Exploring Power Distance, Classroom Activity, and the International Classroom Through Personal Informatics

David Gerritsen, John Zimmerman, Amy Ogan

Human-Computer Interaction Institute,
Carnegie Mellon University, Pittsburgh, USA
{dgerrits, johnz, aeo}@cs.cmu.edu

Abstract. Research shows the benefits of active learning in American college classrooms. International graduate students in American universities may face difficulties in teaching students with different cultural dispositions. The current research uses power distance to explore cultural juxtapositions in classrooms and personal informatics design to propose an adaptive system for cultural acquisition. The work shows that even though instructors are aware of the distinctly Western value of speaking up in class, they do not employ it in their own classes. They show surprise at the amount of time they spend lecturing, but they express ambivalence about the importance of vocal contributions from the students. We describe a technical system design that supports the development of cultural fluency by providing ITAs with feedback such as visualizations of time spent lecturing and suggestions for strategy selection in culturally challenging scenarios. The system would reflect changes in classroom activity over time as a way for TAs to reflect on their own professional development.

Keywords: Power distance, international teaching assistants, classroom activity, personal informatics

1 Introduction

Research in the learning sciences has recently produced an explosion of experimental evidence that college students benefit from less lecture and more student activity. This evidence exists even for content-heavy science, technology, engineering, and mathematics (STEM) classes where instructors have traditionally emphasized the importance of covering and memorizing facts rather than exploring, curating, and constructing knowledge. Most of these studies have taken place in American classrooms and have not addressed questions of cultural dimensions of learning and teaching. Meanwhile, the number of international graduate students teaching introductory STEM classes in American universities continues to grow. These students tend not to have experienced the cultural shift toward active learning and its concomitant decrease in social distance to figures of authority that is familiar to most students from the U.S. This can lead to challenges for international graduate students in the U.S. when they are required to teach American students.

The CATS community has a history of developing systems to improve education and cultural awareness. We build on this line of research by focusing on new design methods that frame the instructor as both the learner and the agent of change in the classroom. Using methods from *Personal Informatics* (PI), we explore the state of international teaching assistants (ITAs) leading STEM classes in an American university, and propose a system that potentially simplifies the implementation of active learning in order to more fully engage students.

PI is an approach to behavior change and maintenance that gathers user data and generates digital artifacts for reflection, such as visualizations of change toward a behavioral goal. Very little research has looked at its value in education, and none has attempted to use it to better understanding culture. It incorporates methods of contextual design and development that may be valuable in improving educational outcomes while investigating culturally adaptive interactions.

To assess the feasibility of this line of research and development, we carried out several overlapping activities: classroom observation of ITAs in action in order to understand the context need for adaptive instruments, surveys and interviews in order to understand how ITAs might make sense of classroom behavior, and data visualization feedback for ITAs in order to understand and explore the potential interface for a PI system. Finally we constructed and evaluated a prototype classroom detection system to investigate if we could sense relevant behaviors.

We confirmed that ITAs' knew of the cultural value of classroom activity, yet their recitations were almost completely based on lecture, with little student participation. They were open to more classroom activity, but with some reservations. They shared an interest in monitoring their teaching behaviors and aligning their performance with expert models. Also, our technical system functioned with 85% accuracy. We propose that these findings support further investigation of PI methods for investigating and supporting the acquisition of cultural fluency in unfamiliar educational contexts.

2 Background

Several decades of research in U.S. higher education has produced a wealth of studies showing the benefits of active learning compared to passive lecture and fact memorization [1, 2, 3, 4]. These studies have investigated and advocated active learning tactics such as *think-pair-share* and *cooperative learning*, showing that students improve academically, socially, and psychologically [1, 4]. Like most education research, the studies tend not to include considerations of cultural dimensions of learning. Cultural dimensions of instructors and learners in American universities are poorly understood. Given the evidence that different cultures have different valuations of student activity in the classroom [5, 6], the call for increased student participation may create a tension when it fails to address how international instructors perceive and value active learning practices. This situation deserves attention as the number of international graduate students teaching STEM classes in the U.S. continues to grow [7].

One way to orient the conversation about cultural differences in praxis is to frame it in terms of power distance [8, 9]. Higher and lower national indices of power dis-

tance (PDI) attempt to describe the level of deference that individuals express toward members of higher and lower social status. Given the long history of measurable social distance between Asian students and American instructors [5, 6], power distance is a reasonable construct with which to study classroom practices. It seems to have a direct mapping to the differences students exhibit as a function of cultural orientation to learning [9]. A low PDI score of 40 in the U.S., compared to 77 and 80 in India and China [8], may partially explain these students' general tendencies to speak or remain silent when they attend American university classes, regardless of how well they know the material [6].

This distance is becoming increasingly important to address. International enrollment to American graduate schools has grown since 2005, with the most recent report showing a 17% jump in enrollment to engineering schools and a 40% increase in graduate students from India [7]. These students often fund their education by teaching small classes that act as a supplement to large introductory STEM courses. These small classes, normally called *recitations*, allow groups of undergraduates from a large class to review course material and interact more closely with each other and an expert instructor.

Although many states require ITAs to pass an oral proficiency exam before teaching, there is little support for developing cultural fluency (or even general teaching skills). In other domains, such as health and finance, PI has recently emerged as a technique for motivating changes in behavior [10–13] with only a small investment of time or conscious effort on the part of the user. It is a new class of socio-technical system based on self-monitoring through data visualization [14]. The process helps motivate people to make new decisions by increasing their awareness of behaviors that are normally obscure and hard to observe, such as encouraging more activity by showing people a record of how much (or how little) they move throughout the day. That awareness is a critical step in the process of making changes [12]. These systems have gained popularity due to advances in wearable technology and smartphones. Current PI systems can track a user's number of steps [10], hours and quality of sleep [15], levels of glucose in relation to food intake [16], consumption of non-renewable goods [17], and many more important activities that are hard to monitor without technological assistance.

Research investigating how people use and make sense of PI systems produced a five-stage model of behavior change that applies to a large number of general cases [14]. The model (Preparation, Collection, Integration, Reflection, and Action) describes the types of data users collect, the integration of data collection and reflection into a daily routine, and the transition from reflection to goal setting. The framework provides a list of barriers and design recommendations for each stage. Researchers have recently proposed that incorporating this framework into adaptive training systems may improve classroom interactions [18], but only one project has evaluated such an application. The Live Interest Meter is a PI system that tracks student engagement through a mobile app and provides data visualization to the instructor. It shows the potential to increase audience engagement and instructor responsiveness [19], but at the cost of increased cognitive demand by relying on live manual data input. Our system advances this work by investigating automatic detection of the

presence of classroom features that may indicate enhanced learning, such as peer-to-peer interaction and student participation, both of which have been shown to correlate with students' critical thinking in American universities [20], and both of which would likely be difficult for cultural non-natives to enact in their classrooms [21]. Additional strategies for involving students include the use of student names, asking students to elaborate on ideas, and asking deep questions [22].

AIED work has addressed professional development for teachers by means of student tracking and data visualization [23, 24], but these systems have focused on online learning or blended classrooms, and did not offer instructors guidance on how to enact change in a live classroom. Other systems have attempted to visualize student participation (e.g., [25, 26]), but these have been deployed to support students' own self-reflection rather than to support the instructor, and only in online applications where participation can be tracked through clickstream data.

In our work, we advance the state of the art by focusing on the instructor as the primary agent of change. We focus on student participation in class as an achievable goal that is likely to provide academic benefits to students and cultural fluency for ITAs. The current stage of the work includes classroom observations and iterative phases of design for the adaptive system. Specifically, we wanted to answer the following research questions:

1. Do ITAs from a culture with a high PDI encourage active classrooms?
2. Are ITAs open to adapting their teaching style to an unfamiliar cultural context?
3. Are ITAs open to using PI to set and reflect on goals for their teaching?
4. Can we easily and inexpensively sense and create visualizations of classroom activity in terms of TA and student interactions?

3 Method

To answer the research questions, we recruited 5 ITAs, observed them teaching, issued surveys, conducted interviews, and showed them visualizations of their classroom data. We also developed a prototype technical system to detect instructor talk, student talk, and silence.

The TAs were all from India, male, and in their mid-twenties. India has a relatively high PDI (77) compared to the U.S. (40). Each TA had similar levels of teaching experience and content knowledge. None of them had received pedagogical training by the institution or the professor in charge of the course. We observed six to seven sessions of each TA's weekly course, a sophomore level computer science recitation, for a total of 32 sessions. We logged behaviors that would adduce attempts to engage active learning. We inferred activity from frequency and duration of student talk, as opposed to TA talk and silence. We logged the time and locus of all spoken contributions in order to extrapolate episodes of discussion vs. passive lecture.

We surveyed and interviewed the ITAs about their teaching experiences in and perspectives on American classrooms. The survey collected theoretical orientations toward cultural dimensions of learning via items such as demographics, definitions of terms (e.g., "classroom contribution"), and perceived locus of responsibility for learn-

ing (e.g., instructor, student, or a combination). We met with each TA three times during the semester (totaling 2.5 – 4 hours per TA) to discuss their survey responses, their perspectives on and motivations for teaching, and to explore their own teaching behaviors with data visualizations.

The data visualizations were initial sketches of what might exist in a PI system. These were meant as a probe for discussion that allowed TAs to reflect on the behaviors they most wanted to capture and view. This is a common technique in the design of new computing systems when there are no design patterns or social conventions to inform the design space [i.e., 27]. We gathered reactions to the visualizations, and redesigned them after each round of feedback. We also probed TAs on their willingness to try new teaching techniques, such as praising students, using students' names, encouraging elaboration, and asking difficult questions. To analyze the results we transcribed the interviews and iteratively searched for areas of strong agreement and disagreement amongst the participants' comments.

Finally, we developed an initial prototype system for a feasibility study, following a typical user-centered design process. We synthesized a set of system needs from the observations and interviews and proposed a minimal set of detection requirements. We developed a prototype system with two Microsoft Kinects and tested it with 20 students and a 60-minute lecture that included various kinds of classroom talk. We hand-coded the audio data with discrete categories of *instructor talk*, *no talk*, and *student talk*. Periods when students talked simultaneously were coded as *student talk*. We tested these categories against the Kinect's angle detection, confidence calculation, and audio amplitude, i.e., whether or not the device picked up sound and if so, where in the room it originated.

4 Findings

Exploring the presence of classroom activity, we observed that ITAs conducted nearly all recitation sections as lectures covering a subset of slides from the most recent primary course lecture. Instructor talk dominated the class, taking up 91.97% of class time (SD=3.6%). Student talk took up only 5.25% of class time on average (SD=2.3%), and the length of their contributions averaged 6.2 seconds (Median=3.4, SD=12.6). The most common prompt for student participation was to ask the class, "Do you have any questions?" The resulting patterns of speech were as follows:

1. TA-talk | silence | TA-talk
2. TA-talk | silence | Student-talk | TA-talk
3. TA-talk | silence | Student-talk | Student-talk

TAs were the first to speak after 85% of their pauses (SD=.088) (pattern 1). 13% of the time (SD=.088) students responded, followed by the TA again (pattern 2). These student contributions were typically brief. 2% of the time (SD .02) a different student contribution followed immediately from a prior student (pattern 3).

Student-student interactions were rare. From an active learning perspective, these interactions are useful as students build on each other’s ideas. These conversations

were typically animated discussions of the course content that took place in the few minutes before class began. TAs usually called a stop to such interactions in order to begin the lecture, and over the course of the semester most students stopped talking as soon as the TA entered the room. This matched an overall pattern of decreasing student talk (and attendance) for most classes over the semester.

ITAs did express that student participation was important to them, but they defined participation as *students asking or answering questions*. They used that information for diagnosis. TA-2: "*If you … don't answer [a question asked by the instructor] there is no way for a teacher to know whether you are understanding what he is teaching or what is going on*." Nevertheless, the TAs made lecturing their priority, and student questions were a distraction from this goal. TA-5: "*Maybe I might want to involve their participation a bit more than what it is, but I also fear by doing so [that I won't] be able to complete the contents*."

To explore ITAs' positions on the cultural dimensions of the American classroom, we asked about their explanation for student silence (pattern 1). They speculated that students already understood the content, only had specific questions about their own work, feared appearing dumb, or that they would rather check with peers. When asked how one might increase participation, there were two types of response: ask students if they have questions (TA-1: "*Probably I should ask more times if they have questions*."), and push student to respond to recall questions (TA-3: "*I'll say … at least take a guess … I'm sure that one of them will say something*.").

Viewing visualizations of their teaching helped to assess the TAs' stance toward adopting new cultural strategies. At times these graphs triggered immediate motivation for change. When TA-1 saw he talked 99% of the time in the preceding class (Fig. 1), he shared that an interactive class was important to him and that he wanted to include the students more. Yet when he later viewed four weeks of data revealing that he never spoke less than 95% of the time (Fig. 2), he became frustrated with the students. "*I would prefer if the class had more [student participation]. I keep asking if there are any questions, but no one speaks so, I cannot help this one*."

We probed TAs about their attitudes toward culturally specific strategies for teacher-student interaction. TAs generally agreed that lengthening the pause after asking students a question might be useful and expressed a familiarity with the idea. They showed interest in the tactic of pausing after a student stops talking, and were surprised that it might be valuable. When asked about asking students to elaborate, they expressed skepticism, sharing that students should only elaborate when the instructor does not understand them. We probed them on asking students deep questions from course content as opposed to simple recall questions. This met with mixed reactions. Most worried that asking hard questions would reduce the time needed to cover the material, and all were reluctant to slow down class. TA-5 described his technique of asking content questions in order to highlight important concepts, but only when the questions could be answered rapidly.

We raised the idea of calling on students by their name and of praising their contributions as approaches to create a supportive environment for student participation. Most TAs agreed that these ideas would help students feel valued and might improve their confidence in the learning process, but none of them were willing to employ

these techniques. They worried that they might call a student by the wrong name and feel embarrassed, or that calling on a student directly might make them feel picked on. TA-2 shared that calling on specific students would point out that the student had not been speaking and that this might generate shame.

After looking at many visualizations of their classroom behaviors, including talk time, distribution of student participation, number of unique speakers per class, proportions of each event type per class, changes in rates across multiple classes, timelines of event types, and more, almost all TAs expressed an interest in eliciting more student talk, but each spoke about wanting explicit goals for different behaviors. How much is the *right* amount of student and TA talk? How long should the TA wait after asking a question? Are enough of the students participating? Most also asked how their individual data compared to the other TAs in the course. They were all open to the idea of using a PI system to empirically answer these kinds of questions.

Finally, as a first technical step towards a PI system, we built a prototype detector for speaker events meant to identify three states of classroom discourse that would indicate interesting patterns of events when viewed in sequence: (i) instructor speaking (in front of class), (ii) student speaking (from seats), and (iii) no one speaking for at least one second. Researchers have previously had success using microphone arrays for speaker localization [e.g. 28], a process that triangulates the angle of a noise source in relation to microphones placed in a line (the array). We chose to use the Microsoft Kinect, an inexpensive commodity device with a robust microphone array, a developers' kit, and a support community for software development.

In our 60-minute test of various kinds of classroom talk, we evaluated the accuracy of a single Kinect on one side of a classroom and the inclusion of a second Kinect at the front of the room facing the students. We used a *Nominal Logistic Fit for Categories* test (JMP V.10.0) with standard output from the device (angle detection and confidence), and were able to discriminate between students and the instructor with high accuracy (Table 1). We expanded the test to also detect silence by including average amplitude for each second of recorded audio as an input variable. This reduced accuracy overall, but much of that loss was amended by the inclusion of a second Kinect.

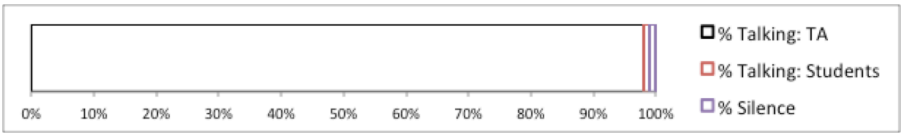

Fig. 1. TA-1's first day of recorded data.

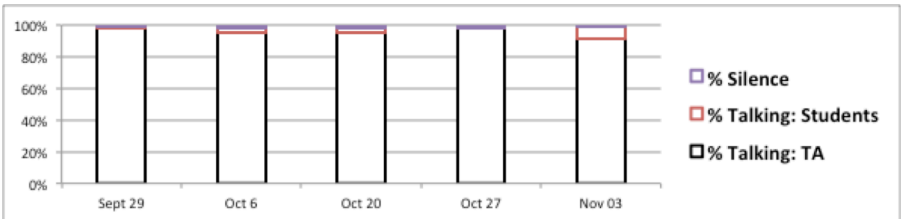

Fig. 2. Aggregate talk time for TA-1 across four classes.

Table 1. Accuracy of Kinects detecting instructor talk, student talk, and no talk

	1 Kinect	2 Kinects
Student/Instructor	94.78%	95.36%
Student/Instructor/Silent	77.70%	85.44%

5 Discussion and Conclusion

Our research explored classroom activity in a university STEM course taught by graduate students from a country with a PDI higher than the host country. We used design methods from PI to better understand the perspectives of ITAs who teach in an unfamiliar cultural context. This process led to the development of a prototype system for identifying levels of classroom activity based on speech events that could indicate higher order discourse phenomena. Our findings suggest that ITAs and their students may benefit from an adaptive feedback system built on measuring levels of classroom activity, and that international instructors would be open to using such a system.

ITAs were open to varying degrees of active learning techniques in their own classrooms. Some were easy for them to imagine using (e.g., pausing after students talk), and others were harder to accept (e.g., asking for elaboration). They showed reluctance to decrease the amount of time spent "covering" critical course material, yet they all valued when students got involved in the lecture. These tensions are clues that an adaptive system for cultural training may need do more than measure and report on behavior, but also provide scaffolding for implementing relatively low-cost active learning strategies, such as *think-pair-share*. The next step would be to assess the user's knowledge and stance toward different contextual behaviors and provide individualized instruction and adding more advanced scaffolding prompts as the TA becomes ready for them. Future research would need to navigate this complex space. To refine the detection system further and more easily differentiate between user states, it would be possible to include machine learning and more factors than we currently use, such as Kinect error rates, classroom details, pitch fluctuations and filters, and so on. With more tuning the system might identify individual speakers, leading to reflection opportunities based on individual student speaking patterns. Turn detection at this level could point out disproportionate properties of classroom talk, such as a group of dominant speakers.

There are aspects of the classroom that the proposed system would not be able to detect. ITAs were curious about whether they had lectured for "too long." They made reasonable requests, such as seeing when they had made a "good" explanation, or if students understood the material. A fully operational PI system would necessarily need supplemental human input to provide such feedback, which is already standard practice in current systems: much like annotating the quality of a recent jog when using a fitness-tracking app, our proposed system could request post-class assessments from students or the TA. Some TAs remarked that it would be a simple procedure to personally label the broad topic of the class, or the context of specific pauses throughout the lecture if they were able to review the data and access the audio. Alt-

hough previous PI systems have not explored user input this deeply, such interactions would be possible to implement, and may be critical for system design.

Our study only observed one genre of recitation, but there are many others. It is critically important to assess how much the observed behaviors in this study were an artifact of culture, context, or simply being new to teaching. In our current work we are performing additional observations of a broad selection of classroom contexts taught by students from many different cultural backgrounds in order to assist in making these distinctions.

Research in professional development for teachers might note that our work did not address the quality of interactions, but only quantity and abstract patterns of discourse. As a first step, we argue that any increase in student talk would more closely align with the cultural context of the U.S. classroom, although in the future quality may prove to be a critical area of investigation. Currently, however, the space of cultural acquisition for graduate students and the professional development of novice instructors is under-investigated, and thus this early work makes a contribution.

The implications of this research are important in their potential to address the lack of research in supporting the cultural fluency of ITAs in a challenging new environment. Our work shows preliminary evidence that PI could be an approach to support reflection on classroom dynamics and an opportunity to adaptively expand an instructor's set of pedagogical tools. The impact of the work points to a better experience for international graduate students and potentially better learning for their students.

6 Acknowledgements

This work is supported in part by Carnegie Mellon University's Program in Interdisciplinary Education Research (PIER) funded by grant number R305B090023 from the US Department of Education.

7 References

1. Faust, J. L., Paulson, D. R.: Active learning in the college classroom. Journal on Excellence in College Teaching, 9(2), 3–24 (1998)
2. Michael, J.: Where's the evidence that active learning works? Advances in Physiology Education, 30(4), 159–67 (2006)
3. Lujan, H. L., DiCarlo, S. E.: Too much teaching, not enough learning: what is the solution? Advances in Physiology Education, 30(1), 17–22 (2006)
4. Freeman, S., Eddy, S.L., McDonough, M., Smith, M.K., Okoroafor, N., Jordt, H., Wenderoth, M.P.: Active learning increases student performance in science, engineering, and mathematics. Proc. Natl. Acad. Sci. U. S. A. 111, 8410–5 (2014)
5. Nisbett, R.: The Geography of Thought: How Asians and Westerners Think Differently... and Why. Simon and Schuster. (2010)
6. Liu, J.: Asian students' classroom communication patterns in US universities: An emic perspective. Greenwood Publishing Group. (2001)
7. Allum, J. R.: Findings from the 2013 CGS International Graduate Admissions Survey Phase III : Final Offers of Admission and Enrollment, 20036 (2013)

8. Hofstede, G., Hofstede, G. J., & Minkov, M.: Cultures and organizations: Software of the mind: 3rd edition. McGraw-Hill Professional. (2010)
9. Wang, M.: Designing online courses that effectively engage learners from diverse cultural backgrounds. British Journal of Educational Technology, 38(2), 294–311 (2007)
10. Lin, J.J., Mamykina, L., Lindtner, S., Delajoux, G., Strub, H.B.: Fish'n'Steps: Encouraging physical activity with an interactive computer game. UbiComp 2006. pp. 261–278. Springer-Verlag (2006)
11. Hsieh, G., Li, I., Dey, A., Forlizzi, J., Hudson, S.E.: Using visualizations to increase compliance in experience sampling. Proc. UbiComp '08. 164 (2008)
12. Li, I., Dey, A.K., Forlizzi, J.: Using context to reveal factors that affect physical activity. ACM Trans. Comput. Interact. 19, 1–21 (2012)
13. Levy, L., Jones, B., Robertson, S., Price, E.D.: Health Mashups: Presenting Statistical Patterns between Wellbeing Data and Context in Natural Language to Promote Behavior Change. ACM Trans. Comput. Interact. 20, (2013)
14. Li, I., Dey, A., Forlizzi, J.: A stage-based model of personal informatics systems. CHI 2010. pp. 557–566. ACM, Atlanta, Georgia, USA (2010)
15. Kay, M., Choe, E.K., Shepherd, J., … Kientz, J.A.: Lullaby: A Capture & Access System for Understanding the Sleep Environment. UbiComp '12. ACM, Pittsburgh, USA (2012)
16. Mamykina, L., Mynatt, E.D., Kaufman, D.R.: Investigating health management practices of individuals with diabetes. Proc. CHI '06. 927 (2006)
17. He, H.A., Greenberg, S., Huang, E.M.: One size does not fit all: Applying the transtheoretical model to energy feedback technology design. CHI 2010. 927–936 (2010)
18. Muller, R., Long, Y., Koedinger, K.R.: Personal informatics for self-regulated learning. In: Personal Informatics in Practice: Improving Quality of Life Through Data — CHI 2012 Workshop. CHI, Austin, TX (2012)
19. Rivera-Pelayo, V., Munk, J., Zacharias, V., Braun, S.: Live interest meter: Learning from quantified feedback in mass lectures. Proc. 3rd Int. Conf. on Learn. Anal. and Knowledge. 23–27. ACM, Leuven, Belgium (2013)
20. Smith, D.G.: College classroom interactions and critical thinking. J. Educ. Psychol. 69, 180–190 (1977)
21. Gerritsen, D., Zimmerman, J., & Ogan, A.: A theoretical approach to the development of critical incidents for cultural training. 5th Int. Workshop on Culturally-Aware Tutoring Systems. 51–60 (2014)
22. Nunn, C. E.: Discussion in the college classroom: Triangulating observational and survey results. J. of High. Educ., 243–266 (1996)
23. Mazza, R., Milani, C.: Exploring Usage Analysis in Learning Systems: Gaining Insights From Visualisations. AIED'05 Work. Usage Anal. Learn. Syst. 65–72 (2005)
24. Scheuer, O., Zinn, C.: How did the e-learning session go? The Student Inspector. Proc. Conf. Artif. Intell. Educ. 487–494 (2007)
25. Janssen, J., Erkens, G., Kanselaar, G., Jaspers, J.: Visualization of participation: Does it contribute to successful computer-supported collaborative learning? Comput. Educ. 49, 1037–1065 (2007)
26. Smith, M.K., Vinson, E.L., Smith, J. a., Lewin, J.D., Stetzer, M.R.: A Campus-Wide Study of STEM Courses: New Perspectives on Teaching Practices and Perceptions. Cell Biol. Educ. 13, 624–635 (2014)
27. Davidoff, S., Lee, M., Dey, A., Zimmerman, J.: Rapidly Exploring Application Design Through Speed Dating. UbiComp 2007. 429–446 (2007)
28. Valin, J., Rouat, J., Dominic, L.: Robust Sound Source Localization Using a Microphone Array on a Mobile Robot. Proc. IEEE/RSJ (IROS); p. 1128–1233. (2003)