Agora: Motivating and Measuring Engagement in Large-Class Discussions

Hedayat Zarkoob
University of British Columbia
Vancouver, Canada
hzarkoob@cs.ubc.ca

Siddharth Nand
University of British Columbia
Vancouver, Canada
sidnand@student.ubc.ca

Kevin Leyton-Brown
University of British Columbia
Vancouver, Canada
kevinlb@cs.ubc.ca

Giulia Toti
University of British Columbia
Vancouver, Canada
gtoti@cs.ubc.ca

ABSTRACT

Cold calling effectively incentivizes all students to actively prepare contributions to a class discussion, but some find it terrifying. Rewarding voluntarily speaking in class is less off-putting, and can be valuable for students who participate; however, it can allow a large fraction of the class to disengage. Agora is an open-source app designed to serve as a middle ground between these extremes, with the added benefit that it automatically produces an assessment of each student's engagement. The key ideas are to give students control over whether their hand is raised or lowered, to choose randomly among students with raised hands, and to give participation credit to all students who were considered every time a speaker is chosen. The system has various other features to facilitate deployment in large classes including multiple queues to support concurrent questions on different topics; a message board to allow students to communicate discretely with the instructor; and polling. We deployed the system in three offerings of a large undergraduate class and demonstrate its effectiveness in terms of learning outcomes, gender balance in participation, and student satisfaction.

CCS CONCEPTS

• **Applied computing** → **Collaborative learning**.

KEYWORDS

In-class Participation, Educational Technology

ACM Reference Format:
Hedayat Zarkoob, Siddharth Nand, Kevin Leyton-Brown, and Giulia Toti. 2024. Agora: Motivating and Measuring Engagement in Large-Class Discussions. In *Proceedings of the 2024 Innovation and Technology in Computer Science Education V. 1 (ITiCSE 2024), July 8–10, 2024, Milan, Italy.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3649217.3653540

ITiCSE 2024, July 8–10, 2024, Milan, Italy

ACM ISBN 979-8-4007-0600-4/24/07…$15.00
https://doi.org/10.1145/3649217.3653540

1 INTRODUCTION

Active learning, a pedagogical approach that emphasizes students' participation in the classroom, has been widely embraced in higher education. Indeed, it offers many documented benefits for students [2, 17, 29, 30]. Whole-class discussion, where the instructor engages students by inviting them to ask and answer questions, is possibly the most ubiquitous form of active learning. It is well established that such discussions improve learning outcomes [6, 19]. However, it is far from trivial to facilitate a class discussion in a way that maximizes its benefit to as many students as possible. The crucial active learning benefit a student receives from a class discussion is trying to formulate a response—whether or not they actually get the chance to speak out loud. Notably, offering students the opportunity to formulate a response is scalable even to large class settings, even though actual speaking opportunities are necessarily limited. However, when no immediate rewards are provided, a significant fraction of students may choose to participate very little, or not at all [22, 32, 35], and substantially miss out as a result. Even worse, specific groups of students shy away from class discussions more frequently than others. When compared to men, women tend to speak less frequently, and with less confidence [1, 10, 12, 40]. This seems to go beyond personal choice; various studies have provided evidence that instructors can subtly (and often unconsciously) favor male students, asking them better questions and providing them with more feedback [8, 34]. Differences in participation between men and women tend to become more pronounced as the percentage of men in the classroom increases [38], a particular concern for Computer Science classes. First-generation students [37] and non-native speakers [20] have also been found to be less likely to voluntarily contribute to in-class discussions.

To maximize students' incentives to actively formulate responses and thus to benefit from active learning, some instructors employ *cold-calling*, a strategy in which students are randomly called upon to answer questions. Cold calling can be very effective in engaging students and driving their learning. Most students will engage with the material and formulate an answer, whether or not they get to share it with the rest of the class. Studies indicate that in classes with frequent cold-calling, significantly more students volunteer to answer questions, and their comfort in participating in class discussions increases over time [14]. Moreover, cold-calling can increase voluntary participation for both men and women, without causing disproportional discomfort for either gender [15]. However, some students find the constant risk of being cold-called to be terrifying, and shy students can even choose to skip class [5, 25]. These negative effects can be especially pronounced for students who already face structural disadvantages, such as first generation students and non-native speakers [21, 28]. In response, alternative approaches have been proposed. For example, "warm calling" provides students with advance notice and the option to opt out of participating in front of the class [28]; instructors can also share a list of possible questions before the lecture [23].

More typically—particularly in the Computer Science context—students are never called upon unless they volunteer to speak, but are given some kind of grade incentive for participating. Indeed, the literature shows that students are very responsive to changes in participation incentives [33], especially if participation is a component of course grades [13]. Krohn et al. [24] compared the effects of different participation conditions and found that the introduction of participation credit increased balance in participation among students; interestingly, the same effect was observed both when student participation was recorded and self-reported. Foster et al. [16] showed that participation could be increased in initially low-responding students by providing credit for recording class comments on note cards. Sommer and Sommer [36] rewarded students for participation only on alternating days, but showed that participation increased in non-rewarded classes as well. McCleary et al. [26] generalized this approach, introducing a randomized approach to participation credits. In particular, they assigned participation credits only on *randomly* selected days, observing a significant reduction in non-participation rates and an increase in credit-level participation. Boniecki and Moore [4] implemented a token economy, where participating students earned tokens that could be exchanged for extra credit. The study observed participation levels in an introductory psychology class before, during, and after implementing the token system, noting a substantial increase in participation during the token economy that persisted after its removal. Overall, participation incentives have a positive and well-documented impact on fostering inclusive and engaged learning environments. However, anyone who has taught using a variant of this methodology is well aware that many students simply disregard the incentives and disengage, losing out on the benefits of active learning as a result.

Various technological tools (notably, clickers) have been introduced with the aim of broadening class participation, particularly in very large classrooms. Clickers are tools that allow every student in a class to answer live questions, typically in a multiple-choice form. They are typically well received by students and contribute to understanding and retention of material [3, 7, 31, 39]. They are often used to facilitate other active learning strategies, such as Think–Pair–Share [3, 9]. For students who carry a smartphone with them to class, purchasing a physical clicker is no longer necessary. Some polling tools are even designed with the specific goal of increasing engagement via a quiz-like appearance (*Kahoot, Quizizz, Mentimeter* are a few examples). Despite their benefits, clickers and online polling apps are not replacements for whole-class discussions. Quinn [31] expressed the concern that students may hide behind clickers rather than speaking in class, and as a result may not practice professionally useful skills. Clickers also necessarily reduce discussion to a small, discrete set of predefined choices, limiting both the depth and the spontaneity of engagement opportunities.

This paper unifies these three themes: incentivizing as many students as possible to prepare answers to every question in a rich, oral discussion; lowering the fear factor through voluntary participation; and leveraging technology to scale to large classrooms. The key idea is moving from incentivizing participation—which is an inherently scarce resource that grows scarcer with class size—to incentivizing *willingness to participate*. We propose a kind of opt-in cold calling, where students dynamically decide whether to volunteer themselves for consideration as a class discussion progresses, and all such students receive participation rewards. Of course, no such system could feasibly be implemented without technology. We thus (in Section 2) present *Agora*, a web/smartphone-based participation app, that realizes our vision. Agora presents a simple interface allowing students to "raise their hand" and "lower their hand" freely as class progresses. Every time the instructor wants to hear a student speak, they push a button and Agora both randomly chooses a single speaker and grants participation credit to *all* students who put themselves at risk of being called upon. Beyond its core mechanic, Agora has a mature set of features that provide auxiliary benefits in a large class. It allows students to raise hands in multiple "lists" to allow for side conversations; offers a live message board that lets students discretely communicate with the instructor without disrupting the class; and supports live polling, both graded and anonymous. It allows students to authenticate using standard university credentials via SAML 2, allowing grades to be tracked reliably. And it is both easily customizable and freely available as open-source software, making it appropriate for a wide variety of deployments.[1] We deployed Agora ourselves in three previous iterations of a large Computers and Society course. We describe the results in Section 3, demonstrating that it led to tangible improvements in students' learning outcomes, benefited gender diversity, and improved student satisfaction as measured via a survey.

2 THE AGORA SYSTEM

We now introduce Agora and its features in more detail and describe the rationale for some of the design choices that we made. We begin by explaining how an instructor can use Agora to call on students in a semi-randomized way. We then discuss various capabilities offered by the system: calling on students from multiple lists; grading students based on their participation in the system; the option for students to send live messages to the instructor; and the system's support for both graded and ungraded polls.

2.1 Calling On Students

Students' core interaction with Agora is through pressing a button to "raise their hand" every time they want to contribute to a discussion (Figure 1 Right), or to press the button again to "lower their hand". Every time the instructor wants to call on someone, they press a button in their own interface (see Figure 1 Left). As a result, one student among those with their hand raised is selected at random by the system. This student's name and photo (if available) are shown to the instructor (so that they may proceed to call them).

[1] Agora's implementation is open-source; see https://github.com/hezar1000/agora.

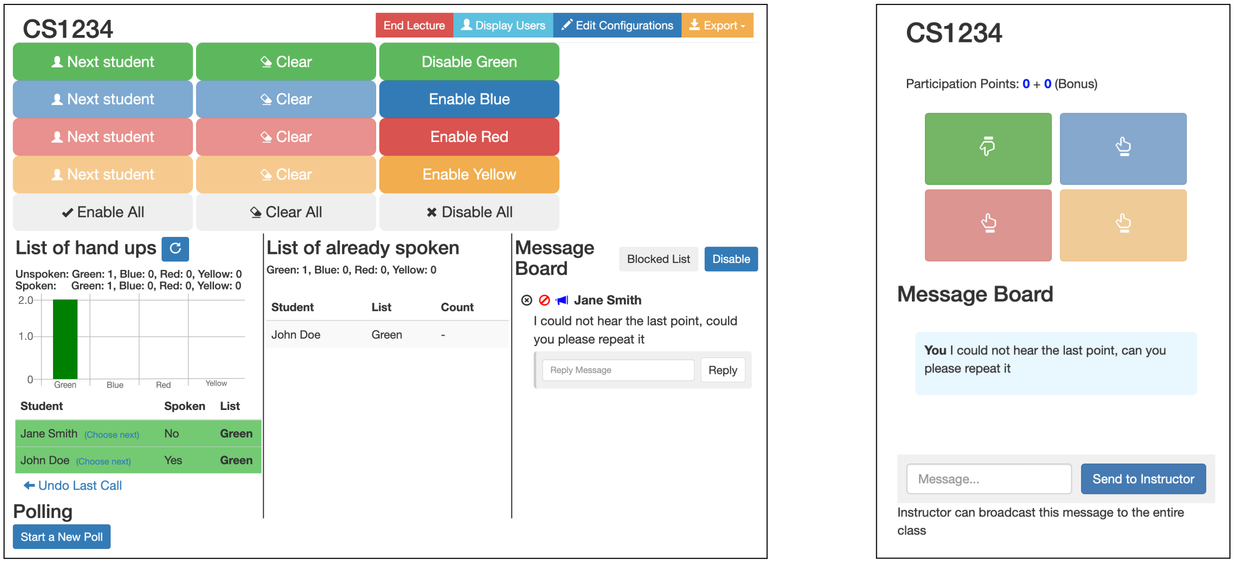

Figure 1: The instructor dashboard (left) and student dashboard (right).

Every other student whose hand was raised at the moment of the instructor's "call" is also recorded.

The instructor's dashboard further displays a list of all students who have their hands up—indicating for each student whether or not they have previously spoken in the same class—and a log of students who previously spoke. By default, the former list is sorted according to the order in which students raised their hands, with the earliest hand raise listed first. The instructor can use this information to ensure that a sufficiently large number of students are willing to speak before pressing the call button. The log of students who have already spoken is sorted in descending chronological order. This is intended to make it easy for the instructor to refer back to the names of students who recently spoke, facilitating calling back to their remarks as the discussion progresses and thereby learning students' names.

When the instructor presses the "call" button, by default, Agora first checks if any student with a raised hand has not yet spoken during the current lecture. If one or more of such students exist, Agora eliminates all other students from consideration; otherwise, it keeps all students in consideration. Once the poll of students from which to call has been adjusted, Agora computes a probability distribution from which to sample, giving a boost to students who have been less active in previous lectures. The benefit of favoring students who have not spoken recently is twofold: it increases the diversity of voices in the classroom and it reduces the incentive for students to only raise their hands during popular questions, hoping that they would not be the one being called. More formally, Agora assigns a weight to each student i of $1/\log_2(\max(c_i, 2))$, where c_i is the number of times i has been chosen to speak so far in the course. Agora then normalizes these weights and samples from the resulting probability distribution.

The instructor has an alternative to the "call" button: they can also manually select any student with a raised hand. From the students' side and with regards to recording participation for all students, this acts just as if the instructor had pressed the "call" button and the same student had been chosen at random. This feature is designed to give flexibility to the instructor, e.g., if they want to recognize the first person who raised their hand, if they want to hear from someone who rarely speaks, or if they know that a particular student has experience relevant to the question.

2.2 Multiple Lists

Agora provides up to four different "raise hand" buttons, color-coded as Green, Blue, Red, and Yellow (although if undesired, certain buttons can be turned off in the instructor interface). Each button corresponds to a different list; the instructor can e.g., press the Blue "call" button and randomly identify a student from that list using the logic described above, and with participation credit awarded to all of the students on (only) the Blue list. We envision at least two use cases for this functionality.

First, the lists allow side conversations without clearing the list of raised hands from the main discussion. For example, as a class discussion is proceeding using the Green list, a student might ask a question. The instructor might want to ask if any other students know the answer to this question. In this case, the Green list will be full of students with other points to make in response to the instructor's original prompt, so the instructor might instead direct students who'd like to address the new question to raise their hands in the Blue list. Once that conversation wraps up, the instructor can proceed back to the Green list.

Second, multiple lists can be used to elicit students' opinions on a multi-dimensional topic. For example, an instructor may want to hear arguments both in favor of and against a given proposition. In such a case, they can ask students to raise their hands in the appropriate lists and then iterate over the lists as appropriate to elicit different opinions. Note that Agora also offers a polling feature that covers a distinct but related use case; see below.

Raise your hand at least this many times	*Get this many points*
0	0
1	5
2	6
3	6.5
4	7
The median	8
The maximum	10

Table 1: A per-class grading scheme we used with Agora. The median and maximum were taken among students who participated at least once in the class.

2.3 Grading

For every class section and for every student, Agora records a count of the number of times the instructor called on the class and the given student had their hand raised. In our own teaching, we use this information as the basis of a class participation grade, applying a nonlinear formula that gives decreasing numbers of marginal grades for subsequent marginal hand raises (and thereby gives stronger marginal incentives for participation to shy students); this scheme is presented in Table 1. We average students' participation grades across the whole term to obtain their final grades, dropping a prearranged number of classes to allow for absences. We make it clear in our class conduct rules that the use of Agora when not physically present in the classroom is an academic honesty violation; we note to students that Agora's bias towards students who have spoken less often makes such behavior risky. Because Agora's selection mechanism makes it very unlikely (at least in a large, active class) that a student would be selected a second time after speaking in class, we recognize that such a student could "safely" raise their hand during every subsequent call and get the participation points. To prevent students from needing to game the system, we automatically award "bonus hand raises" for every call that follows an occasion when a student speaks in class. Of course, all of these system parameters can be adjusted.

Instructors can choose to assign different numbers of points for participation in different lists. They can also configure the system so that the very first hand raise by each student is worth more than subsequent ones, again potentially differing across lists. This allows instructors to incentivize shy students to risk participation at least once in each class. Lists can also be configured to award no credit at all; for example, the red list could be used for raising logistical questions for which participation credit is not granted.

2.4 Live Message Board

In sufficiently large classes, the just-described mechanism for handling logistical questions will not scale well: the instructor will not want every logistical question to interrupt the whole class. We thus designed a live message board in which students can communicate only with the instructor. This feedback mechanism is designed especially for flagging logistical issues and misunderstandings in a minimally disruptive way. Messages appear in the instructor interface in real time. The instructor may simply choose to react verbally to useful messages and to ignore others. Agora also allows them to send text replies to the student who wrote a given message; to broadcast useful messages to the whole class; and to block specific students from sending further messages for the rest of the lecture. If a TA is present in class, they can also reply to student messages in real time. Of course, the message board can be disabled entirely if an instructor does not want to use it.

2.5 Polling

Agora allows the instructor to create polls having a custom number of choices and a custom prompt for each choice. The instructor can choose whether to save both students' answers (for simple auto-grade quizzes) and IDs or only students' IDs without answers (for polls that are graded only for participation and/or for which students want to know that their responses are anonymous). Once a poll starts on Agora, the instructor is shown a diagram summarizing live results. When the poll ends, both the instructor and students receive a diagram summarizing the final results; this is also saved as an image for the instructor. Additionally, the instructor can later export raw data corresponding to poll results as a CSV file.

3 EVALUATION

We deployed Agora in three in-person undergraduate-level computer science classes: two sections (having different instructors) in Fall 2022 and one in 2023, having 74, 120, and 115 students respectively. These classes focused on the interplay between information technology and society and followed a flipped classroom design, providing students with many opportunities to engage in class discussions and debates. Ten percent of students' total grades came from in-class participation, which we tracked via Agora using the grading scheme previously described in Section 2.3. We also gave students a reduction in an assigned peer grading obligation for each class in which they raised their hands at least once, which we did to incentivize attendance. A similar version of the course, not using Agora, ran in Fall 2018 and Fall 2019. During these courses, an in-class teaching assistant manually recorded the name of every student who was chosen by the instructor to speak in each class. During these two course offerings, students were also offered the opportunity to participate offline through Piazza, which was awarded half a participation credit per week. We applied for and received human-subject approval to study all of this data, and furthermore to access the self-reported genders of all students in our class, as recorded in the university database. (We note that this data regrettably only allows two options for gender: male and female.) To further enrich our data, we integrated students' grades with their participation data and conducted a survey asking students about their experience with Agora. After each course was completed, we contacted students explaining the details of this study and giving them the opportunity to withdraw all of their data from it; in total, only one student did so.

3.1 Impact on Learning Outcomes

3.1.1 Participation vs. Final Grades. We examined the correlation (as Pearson correlation coefficient) between the number of times students raised their hands in class and their final exam performance in Fall 2022 and Fall 2023 (see Table 2). In Fall 2022, Section

Class	*Correlation Coeff.*	*p-value*
Fall 2022 Section 1	0.0203	0.8310
Fall 2022 Section 2	0.2475	0.0389
Fall 2023	0.2422	0.0142

Table 2: Pearson correlation coefficients between students' participation and final exam performance.

Time Period	*Correlation Coeff.*	*p-value*
Midterm 1 until Midterm 2	0.1521	0.1270
Midterm 2 until the Final	0.0681	0.4968
Midterm 1 until the Final	0.1954	0.0491

Table 3: Pearson correlation coefficients between students' participation and performance improvement in Fall 2023.

2, and Fall 2023, we observed a correlation between the number of hand raises and students' final grades. Section 1 of Fall 2022 was different; we observed a much smaller positive correlation that was not statistically significant (p-value > 0.05).

3.1.2 Participation vs Grade Improvement. A correlation between the number of times students raised hands and their final performance could have been due to factors besides active participation. For example, rather than participation causing a performance benefit, high-performing students might have tended to participate more in class. To more deeply investigate the hypothesis that in-class participation benefits students, we looked at the percentage improvement of students' grades throughout the course in Fall 2023. (The Fall 2023 class had 2 midterms and a final exam; in contrast, the Fall 2022 sections replaced the midterm exams with a single take-home essay that does not lend itself to the same analysis.) We observed a positive correlation between the amount by which students' grades improved and their level of participation in class between Midterm 1 and Midterm 2. We observed a similar but smaller correlation between Midterm 2 and the final (see Table 3).

3.2 Balance in Participation

In the three courses in which Agora was deployed, 100% of students ended up raising their hands at least once in class in all three offerings. 92% of students spoke at least once in Fall 22 Sec 1, 96% in Fall 22 Sec 2, and 91% in Fall 23. We note that these numbers by themselves may be astounding to some readers; they are certainly not the norm at our own university among Computer Science courses having between 74 and 120 students.

3.2.1 Inequality of Distribution. We used the Gini coefficient [18] to measure how equally students participated in our classes. The Gini coefficient is a widely used measure of inequality, applicable to any distribution, including the distribution of total participation opportunities among students in our classes. The coefficient ranges from 0 (perfect equality: all students participate the same amount) to 1 (perfect inequality, one student receives all the participation points and all others receive none).

Class	*Gini Coeff.*	
Fall 2018	0.478	Without Agora
Fall 2019	0.561	
Fall 2022 Section 1	0.429	With Agora
Fall 2022 Section 2	0.386	
Fall 2023	0.383	

Table 4: Gini coefficient of students participation.

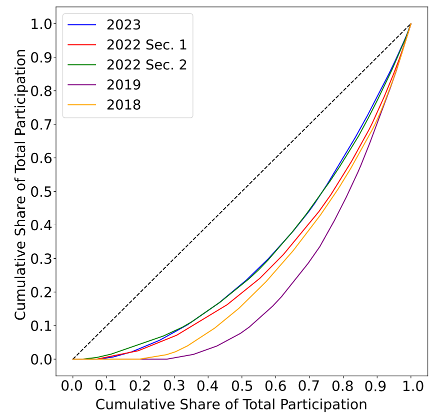

Figure 2: Lorenz curve of students' participation distribution

As seen in Table 4, both sections in Fall 22, and Fall 23 have Gini coefficients smaller than those for the classes in Fall 2018 and Fall 2019, indicating that Agora distributed opportunities to speak more equally among students than before its deployment. We also conducted significance tests by bootstrapping our datasets (resampling our data with replacement to generate 10,000 new, synthetic classes of the same size) to check whether these differences were statistically significant. We tested the null hypothesis that each of Fall 22 and Fall 23's Gini coefficients was greater than or equal to those of Fall 18 and 19 respectively. Choosing our p-value at 0.05, we found these differences to be significant except for Fall 22 Section 1, for which we were unable to reject the null.

We use Lorenz curves to visually present the distribution of participation in our classes. The closer a Lorenz curve is to the $y = x$ line, the more equal participation is among students. (Indeed, the Gini coefficient is the ratio of the area between $y = x$ and the Lorenz curve, divided by the total area under the $y = x$ line.) As seen in Figure 2, the classes in which Agora was deployed achieved more equitable distributions across the board.

3.2.2 Participation by Women. Gender bias and inequality are longstanding concerns in Computer Science, with female students known to have lower participation levels in computer science courses [11, 27]. Table 5 summarizes the population of women in our classes and the fraction by which they contributed to total participation. Ideally, we would want these fractions to be as similar to each other as possible in each year, although we also expect a certain amount of random variation from year to year. In order to more deeply understand whether women in our classes participated

Class	*% women*	*% of times spoken by women*	
Fall 2018	42	40	Without Agora
Fall 2019	45	51	
Fall 2022 Section 1	29	33	With Agora
Fall 2022 Section 2	40	37	
Fall 2023	41	42	

Table 5: % of women in each course and % of times they have spoken out of total students' interventions.

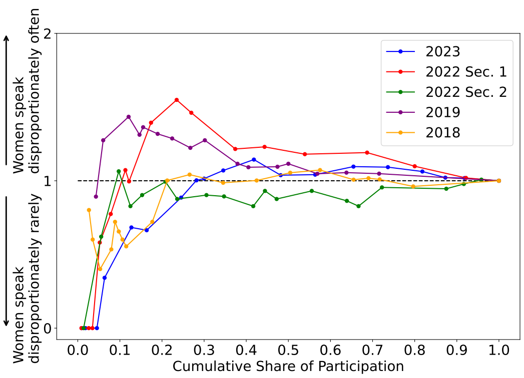

Figure 3: Participation by female students

at above or below their level of representation, we looked at the cumulative share of participation in our classes (Figure 3). In Fall 2018 and 2019, the instructor reports having made a conscious effort to maintain gender balance in class discussions. In contrast, Agora had no way of having any knowledge about students' genders. However, it managed to choose students so that women's participation was close to their level of representation. More specifically, in all classes the very most active (e.g., top 5% of) participants were men. In all classes other than 2022 Sec 2, the vast majority of women participated above their level of representation: i.e., the typical woman participated more actively than the typical man. Overall, we can conclude that there was a small number of highly engaged men, but, in the remaining population, women were actually more likely to participate. Finally, we can see that the students least likely to engage were overwhelmingly men. In 2022 Sec 2, women's participation remained below the population percentage across the cumulative distribution, but nevertheless did remain close to parity beyond the most engaged 10% of students. We do note that in 2022 Sec 1 had a male instructor and Sec 2 had two female instructors; conversely, Sec 2 had only 29% women vs 40%. It is possible that some combination of a selection effect among female students and the impact of being more substantially minoritized played a role in the difference observed here.

3.3 Survey Results

We conducted an optional survey at the end of our courses, seeking students' feedback on their experiences in the course, including the use of Agora. The survey was sent to all students in the Fall 22 and Fall 23 sections (N=309) and received 31 total responses. We

	SD	*D*	*N*	*A*	*SA*	*Total Responses*
Q 1	4	6	0	6	15	31
Q 2	2	3	0	12	8	25
Q 3	6	7	0	5	4	22

Table 6: Survey Results: SD = Strongly Disagree; D = Disagree; N = Neutral; A = Agree; SA = Strongly Agree.

collected both broad and highly specific impressions about Agora via between 4 and 10 questions (varying across years), each using a 5-point Likert scale. We highlight three of the most interesting questions here. In response to a question about how easy they found it to use Agora, 68% of participants either agreed or strongly agreed that the app was easy to use (Q1). To a question asking whether they would have participated less in class without Agora, 80% of participants either agreed or strongly agreed that they would have participated less (Q2). When we asked whether Agora helped students overcome their anxiety of raising hand to talk in class, 41% of students agreed that Agora was helpful for this purpose (Q3).

4 CONCLUSIONS

Agora is an open-source app that encourages students to engage in active learning by incentivizing them to prepare answers to questions under collective class discussion, while also giving them control over when to participate. Students decide when to raise or lower their hands; the system chooses among students who opt in (skipping students who have already spoken in the current class and biasing towards students who have spoken rarely in previous classes) and awards participation credit to all students whose hands were raised. Agora offers a range of additional features designed to aid in its practical deployment. We empirically evaluated the system across three course offerings and found that active participation correlated both with high grades and with grade improvement between the midterm and final; that Agora yielded generally balanced participation ratios between men and women; and that students reported satisfaction with the system in a survey. We note some limitations of our study: notably, our assessment of learning outcomes was correlational rather than causal (ideally, two identical sections would run in parallel, one with Agora and one without); and our assessment of gender was restricted to a binary concept of gender due to our reliance on the registrar's data. Moreover, only about 10% of the students responded to our survey, making it impossible to make a statistically robust argument. In future work, we intend to expand our study by exploring different grading schemes to map participation points to course credit. Additionally, we plan to extend our study on distribution balance to include other underrepresented groups, such as racially minoritized students.

5 ACKNOWLEDGMENTS

This work was funded by an NSERC Discovery Grant, a DND/NSERC Discovery Grant Supplement, a CIFAR Canada AI Research Chair (Alberta Machine Intelligence Institute), and awards from Facebook and Amazon Research. Research use of our data was authorized by UBC's Behavioural Research Ethics Board (BREB #H22-01652).

REFERENCES

[1] S. M. Aguillon, G. Siegmund, R. H. Petipas, A. G. Drake, S. Cotner, and C. J. Ballen. 2020. Gender differences in student participation in an active-learning classroom. *CBE—Life Sciences Education* 19, 2 (2020), ar12.

[2] J. Allsop, S. J. Young, E. J. Nelson, J. Piatt, and D. Knapp. 2020. Examining the Benefits Associated with Implementing an Active Learning Classroom among Undergraduate Students. *International Journal of Teaching and Learning in Higher Education* 32, 3 (2020), 418–426.

[3] E. Bojinova and J. Oigara. 2013. Teaching and Learning with Clickers in Higher Education. *International Journal of Teaching and Learning in Higher Education* 25, 2 (2013), 154–165.

[4] K. A. Boniecki and S. Moore. 2003. Breaking the silence: Using a token economy to reinforce classroom participation. *Teaching of Psychology* 30, 3 (2003), 224–227.

[5] M. Broeckelman-Post, A. Johnson, and J. Reid Schwebach. 2016. Calling on Students Using Notecards: Engagement and Countering Communication Anxiety in Large Lecture. *Journal of College Science Teaching* 45, 5 (2016).

[6] S. D. Brookfield and S. Preskill. 2012. *Discussion as a way of teaching: Tools and techniques for democratic classrooms.* John Wiley & Sons.

[7] M. Camacho-Miñano and C. del Campo. 2016. Useful interactive teaching tool for learning: Clickers in higher education. *Interactive Learning Environments* 24, 4 (2016), 706–723.

[8] K. Canada and R. Pringle. 1995. The role of gender in college classroom interactions: A social context approach. *Sociology of Education* (1995), 161–186.

[9] J. V. Canino. 2015. Comparing student performance in thermodynamics using the flipped classroom and think-pair-share pedagogies. In *2015 ASEE Annual Conference & Exposition.* 26–376.

[10] A. Caspi, E. Chajut, and K. Saporta. 2008. Participation in class and in online discussions: Gender differences. *Computers & Education* 50, 3 (2008), 718–724.

[11] J. M. Cohoon. 2007. Gendered experiences of computing graduate programs. *ACM SIGCSE Bulletin* 39, 1 (2007), 546–550.

[12] G. Crombie, S. W. Pyke, N. Silverthorn, A. Jones, and S. Piccinin. 2003. Students' perceptions of their classroom participation and instructor as a function of gender and context. *The journal of higher education* 74, 1 (2003), 51–76.

[13] K. E. Czekanski and Z. R. Wolf. 2013. Encouraging and Evaluating Class Participation. *Journal of University Teaching and Learning Practice* 10, 1 (2013), 7.

[14] E. J. Dallimore, J. H. Hertenstein, and M. B. Platt. 2013. Impact of cold-calling on student voluntary participation. *Journal of Management Education* 37, 3 (2013), 305–341.

[15] E. J. Dallimore, J. H. Hertenstein, and M. B. Platt. 2019. Leveling the playing field: How cold-calling affects class discussion gender equity. *Journal of Education and Learning* 8, 2 (2019), 14–24.

[16] L. N. Foster, K. R. Krohn, D. F. McCleary, K. B. Aspiranti, M. L. Nalls, C. C. Quillivan, C. M. Taylor, and R. L. Williams. 2009. Increasing low-responding students' participation in class discussion. *Journal of Behavioral Education* 18 (2009), 173–188.

[17] S. Freeman, S. L. Eddy, M. McDonough, M. K. Smith, N. Okoroafor, H. Jordt, and M. P. Wenderoth. 2014. Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the national academy of sciences* 111, 23 (2014), 8410–8415.

[18] C. Gini. 1912. *Variabilità e mutabilità: contributo allo studio delle distribuzioni e delle relazioni statistiche.[Fasc. I.].* Tipogr. di P. Cuppini.

[19] M. Hajhosseini, S. Zandi, S. Hosseini Shabanan, and Y. Madani. 2016. Critical thinking and social interaction in active learning: A conceptual analysis of class discussion from Iranian students' perspective. *Cogent Education* 3, 1 (2016), 1175051.

[20] A. C. Hlas, K. L. Schuh, and S. M. Alessi. 2008. Native and non-native speakers in online and face-to-face discussions: Leveling the playing field. *Journal of Educational Technology Systems* 36, 4 (2008), 337–373.

[21] S. Hood, N. Barrickman, N. Djerdjian, M. Farr, R. J. Gerrits, H. Lawford, S. Magner, B. Ott, K. Ross, H. Roychowdury, et al. 2020. Some believe, not all achieve: the role of active learning practices in anxiety and academic self-efficacy in first-generation college students. *Journal of microbiology & biology education* 21, 1 (2020), 10–1128.

[22] J. R. Howard and A. L. Henney. 1998. Student participation and instructor gender in the mixed-age college classroom. *The Journal of Higher Education* 69, 4 (1998), 384–405.

[23] S. H. Kollat. 2018. The Case for Cold-Calling: Initiating Active Participation in the Classroom. *College Teaching* 66, 3 (2018), 168–169.

[24] K. R. Krohn, K.B. Aspiranti, L. N. Foster, D. F. McCleary, C. M. Taylor, M. L. Nalls, C. C. Quillivan, and R. L. Williams. 2010. Effects of self-recording and contingent credit on balancing participation across students. *Journal of Behavioral Education* 19 (2010), 134–155.

[25] D. Levy. 2014. Cold Calling and Web Postings: Do They Improve Students' Preparation and Learning in Statistics?. *Journal of the Scholarship of Teaching and Learning* 14, 5 (2014), 92–110.

[26] D. F. McCleary, K. B. Aspiranti, L. N. Foster, C. A. Blondin, C. E. Gaylon, J. S. Yaw, B. N. Forbes, and R. L. Williams. 2011. Balancing participation across students in large college classes via randomized participation credit. *The Journal of General Education* 60, 3 (2011), 194–214.

[27] P. Medel and V. Pournaghshband. 2017. Eliminating gender bias in computer science education materials. In *Proceedings of the 2017 ACM SIGCSE technical symposium on computer science education.* 411–416.

[28] K. J. Metzger and Z. Via. 2022. Warming Up the Cold Call: Encouraging Classroom Inclusion by Considering Warm-& Cold-Calling Techniques. *The American Biology Teacher* 84, 6 (2022), 342–346.

[29] J. Michael. 2006. Where's the evidence that active learning works? *Advances in physiology education* (2006).

[30] R. M. Millis, S. Dyson, and D. Cannon. 2009. Association of classroom participation and examination performance in a first-year medical school course. *Advances in physiology education* 33, 3 (2009), 139–143.

[31] A. Quinn. 2010. An exploratory study of opinions on clickers and class participation from students of human behavior in the social environment. *Journal of Human Behavior in the Social Environment* 20, 6 (2010), 721–731.

[32] K. A. Rocca. 2010. Student participation in the college classroom: An extended multidisciplinary literature review. *Communication education* 59, 2 (2010), 185–213.

[33] J. M. Rogers. 1997. Class participation: Random calling and anonymous grading. *Journal of Legal Education* 47, 1 (1997), 73–82.

[34] M. Sadker and D. Sadker. 2010. *Failing at fairness: How America's schools cheat girls.* Simon and Schuster.

[35] K. Schnitzler, D. Holzberger, and T. Seidel. 2021. All better than being disengaged: Student engagement patterns and their relations to academic self-concept and achievement. *European Journal of Psychology of Education* 36 (2021), 627–652.

[36] R. Sommer and B. A. Sommer. 2007. Credit for comments, comments for credit. *Teaching of Psychology* 34, 2 (2007), 104–106.

[37] K. M. Soria and M. J. Stebleton. 2012. First-generation students' academic engagement and retention. *Teaching in higher education* 17, 6 (2012), 673–685.

[38] H. E. Tatum, B. M. Schwartz, P. A. Schimmoeller, and N. Perry. 2013. Classroom participation and student-faculty interactions: Does gender matter? *The Journal of Higher Education* 84, 6 (2013), 745–768.

[39] C. Whitehead and L. Ray. 2010. Using the iclicker classroom response system to enhance student involvement and learning. *Journal of Education, Informatics and Cybernetics* 2, 1 (2010), 18–23.

[40] M. Younger, M. Warrington, and J. Williams. 1999. The gender gap and classroom interactions: reality and rhetoric? *British journal of Sociology of Education* 20, 3 (1999), 325–341.