher ratings in a distributed or F2F course?

**RQ5.** Was there more discussion in a distributed or F2F course?

### 4.1 Qualitative evaluations

This section compares distributed offerings of this course with the previous F2F offerings. We were pleasantly surprised to find that instead of merely maintaining quality, by several measures, the distributed offerings of this course were better than the F2F offerings. This section considers comments from the students (RQ1), instructor assessments of the quality of the projects (RQ2), and the level of engagement by the students (RQ5).

GMU obtains anonymous student comments for each course through a standardized evaluation form. Student affective or *felt* outcomes that emphasize student evaluations are widely used [8]. Studies show that student ratings often agree with other measures such as learning and expert evaluations [7].

In F2F versions of the course, students often commented on the difficulty of staying engaged (“*no concentration during class*,” “*can't keep up with discussions*”). They also expressed disappointment in the amount of feedback on experimental designs (“*the professor didn't get it*,” “*other students were not interested*”). In the distributed versions, however, comments emphasized the quantity and quality of the feedback (“*professor and students*,” “*completely re-designed my project*”). Another interesting finding is that Swedish students often recommend it to students outside the University. PhD students from six different Swedish universities have so far enrolled in the course at University of Skövde, which is unique for the University.

Instructors are also positive about the course (RQ2). It is worth emphasizing that neither instructor was trying to promote an online learning agenda. We found that it was easier to assess individual performance in the online version of this course. The major components are the semester-long projects that are unique to the individual, and participation in the paper discussions. In the face-to-face versions, it was hard to measure participation because the discussions were free-form and impromptu, making it hard to measure who said what. In the online asynchronous course, all comments are recorded and available for later review.

Both by spreading our effort throughout the week and by sharing the load, the online course took less effort and was much less stressful. Preparing for a 2.5 hour course and focusing for the entire class meeting can be exhausting. The online version of the course relieved us from the “tyranny of the clock.”

The asynchronous and discussion-driven nature of the course is more rewarding for the instructors. We have time to reflect and to formulate our answers and feedback more carefully than in a classroom or an online **synchronous**

Table 2: Course projects that resulted in published papers

| Course type | Students | Published | % published |
|-------------|----------|-----------|-------------|
| F2F         | 62       | 14        | 22.5%       |
| Distributed | 35       | 18        | 51.4%       |

seminar. Not only does this mean that students get better answers and more helpful feedback, it is more satisfying for instructors. Moreover, instructors with different background knowledge can contribute with more information when needed. By doing so, the instructors get to learn more.

The instructors have collectively more than 55 years of experience with F2F courses and the level of engagement by the students that we see in the distributed courses is much higher than we have ever seen in F2F courses. The discussions are longer and deeper, showing much more insights and reflections. The students challenge the instructors and force them to reflect and learn. Another sign of engagement is the throughput we see among the Swedish students. Swedish students can only get a pass or fail and have to get a pass for all parts of the course to get a grade. It is therefore common in Swedish F2F courses for student to fail or resubmit assignments long after the course finishes. Unusually, all Swedish students have successfully finished the distributed course on time.

## 4.2 Quantitative assessments

Quantitative metrics include the number of course projects that resulted in published papers (RQ3), assessed grades of the student projects (RQ3), the time spent discussing research papers (RQ5), and standardized student course evaluations (RQ4).

Although not a requirement or a specific learning goal, students do real research and we hope that some of their projects will be published. At the end of the semester, part of our assessment is what it would take to make the paper publishable. Our assessments include submission-ready as is (very rare), submission-ready after revision, needs more analysis for publication, needs more subjects or data for publication, and not publishable. The raw numbers about publications were already given in Table 1. Table 2 summarizes the number of papers published, possibly after revision, omitting the 2019 edition of the course as it is still in-progress.

Although the small numbers make statistical testing problematic at best, the difference is remarkable. Less than a quarter of projects in the F2F versions of the course led to published papers, whereas slightly more than half of the projects in the distributed version did. This was unexpected, and it is impossible to have the kind of experimental controls in this reflective study to definitively answer why. But we have some theories. As documented below, discussions of research papers were richer and deeper, and we suspect this means students learned more. Because professors and

students can reflect and spread the work throughout the week, they gave much more feedback on their experimental design. Indeed, several of the experimental designs and draft papers had significant feedback from several classmates, which did not happen in the face-to-face courses.

Our next quantitative assessment is based on the grades assigned to the projects. Universities in Sweden do not assign letter grades to graduate courses, so this measure only includes George Mason students. Table 3 summarizes aggregate grades from the F2F versions and the distributed versions of the course. Although grades in the distributed versions of the course were higher than in the F2F versions, the differences are small. In addition, grading research papers is inherently subjective. Thus, we feel that we can conclude at most that the students did not do worse in the distributed versions of the course (our original goal).

We next want to compare the discussions. This requires comparing oral discussions, which are naturally measured by time, with written discussions, which are naturally measured by word count. Speaking coaches claim that people speak about 150 words per minute (*wpm*) when reading aloud, between 125 and 150 wpm when giving a rehearsed speech, and about 110 wpm during conversation [10]. We were not able to find published estimates of wpm during extemporaneous or impromptu discussions on technical matters. In these situations, speakers are struggling to turn complex and abstract ideas into words, so they may speak slower.

To approximate this, we counted words during six technical (computer science) seminars. We counted 37 one-minute intervals during the presentations, and then 21 questions and answers. Many of these were less than one minute. In this sample, speakers averaged 85 wpm during the presentations and speakers and audience averaged 75 wpm during the Q&A. Therefore, we assert that technical discussions can be estimated to occur at a rate between 75 wpm (Q&A) and 110 wpm (conversational).

Offutt attempted to quantify participation of students in the 2008 F2F version of the course by having a GTA record the number of minutes used by each student when initially presenting each paper, and the number of minutes used during discussion. We also counted the number of words in summaries, dissents, and responses on Piazza during the 2012, 2015, and first half of the 2019 course. Table 4 summarizes these numbers. We added dissents during the 2012 course, so they only appear for the online versions.

In addition, the F2F version averaged six responses per paper, whereas the online versions average 13 responses per paper. Although the wordcounts are necessarily rough calculations, Table 4 clearly shows that the online version of this course leads to more participation. Even at the most pessimistic 110 wpm (conversational) measure, the online version had almost twice as much discussion. We believe the 85 wpm measure is more accurate.

Finally, the course evaluations at George Mason include

Table 3: Grades for F2F and distributed course offerings

| Year                       | Students | A     | A-    | B+   | B     | B-   |
|----------------------------|----------|-------|-------|------|-------|------|
| 1994                       | 12       | 8     | 2     | 0    | 2     | 0    |
| 1996                       | 15       | 8     | 0     | 0    | 6     | 1    |
| 2001                       | 9        | 5     | 2     | 0    | 2     | 0    |
| 2005                       | 11       | 8     | 1     | 1    | 0     | 0    |
| 2008                       | 15       | 12    | 3     | 0    | 0     | 0    |
| 2012                       | 9        | 7     | 2     | 0    | 0     | 0    |
| 2015                       | 14       | 10    | 4     | 0    | 0     | 0    |
| <b>Total F2F</b>           | 62       | 41    | 8     | 1    | 10    | 1    |
| <b>Total distributed</b>   | 23       | 17    | 6     | 0    | 0     | 0    |
| <b>Total F2F %</b>         | 62       | 66.1% | 12.9% | 1.6% | 16.1% | 1.6% |
| <b>Total distributed %</b> | 23       | 73.9% | 26.1% | 0%   | 0%    | 0%   |

Table 4: Words used during technical paper discussions. These are averaged over the papers discussed in the class.

|               | Avg words per summary | Avg words per dissent | Avg words all discussions | Avg total number words |
|---------------|-----------------------|-----------------------|---------------------------|------------------------|
| Online        | 1235                  | 618                   | 2367.0                    | <b>4220.0</b>          |
| F2F (minutes) | 12                    | NA                    | 8.5                       | 20.5                   |
| F2F, 75 wpm   | 900                   | —                     | 637.5                     | <b>1537.5</b>          |
| F2F, 85 wpm   | 1020                  | —                     | 722.5                     | <b>1742.5</b>          |
| F2F, 110 wpm  | 1320                  | —                     | 935.0                     | <b>2255.0</b>          |

questions that are scored from 1 (poor) to 5 (excellent). Of particular interest is the question “*my overall rating of the course.*” In the F2F versions, the average rating on this question ranged from a low of 4.25 to a high of 4.52. In the distributed versions of the course, the average rating on this question was 4.83 (2011) and 4.87 (2015). This indicates a clear preference for the distributed version.

## 5. Conclusions

Much of what we found in this course flows from a few key differences. In the traditional F2F versions, students often came to the classroom tired at the end of a work-day. In the distributed version, students **choose** when to participate throughout the week. The F2F version requires extemporaneous and impromptu participation, whereas the asynchronous version allows time for reflective participation. The F2F version limited feedback to a 2.5 hour session, whereas the distributed version allows extended discussion through several days or even weeks. In the F2F version, it was difficult to track who joins discussions and who makes substantial contributions, whereas the distributed version allows discussions to be recorded and counted later.

These differences lead to substantial benefits, as documented above. The basic course structure (in both settings) emphasizes creative problem solving and divergent thinking and encourages peer-learning. The dissents help students see multiple points of view, leading to more thoughtful reflections on the papers. The asynchronous course gave time to reflect, which leads to deeper and more interesting

discussions. Online discussions are not limited by time, thus nobody can dominate the discussions. Some students are more expressive and outgoing when typing online than when speaking in a group setting. The extended time allows for deeper and more meaningful comments. An interesting side-effect is that students can easily compare their participation with other students and then self-adjust their participation level.

We developed several strategies while teaching this course. First, we let students bid on the papers they want to summarize at the beginning of the course, then do our best to assign papers accordingly. Not only do students appreciate picking the papers they summarize, we believe it enables them to write better and deeper summaries. We also found that setting weekly expectations is very important, and we coordinate with weekly instructor meetings. It is also essential that instructors stay engaged and that the students can see that engagement. Early in the semester, much of our feedback is positive reinforcement. This establishes trust and creates a supportive environment, which is harder to do in online courses. Then we gradually shift to giving more constructive feedback and criticism to encourage them to learn more and do more. When we see students not participating, we send intervention emails asking if they need help or if there is a problem. We also use synchronous office hours, either in person or online, when we need high bandwidth discussions.

We also have a few tactics to share. We learned early on to have one piazza note per paper, and then have the

summaries, dissents, and all discussions as responses within that note. We also require short responses to early papers and the online lectures to ensure students read and watched. After the first offering, we archived key emails and notes to reuse in future offerings

We believe this format can be successfully adapted to any discussion class, whether in science, engineering, or humanities. The discussions and diversity of projects allowed for divergent thinking and substantial peer learning. Moreover, the multi-university format offers several interesting possibilities. If there aren't enough students to support a specialized class at one university, they could join the same class at another university, even on another continent. If a class is taught by someone with unique expertise, students elsewhere could join. This format could scale from 20 to 200 students if professors at other universities participated in the feedback and evaluations. This opens the possibility of *crowd teaching*, where a diverse team of professors from 5 or even 50 universities share a collaborative space to increase peer-learning and allow professors with different skill-sets to support each other to better educate all students.

## References

- [1] Eytan Adar, Jaime Teevan, Susan T. Dumais, and Jonathan L. Elsas. The Web changes everything: Understanding the dynamics of web content. In *Proceedings of the Second ACM International Conference on Web Search and Data Mining*, pages 282–291, Barcelona, Spain, February 2009.
- [2] Elaine Allen and Jeff Seaman. Going the distance: Online education in the United States. Technical report ED529948, Babson Survey Research Group, November 2011.
- [3] Lee R. Alley and Kathryn E. Jansak. The ten keys to quality assurance and assessment in online learning. *Journal of Interactive Instruction Development*, 13(3):3–18, 2001.
- [4] Jaclyn Broadbent and Walter L. Poon. Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education*, 27:1–13, October 2015.
- [5] Richard DeMillo. *Abelard to Apple: The Fate of American Colleges and Universities*. The MIT Press, Cambridge, MA, USA, 2011.
- [6] Caitlin Dewey. 36 ways the Web has changed us. Online, March 2014. <https://www.washingtonpost.com/news/arts-and-entertainment/wp/2014/03/12/36-ways-the-web-has-changed-us/>, last access February 2019.
- [7] James A. Kulik. Student ratings: Validity, utility and controversy. *New Directions for Institutional Research*, 109:9–25, January 2002.
- [8] Alf Lizzio, Keithia Wilson, and Roland Simons. University students' perceptions of the learning environment and academic outcomes: Implications for theory and practice. *Studies in Higher Education*, 27(1):28, 2002.
- [9] Barbara Means, Yukie Toyama, Robert Murphy, Marianne Bakia, and Karla Jones. Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies. Technical report, U.S. Department of Education Office of Planning, Evaluation, and Policy Development Policy and Program Study Service, September 2010. <http://www2.ed.gov/rschstat/eval/tech/evidence-based-practices/finalreport.pdf>.
- [10] Peggy Noonan. *On Speaking Well*. William Morrow Paperbacks, 1999.
- [11] Namin Shin. Transactional presence as a critical predictor of success in distance learning. *Distance Education*, 24(1):69–86, 2003.
- [12] Karen Swan. Building learning communities in online courses: The importance of interaction. *Education, Communication & Information*, 2(1):23–49, 2002.
- [13] James Teevan. The Web changes everything. Online Video lecture, July 2011. [http://videlectures.net/acmwebsci2011\\_teevan\\_changes/](http://videlectures.net/acmwebsci2011_teevan_changes/), last access February 2019.