2023 Annual Conference & Exposition
Baltimore Convention Center, MD | June 25 - 28, 2023
The Harbor of Engineering Education for 130 Years
Paper ID #38487

Board 254: Developing Tools, Pedagogies, and Policies for Computer-Based Collaborative Learning Activities

Morgan M. Fong, University of Illinois, Urbana - Champaign

Morgan is a PhD student in the Department of Computer Science at the University of Illinois at Urbana-Champaign. Her current research focuses on developing methods and analyzing cooperative learning in undergraduate computing courses.

Liia Butler, University of Illinois, Urbana - Champaign
Dr. Abdussalam Alawini, University of Illinois, Urbana - Champaign

Dr. Abdussalam Alawini received a doctoral degree in Computer Science from Portland State in 2016. In his Ph.D., he built systems to help scientists manage their file-based datasets by predicting relationships among spreadsheet documents. Passionate about a career in academia, Dr. Alawini joined the University of Pennsylvania in 2016 as a postdoctoral researcher. As a postdoc, he developed data citation and data provenance systems for scientists. Dr. Alawini's research interests are broadly in databases, applied machine learning, and education. He is particularly interested in applying machine learning methods to improve classroom experience and education in general. He is also interested in building next-generation data management systems, including data provenance, citation, and scientific management systems.

Dr. Geoffrey L. Herman, University of Illinois, Urbana - Champaign

Dr. Geoffrey L. Herman is the Severns Teaching Associate Professor with the Department of Computer Science at the University of Illinois at Urbana-Champaign.

Prof. Mariana Silva, University of Illinois, Urbana - Champaign

Mariana Silva is a Teaching Associate Professor in the Department of Computer Science at the University of Illinois at Urbana-Champaign. Silva is known for her teaching innovations and educational studies in large-scale assessments and collaborative learning. She has participated in two major overhauls of large courses in the College of Engineering: she played a key role in the re-structure of the three Mechanics courses in the Mechanical Science and Engineering Department, and the creation of the new computational-based linear algebra course, which was fully launched in Summer 2021. Silva research focuses on the use of web-tools for class collaborative activities, and on the development of online learning and assessment tools. Silva is passionate about teaching and improving the classroom experience for both students and instructors.

Developing tools, pedagogies, and policies for computer-based collaborative learning activities

Abstract

Collaborative learning can improve student learning, student persistence, and the classroom climate. Most of the evidence-based practices for collaborative learning rely on the assumption of face-to-face interactions or asynchronous online activities. In this paper we summarize the milestones, lessons learned, and preliminary research findings for the NSF IUSE project award #2121412 titled "Enhancing Equity and Access Via Digitally-mediated Collaborative Learning Experiences". As part of this project, we have developed tools and pedagogies for synchronous computer-supported collaborative learning activities that can be used in online and in-person classes. More specifically, we will describe 1) the computer-based tools that facilitate group assignments and distribution of tasks; 2) how the tool has been adopted by several courses in different institutions and 3) how it has impacted students' sense of belonging.

1 Introduction

Collaborative learning is an evidence-based instructional practice that has been adopted by many instructors in STEM courses in higher education. Research indicates that engaging students in collaborative activities is associated with increased student persistence, and improves student learning outcomes and retention [1, 2]. Successful and productive collaborations are rarely guaranteed, however they can be greatly improved by a careful design of the task [3] and the use of available technologies, to both promote collaborations among students and support the instructors implementing these activities [4, 5, 6, 7]. However, most of the evidence-based practices for collaborative learning rely on the assumption of face-to-face interactions [5, 6] or asynchronous online activities [8, 9]. Robust tools and pedagogies are lacking for synchronous computer-supported collaborative learning activities that can be offered online, in-person or in a hybrid format.

This project aims to develop technology to support collaborative learning during computer-based activities, and investigate how these tools improve the classroom experience and student's achievement of learning goals.

Computer-based assessment systems such as Gradescope [10] and Canvas [11] have built-in features to create group assessments, but these are not synchronized for real time communication. Other computer-based tools such as CoCalc [12], Google Colab [13] and Google Docs [14] offer the ability of co-editing in real time, but lack the framework for instant grading and feedback.

This research team is not aware of an online tool that combines these collaborative features. In addition, the use of commercial packages limits how instructors can create their assessments, and often does not provide sufficient access to log files that can be used to understand classroom interactions and potential educational research.

PrairieLearn [15, 16] is an open-source online assessment system that encourages students to master content. It was originally developed at the University of Illinois, Urbana-Champaign in 2015, and it is currently adopted in over 100 courses, impacting more than 15,000 students. Since 2020, several other universities started using PrairieLearn, including University of California Berkeley, University of British Columbia, University of Maryland, Michigan State University, University of Michigan and Grand Valley State University. PrairieLearn permits the authoring of question generators, each of which is capable of generating a range of parameterized question instances. It permits a broad range of question types, including but not limited to numeric, graphical, symbolic, programming, and drawing problems. Question instances are automatically graded, and students receive immediate feedback. Instructors have complete control of how they create their assessments, and have direct access to log files that can be used to understand classroom interactions, assessment quality and student outcomes.

One of the goals of this project is to incorporate into PrairieLearn the ability to create collaborative assessments that can be used by students for real-time activities. We want to use the developed group feature to better understand 1) how the scheduling of synchronous computer-supported collaborative learning activities affect access and equity and 2) how we can adapt face-to-face facilitation strategies for use in online environments, especially to enhance equity and access.

2 Tool development

2.1 Auto-graded group assessments

With the implementation of shared assessments, students in the same group share the same assessment view and submission history, and receive the same score. Instructors can form groups selecting from three options: 1) upload a roster of groups with corresponding students; 2) PrairieLearn randomly generates the groups; 3) students create their own groups. Figure 1 illustrates how students can form their own group: one team member decides on a group name, and once they click "Create new group", they will receive a join code that can be shared with other students. Instructors can decide on the minimum and maximum number of students in each group.

2.2 Scaffolded assessments

Figure 2 shows the main page for an example group assessment introducing Random Walk, which consists of three questions, each one increasing in level of difficulty, and relying on the understanding of the previous solution. This scaffolding of content is typical in Process Oriented Guided Inquiry Learning (POGIL)[17] activities, a group-learning instructional strategy adopted by the instructors involved in this project. To unlock the second question, students need to achieve completeness of the first question, here defined by a score of 100%. This locking mechanism

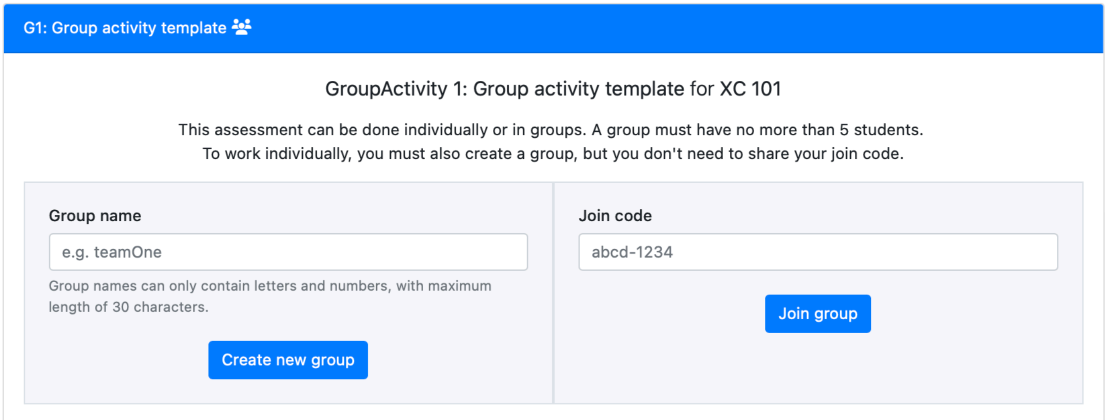


Figure 1: Page where students can create their own groups: one team member decides on a group name, and then shares the join code with others. Instructors decide on the minimum and maximum number of students. In this example, students can complete the assessment in groups of 1 to 5 students (meaning, it can also be completed individually.

prevents students from splitting tasks to work constructively (also known as divide-and-conquer) and instead work collaboratively. This locking feature can be turned on or off, depending on the choice of the instructor.

Question	Value	History	Awarded points
Penny-shaving scheme			
GA 4.1. Random Walk introduction	1		— /1
Modeling the stock market			
GA 4.2. Stock Market - simple model	2		— /2
GA 4.3. Stock Market - Black-Scholes model	7		— /7

Figure 2: Assessment page showing how questions (parts of the assessment) are locked, depending on completion of previous questions. This locking mechanism prevents students from working on different parts of the group assessment.

2.3 Real-time synchronization

The group shared assessments are currently "synced" among all members of the group upon a submission attempt, i.e., after one of the students submit an answer, all the students can view the submitted answer, the submission history, and their shared scores. However, the question interface is currently not yet synced in real-time, thus students are not able to view in real-time if another student is typing an answer (this is a work-in-progress).

However, the real-time synchronization is already available inside PrairieLearn workspaces, which are persistent remote containers via in-browser frontends such as VS Code [18] and Jupyter Lab [19]. As such, students completing group assignments delivered via VS Code or Jupyter Lab can have similar experience as working on a Google Docs, with the addition of auto-grading features, and instant feedback.

2.4 Structured roles

Prior research in face-to-face settings has highlighted the many ways that equity can be improved, including the use of role scripting (such as POGIL) and visual tools to monitor groups' progress and work distribution [17, 20, 21, 22, 23]. Due to the pandemic, more recent studies have explored how these face-to-face strategies should be translated into a web-based environment [24, 25].

We have implemented collaborative learning activities in our courses using the group features in PrairieLearn while encouraging students to use the POGIL roles of Recorder, Manager, and Reflector. The Recorder is the main "driver" who enters most of the answers in PrairieLearn. The Reflector completes a survey at the end of each activity, reflecting on the group's interaction and how the activity itself helped their learning. The Manager coordinates team's efforts, making sure everyone is contributing and following along. Currently these roles are encouraged, but not enforced by the system. Members of each group are required to alternate in these roles such that every student participates in at least one of the roles by the end of a term (this is currently achieved by giving a grade for participation in each role). Section 4.1 presents some of our findings from the use of structured roles in computer-based collaborative learning activities.

We are finalizing the implementation of role assignment within the PrairieLearn platform to give instructors the ability to enforce roles with their corresponding actions. For example, a instructor may enable "Edit" permissions only for the Recorder, while all the other members will have "Viewer" permissions. The Reflector may be the only member with "Viewer" permissions to a reflection survey at the end of the assessment. Figure 3 illustrates how a team of three students can self-select their roles. Instructors can define which roles are available, which ones are required and what are the actions for each role. In this example, the Manager, Recorder and Reflector are required roles, so if a team has only two students, one of them will need to take two roles. The system will perform all the validations before allowing students to start the assessment.

2.5 Group Insights Dashboard

We are developing a Group Insights dashboard to help empower instructors in better understanding how equitable group collaborations are among students in a course. PrairieLearn already has statistics and other information readily available at the assessment-level that can be used to gain insight into student group dynamics. With the new Group Insights dashboard, instructors will have access to metrics that characterize how well students are collaborating across several group assessments (i.e., providing insight into how well student groups appear to be working as a whole, rather than per assessment). Initial features being implemented include displaying whether or not students are working with their assigned groups over a set of group work assessments, displaying submission counts per student compared to their group members across a set of group work assessments, and for courses with structured roles, displaying the number of times each student takes on each role in a group for a set of group work assessments. These insights can help ensure that students are completing group work equitably and according to course policies, as well as help instructors deciding when interventions are needed.

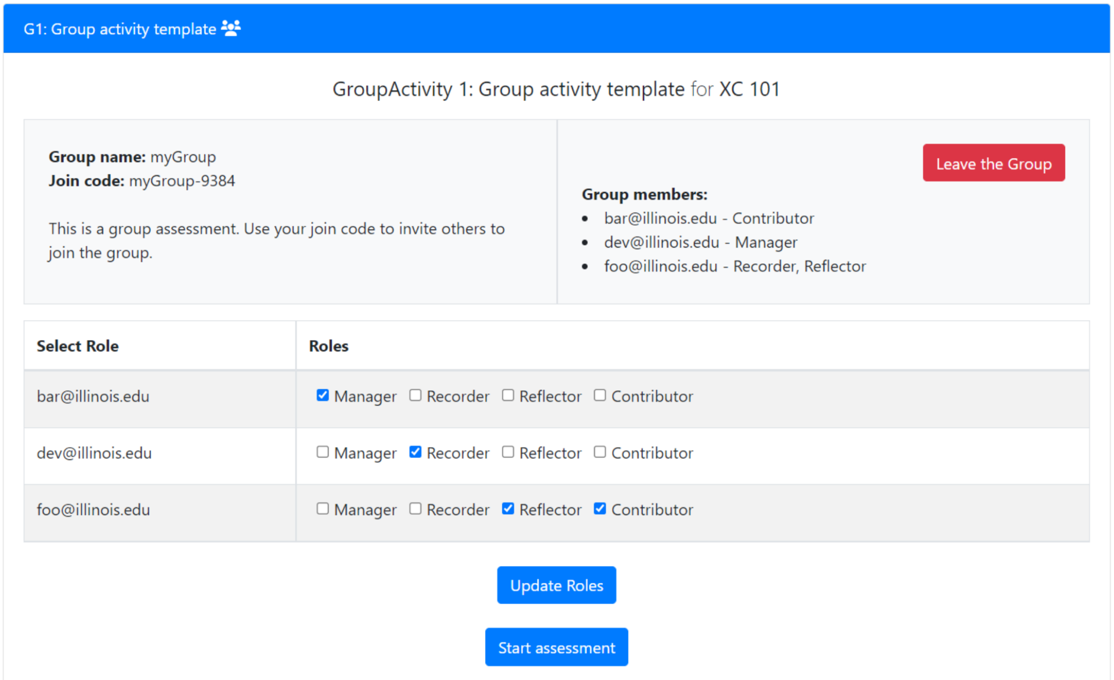


Figure 3: Group initial page, where students can select their roles for the group work. Each role can provide access restrictions (for example View, Edit) in different parts of the assessment.

3 Dissemination and adoption

Our teaching innovations were initially developed to support three undergraduate computer science courses offered at the University of Illinois, Urbana-Champaign (UIUC): Computer Architecture, Numerical Methods and Database Systems. All three courses adopt a flipped classroom format, where students watch pre-recorded videos prior to class, and work on collaborative learning activities during class time. The pedagogies and tools developed for this project have been successfully implemented in online and in-person sections of the three courses.

To our knowledge, the PrairieLearn group features have been adopted by at least 12 other courses in 4 institutions: Statics, Linear Algebra, Differential Equations, Programming Languages and Compilers (UIUC); Computer Network Protocols and Applications, Operating System Fundamentals (York University), Computer Hardware and Operating Systems (University of British Columbia); Computational Science, Computer Science II, Computer Organization and Assembly Language, Operating Systems Concepts, Data Communications (Grand Valley State University).

In Summer 2022, one of the PIs organized a 6-week workshop, name “Incorporating Computing into Engineering Curriculum”, with the goal to provide participant teams from the Grainger College of Engineering at the University of Illinois Urbana-Champaign with resources and tools to facilitate the creation of computational components (exercises, class activities, projects) in their courses. One of the modules of the workshop introduced the collaborative learning tools and pedagogies developed for this project. We will offer this workshop again in Summer 2023, which will be available to participants from other institutions.

4 Impact

4.1 Synchronous versus asynchronous classes

To explore the effects of synchronous versus asynchronous classes, we ran a quasi-experimental research study in two courses: Numerical Methods and Computer Architecture. Policies, assignments, and instructors for the course were kept the same aside from the modality of the courses (synchronous vs. asynchronous) and the scaffolding of role assignments (free-for-all versus structured rolse). We explored the following research questions to better understand the affect of these differences. Methods and more details about findings for this section are available in our prior publication [26].

4.1.1 Research questions

RQ1: What effect do student-scheduled, synchronous classes with free-for-all roles and instructor-scheduled synchronous classes with structured roles have on the student learning experience during collaborative learning activities? In other words, do students score higher or complete assessments faster in one of the conditions?

RQ2: What effect do student-scheduled synchronous classes with free-for-all roles and instructor-scheduled synchronous classes with structured roles have on the equality of the number of students' contributions during collaborative learning activities? In other words, is there less freeloading or dominating in one of the conditions?

4.1.2 RQ1: Effect on Learning Experience

We defined two metrics to describe the quality of the student learning experience during group activities: the performance of submissions and the time to completion. We hypothesize that if group members are helping each other learn, they should make higher performing submissions and they would spend less time to finish an assignment. Analyzing both metrics is important for observing productive collaboration. For example, a team might reduce their time to completion by using a divide-and-conquer rather than collaborative approach, but because team members are not actively helping each other, we would not expect to also see a corresponding improvement in the performance of submissions.

Performance of Submissions: when using PrairieLearn, students can make unlimited submission to the same question without being penalized. Because students were allowed to resubmit answers without penalty until they got them marked as correct, we cannot use the final assessment score as a performance metric, because most teams eventually earned perfect scores. We defined the performance of submissions as the team's average submission score made during an assessment. If a team had a better understanding of the course material, we expect they would achieve higher average submission scores.

By using multi-level modeling, we demonstrated that the submissions by synchronous groups in Computer Architecture were, on average, 4.86 percentage points better than submissions by asynchronous groups. The ICC for this model was 0.44, meaning that just over half of the unexplained variance in the model came from the variance between assessments, and the rest came from the variance between groups working on the same assessment. The standardized effect size was small (0.37) but significant ($p < 0.001$). Likewise the submissions by synchronous groups in Numerical Methods were, on average, 9.63 percentage points better than submissions by "asynchronous" groups. The ICC for this model was 0.22, meaning that most of the unexplained variance in the model came from the variance between assessments, and the rest came from the variance between groups working on the same assessments. The standardized effect size is large (1.31) and significant ($p < 0.001$). In summary, synchronous students performed a half to a full letter grade better on average on their submissions.

Time to Completion: We summed the time between submissions, removing submissions made more than 60 minutes apart. We assumed this amount of time indicated that a team was idle and not actively working on the assessment during that time interval.

By using multi-level modeling, we demonstrate that students in synchronous offering of Computer Architecture completed assessments 10.93 minutes faster on average than "asynchronous" groups. The standardized effect size was medium sized (-0.43) and significant ($p < 0.001$). Summing across all assessments, the medium-effect size translated to synchronous students in Computer Architecture spending 3.98 hours less or 1 full week of instructional time on in-class activities. Likewise, students in the synchronous offering of Numerical Methods completed assessments 16.19 minutes faster on average than "asynchronous" groups. The standardized effect size was large (-1.41) and significant ($p < 0.001$). Summing across all assessments, the large effect size translates to synchronous students in Numerical Methods spending 2.83 hours of instructional time less than students in the "asynchronous" offering (almost a full week) on in-class activities.

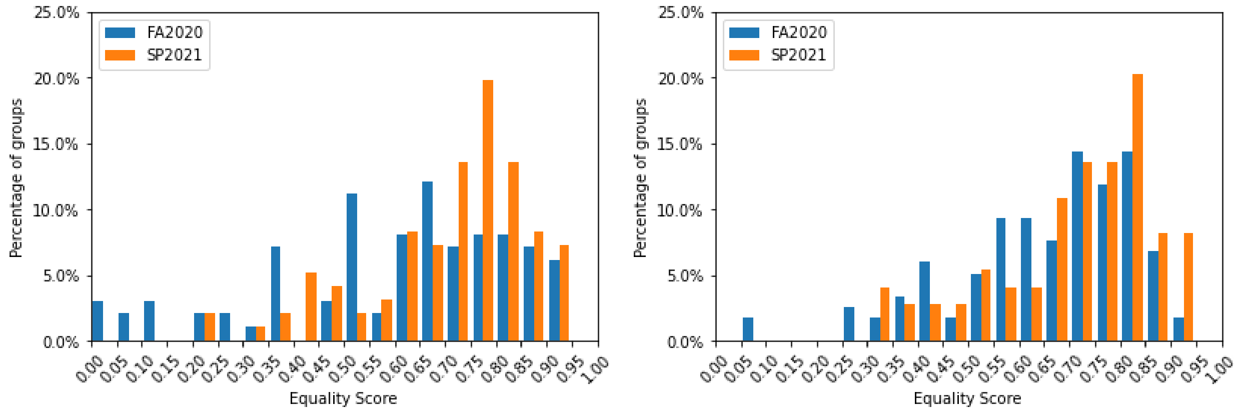
4.1.3 RQ2: Effect on equality of submissions

To define a standard metric for equality, we computed the standard deviation of the percentage of submissions for each team member (σ_p). Because the group size varied between 2–4 and group size affects the maximum possible standard deviation ($\sigma_{\max}$), we normalized our metric by dividing the standard deviation by the maximum possible standard deviation. To improve interpretability, we subtract this quantity from 1 to get the **equality score** e, giving

$$e = 1 - \frac{\sigma_p}{\sigma_{\max}}. \tag{1}$$

Teams with perfectly equal number of submissions from all team members would yield an equality score of 1 and a team where one team member makes all submissions would yield an equality score of 0.

Figure 4a provides a histogram of the equality score for all teams with more than 2 members from Computer Architecture. Figure 4b provides a histogram of the equality score for all teams with more than 2 members from Numerical Methods. A multi-level modeling analysis of these data revealed that the synchronous offering of the courses yielded significantly higher ($p < 0.001$) equality in the number of student submissions with a small effect size ($g = 0.337$).



(a) Computer Architecture (b) Numerical Methods

Figure 4: Histogram of the equality scores for each team for the "asynchronous" (FA20) and synchronous (SP21) offerings.

4.2 A Temporal Analysis of Collaborative Learning Problem-Solving Behaviors

With the increased adoption of collaborative learning pedagogies, instructors must understand their students' problem-solving approaches during these group activities to better design class materials. Among the multiple ways to reveal collaborative problem-solving processes, temporal submission patterns is one that is more scalable and generalizable in Computer Science education. As part of this project, we conducted a temporal analysis of a large dataset of students' submissions during group work assignments in the Database System course.

We used the data log from PrairieLearn's collaborative assessments to extract the timestamp of each student's submissions to a given collaborative problem. Each submission was labeled as quick (Q), medium (M), or slow (S) based on its duration and whether it was shorter or longer than the 25th and 75th percentile. We then applied sequence compacting techniques, sequential pattern mining, and correlation analysis to identify latent patterns that characterize various problem-solving strategies across three database query languages (SQL, MongoDB, Neo4j). The objective of this study is to investigate the potential of temporal information - the amount of time spent on each submission attempt – in uncovering the recurrent patterns in groups' submission sequences. The next step is to perform code analysis of discovered classes of temporal patterns to study the correlations between temporal behavior and collaborative problem-solving strategies.

Figures 5 presents the results of the temporal analysis of students' logs, focusing on submission time across semesters and database query languages. Sequences were extracted and clustered to identify representative submission patterns associated with temporal problem-solving behaviors. Findings revealed three distinct submission patterns that corresponded to novice, advanced beginner, and competent problem-solving behaviors. Novice problem-solving behavior was characterized by numerous submissions with a large portion of slow attempts, whereas advanced beginner problem-solving behavior was characterized by a moderate amount of submissions with several slow attempts in the middle. Competent problem-solving behavior, on the other hand, was characterized by fewer attempts, most of which were medium speed.

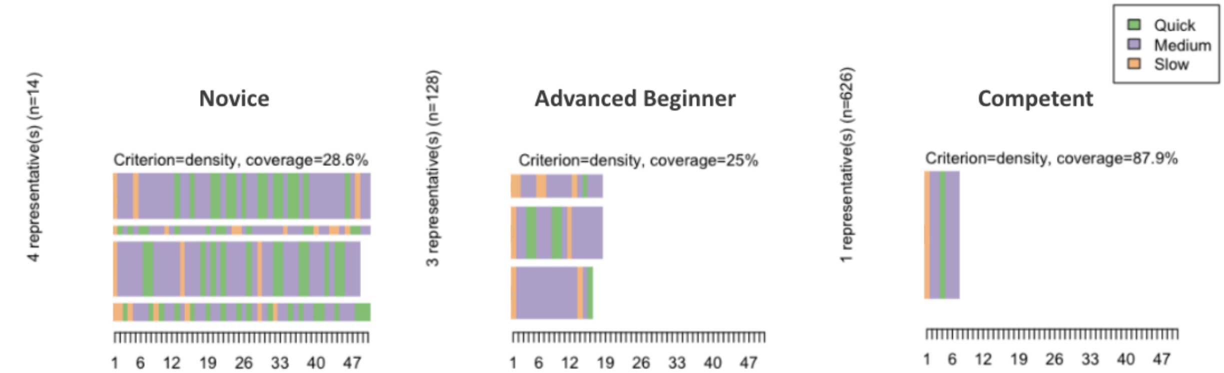


Figure 5: Our temporal analysis of students collaborative assessments log data identified three distinct collaborative classifications that corresponded to novice, advanced beginner, and competent problem-solving behaviors. Under each classification, there could be several temporal patterns. For instance, the *Novice* class has four distinct temporal patterns, each represented by a bar. The height of each bar indicates the number of teams that followed that particular temporal pattern.

4.3 Sense of belonging

Prior to this project, one of the PIs had previously compared the impact of collaborative learning versus lecture-based peer instruction [27], and had found that, on average, students in the collaborative learning section experienced greater sense of belonging compared to their classmates in the peer instruction section. However, one limitation of this study was the size of the collaborative learning section (N=73) compared to the peer instruction section (N=179). One of the goals of this project was to see if the collaborative learning setup could scale to the full classroom with similar improved outcomes in sense of belonging.

Similarly to the past study [27], the project collected students' self-reported sense of belonging in pre- and post-course surveys. Sense of belonging survey items were taken from a subset of questions originally created and validated by Hoffman et al. (2002) [28]. Questions included students' perceived comfort in the classroom (e.g., "I feel comfortable volunteering ideas or opinions in class"), perceived isolation (e.g., "It is difficult to meet other students in class"), and perceived support (e.g., "I feel that an instructor would take the time to talk to me if I needed help"). Questions on the pre-course survey asked students to answer based on prior experiences in past computing courses. Questions on the post-course survey asked students to answer on their experiences in the course.

A total of 10 items were used in surveys. Questions were presented on a 5-point Likert scale using the following options and numeric values: "Strongly disagree" (-2), "Disagree" (-1), "Neither agree nor disagree" (0), "Agree" (1), "Strongly agree" (2). Negatively phrased questions (e.g., "I know very few people in class") were reverse coded. Individual questions scores were averaged together to get a students' individual sense of belonging score between -2 and 2.

Figure 6 shows averages in students' sense of belonging scores by self-identified gender. Only results from the Spring and Fall 2022 semesters are included for Numerical Methods and Database Systems as surveys in prior semesters for these courses were submitted anonymously.

Across pre-course surveys, men on average had higher sense of belonging scores compared to women, however this difference does not seem as pronounced in the Fall 2022 semester. This may be due to other professors' changes at the introductory course level as well as ongoing changes in students' overall perceptions due to COVID. Focusing on post-course surveys, we see increases for both men and women such that the gender difference appears to disappear, except for Fall 2022 in Numerical Methods and Database Systems where women seem to experience greater sense of belonging by the end of the course. These results indicate a positive trend from pre- to post-course for sense of belonging in all three courses. However, due to a lack of a control group for comparison (i.e., students work individually during class), we cannot isolate these trends as a direct effect of group activities.

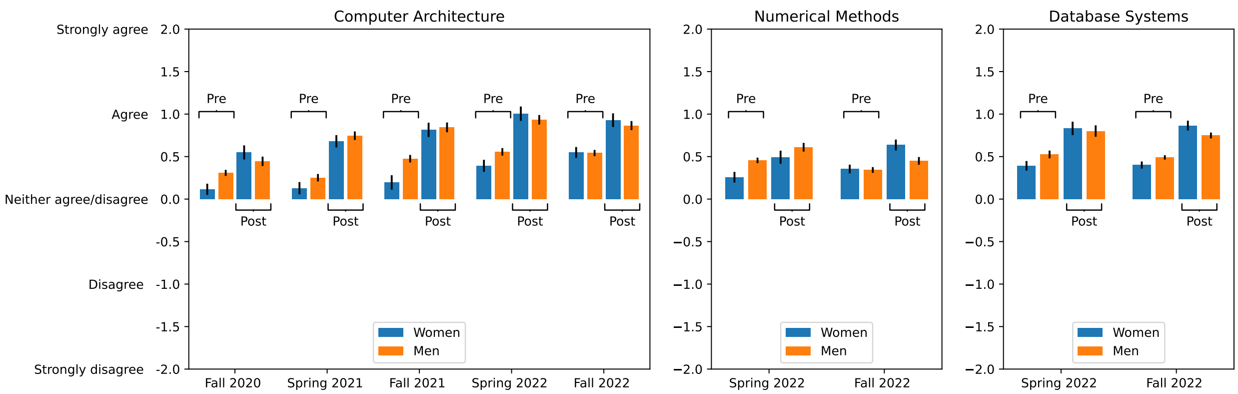


Figure 6: Student's self-reported sense of belonging by self-identified gender. Brackets above bars indicate results from the pre-course survey. Brackets below bars indicate results from the post-course survey. Questions used a 5-point Likert scale from Strongly disagree (-2) to Strongly agree (2).

Based on these results, we wanted to dive deeper into students' experience within their groups. Starting in the Spring 2022 semester, we modified the sense of belonging questions to focus on the group rather than the classroom at large. For example, questions were rephrased to ask about students' perceived comfort in their group (e.g., "I feel comfortable volunteering ideas or opinions in my group"), perceived isolation within their group (e.g., "It is difficult to talk to my group"), and perceived support from their group (e.g., "I feel that a group member would take the time to talk to me if I needed help"). The scores were aggregated in the same manner as the whole class sense of belonging scores. The results for students' self-reported sense of belonging from the post-course surveys comparing whole class to within group is shown in Figure 7. Across all courses, students seem to experience greater sense of belonging within their group compared to the whole class, which is to be expected as they spent a substantial amount of time with their group. For example, students in Computer Architecture usually worked with their groups during in-class activities and on weekly lab assignments, and students in Database Systems also worked with their groups on a semester-long project. Additionally, the difference across courses seems less pronounced when comparing within group averages to whole class averages. The relatively lower within class sense of belonging for Numerical Methods may be due to two reasons: 1) students have relatively less time in class (i.e., the group activities once per week), and 2) Numerical Methods is taught to a wider variety of majors as it is also a course requirement or alternative for other majors and minors.

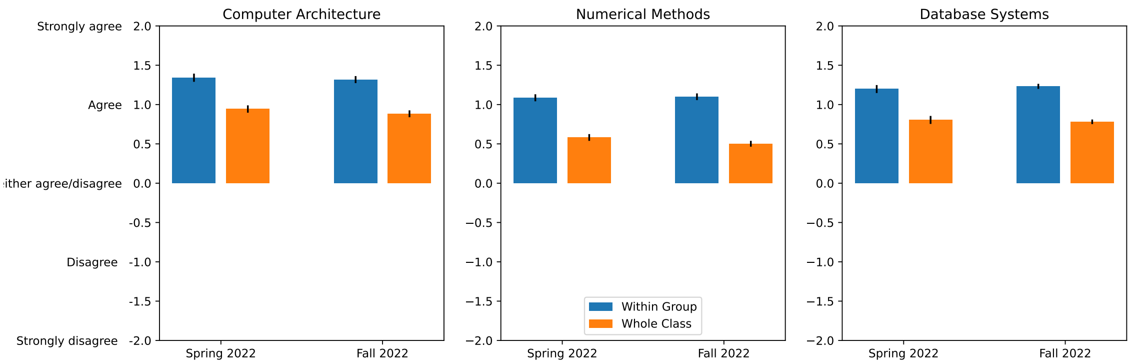


Figure 7: Student's self-reported sense of belonging within their group compared to whole class based on data from Spring and Fall 2022's post-course surveys.

Since Fall 2021, Numerical Methods has offered two sections, one intended for in-person, synchronous group activities and one intended for online, asynchronous group activities. Students in the in-person section were required to attend class, whereas students in the online section could choose when and where to meet. Typically, students in the online section would meet virtually, even if they chose to meet during the official class time. On average, students in the in-person section seem to experience greater sense of belonging in general compared to students in the online section, and these differences seem larger for within group sense of belonging than whole class. The differences in within group sense of belonging may speak to a difference in how students interact and connect with each other in online spaces. The differences in whole class sense of belonging may speak to the relatively smaller amount of time students spend in class (i.e., officially, students are required to meet once per week for group activities, compared to twice a week for Computer Architecture and Database Systems).

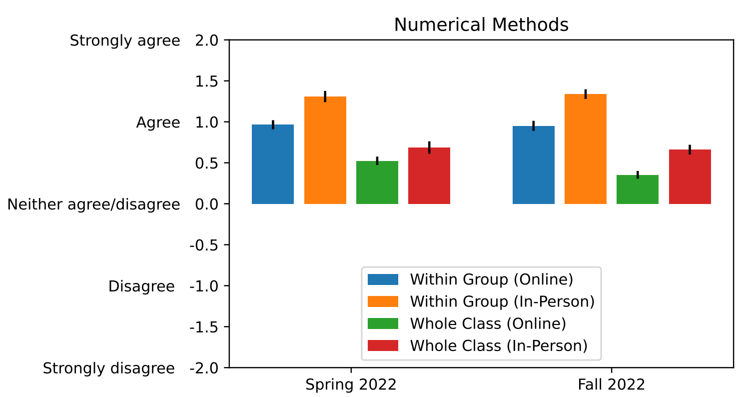


Figure 8: Student's self-reported sense of belonging within their group compared to whole class by enrollment based on data from Spring and Fall 2022's post-course surveys for Numerical Methods.

5 Implications and Future Work

Our team consists of three instructors from Illinois' Computer Science Department who teach courses at different levels of the curriculum: Computer Architecture (a required course for sophomore and juniors), Numerical Methods I (a required course for juniors and seniors), and Database Systems (an elective course for senior and graduate students). This diverse combination of courses provides us with a unique opportunity to study the effects of collaborative learning across CS curriculum on students' classroom experience and sense of belonging. For example, the Database Systems instructor has observed improved group dynamics and reduced resistance to collaborative assessments compared to early implementations of these type of activities in his course. We have also received similar feedback from other instructors outside of this research group (e.g., from one of the instructors of Programming Languages & Compilers, a required course to Computer Science seniors). This improvement may be attributed to students' prior exposure to similar collaborative learning policies in earlier courses, and an overall increase in sense of belonging. More research studies are needed to investigate the longitudinal impact of collaborative learning activities across the curriculum.

In summary, the work developed for this project has impacted thousands of students for the last couple of years at the University of Illinois Urbana-Champaign. We have established a model for teaching computer-based collaborative learning activities that is scalable and feasible. We have tested our models in different contexts, including mathematics, engineering and computer science courses, online and in-person courses, classes offered as small discussion sections, in large lecture halls or flexible style classrooms and led by instructors or graduate/undergraduate teaching assistants. The teaching pedagogies and tools developed for this project are ready to be adopted by faculty from other universities.

References

[1] S. Freeman, S. Eddy, M. McDonough, M. Smith, N. Okoroafor, H. Jordt, and M. Wenderoth, "Active learning increases student performance in science, engineering, and mathematics," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 23, pp. 8410–8415, 2014.

[2] J. Gasiewski, M. Eagan, G. Garcia, S. Hurtado, and M. Chang, "From gatekeeping to engagement: A multicontextual, mixed method study of student academic engagement in introductory stem courses," *Research in Higher Education*, vol. 53, p. 229–261, 2012.

[3] T. Tucker, S. Shehab, E. Mercier, and M. Silva, "Board 50: Wip: Evidence-based analysis of the design of collaborative problemsolving engineering tasks," *Proceedings of American Society for Engineering Education*, 2019.

[4] T. Nokes-Malach, J. Richey, and S. Gadgil, "When is it better to learn together? insights from research on collaborative learning," *Educational Psychology Review*, vol. 27, p. 645–656, 2015.

[5] E. Mercier and S. Higgins, "Collaborative learning with multi-touch technology: Developing adaptive expertise," *Learning and Instruction*, vol. 25, p. 13–23, 2013.

[6] L. Paquette, N. Bosch, E. Mercier, J. Jung, S. Shehab, and Y. Tong, "Matching data-driven models of group interactions to video analysis of collaborative problem solving on tablet computers," *Proceedings of International Conference of the Learning Sciences*, vol. 1, 2018.

[7] S. Shehab, L. Lawrence, E. Mercier, A. Margotta, E.Livingston, M. Silva, and T. Tucker, "Towards the effective implementation of collaborative problem solving in undergraduate engineering classrooms: co-designing guidelines for teaching assistants," *Proceedings of American Society for Engineering Education*, 2020.

[8] A. Solimeno, M. E. Mebane, M. Tomai, and D. Francescato, "The influence of students and teachers characteristics on the efficacy of face-to-face and computer supported collaborative learning," *Computers & Education*, vol. 51, no. 1, pp. 109–128, 2008.

[9] S. Dewiyanti, S. Brand-Gruwel, W. Jochems, and N. J. Broers, "Students' experiences with collaborative learning in asynchronous computer-supported collaborative learning environments," *Computers in Human Behavior*, vol. 23, no. 1, pp. 496–514, 2007.

[10] "Gradescope," https://www.gradescope.com/, Last accessed on 2023-01-29.

[11] "Canvas," https://www.instructure.com/higher-education/products/canvas/canvas-lms, Last accessed on 2023-01-29.

[12] "Cocalc," https://cocalc.com, Last accessed on 2023-01-29.

[13] "Google colab," https://colab.research.google.com/, Last accessed on 2023-01-29.

[14] "Google docs," https://www.google.com/docs/about/, Last accessed on 2023-01-29.

[15] "Prairielearn: mastery learning meets online assessment," https://prairielearn.com/, Last accessed on 2023-01-29.

[16] M. West, G. L. Herman, and C. Zilles, "Prairielearn: Mastery-based online problem solving with adaptive scoring and recommendations driven by machine learning," in *2015 ASEE Annual Conference & Exposition*. Seattle, Washington: ASEE Conferences, June 2015.

[17] "Pogil: Process oriented guided inquiry learning," https://pogil.org, Last accessed on 2023-01-29.

[18] "Visual studio code," https://code.visualstudio.com, Last accessed on 2023-01-29.

[19] "Jupyter lab," https://jupyter.org, Last accessed on 2023-01-29.

[20] R. S. Moog and J. N. Spencer, *Process oriented guided inquiry learning*. American Chemical Society Washington, DC, 2008, vol. 994.

[21] D. Bodemer and J. Dehler, "Group awareness in cscl environments," *Computers in Human Behavior*, vol. 27, no. 3, pp. 1043–1045, 2011.

[22] J. Janssen, G. Erkens, and P. A. Kirschner, "Group awareness tools: It's what you do with it that matters," *Computers in human behavior*, vol. 27, no. 3, pp. 1046–1058, 2011.

[23] J. Janssen and D. Bodemer, "Coordinated computer-supported collaborative learning: Awareness and awareness tools," *Educational psychologist*, vol. 48, no. 1, pp. 40–55, 2013.

[24] N. Joshi and S.-K. Lau, "Effects of process-oriented guided inquiry learning on approaches to learning, long-term performance, and online learning outcomes," *Interactive Learning Environments*, 2021.

[25] C. Kussmaul and H. Hu, "Improving Online Collaborative Learning with POGIL Practices," *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education (SIGCSE)*, 2021.

[26] G. H. G, Y. Jiang, Y. J. Y., S. Poulsen, M. West, and M. Silva, "An analytic comparison of student-scheduled and instructor-scheduled collaborative learning in online contexts," in *Proceedings of the 2022 American Society for Engineering Education Conference*, 2022.

[27] G. L. Herman and S. Azad, "A comparison of peer instruction and collaborative problem solving in a computer architecture course," in *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, ser. SIGCSE '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 461–467. [Online]. Available: https://doi.org/10.1145/3328778.3366819

[28] M. Hoffman, J. Richmond, J. Morrow, and K. Salomone, "Investigating "sense of belonging" in first-year college students," *Journal of College Student Retention: Research, Theory & Practice*, vol. 4, no. 3, pp. 227–256, 2002.