

The Student Prompt: Student Feedback and Change in Teaching Practices in Postsecondary Computer Science

Lecia Barker and Jane Gruning

School of Information

University of Texas at Austin

Austin, Texas, USA

lecia@ischool.utexas.edu, jane.gru@gmail.com

Abstract—Research on retaining students in computer science is clear: pedagogical and curricular interventions that are experienced by all students are the most effective. Yet understanding what brings faculty to modify their teaching practices is scant. This study presents findings from a study based on interviews with 66 computer science professors and instructors in the U.S. Faculty described various sources of student feedback that prompt them to change their teaching practices. Though not originally an interview topic, every instructor mentioned ways in which student feedback instigated change. Faculty described several sources of feedback that influence teaching choices, with the goal of trying to improve their teaching. These include formalized evaluations; feedback inferred from student performance; direct requests from students; and students' nonverbal behavior. As sources of evidence for making decisions about teaching practices, each of these is problematic. We describe what makes them problematic and offer solutions described by faculty interviewed.

Keywords—*change in pedagogy; teaching change; student feedback; student evaluations of teaching*

I. INTRODUCTION

What brings postsecondary faculty to change their teaching practices? The answer to this question is important for a number of reasons. For one, technology is changing the possibilities of teaching and learning. Online courses— notwithstanding the mythology around massive open courses (MOOCs)[1]—are the teaching and learning medium of choice for many faculty and students. In the effort to reduce costs of traditional university settings and to reach students who would not otherwise be able to take a course, distance courses are being offered and taken at increasing rates. In face-to-face courses, a variety of teaching techniques and information technologies are used to enhance learning and increase student engagement. In addition, government agencies and foundations have provided a great deal of funding for the development and evaluation of a variety of teaching innovations, not limited to technology. Communicating the availability of teaching techniques does not necessarily translate to adoption. Another critical reason for studying change in faculty teaching practice is research on retention: a robust body of empirical studies demonstrates that classroom teaching

practices are among the strongest predictors of retention of students in computer science, particularly women [2]–[13]. Applying research on what brings faculty to change their teaching practices could increase women's representation in this important field.

Research and theory that addresses what brings about teaching change generally focus on institutional factors and the characteristics of individuals. Institutional factors include organizational structure, culture, climate, and the reward system [14], [15]. Individual characteristics include awareness of and receptiveness to new ideas, values and attitudes related to specific innovations, perceptions of costs and benefits relative to existing practices, and attitudes toward gender issues [16]–[19]. Not generally addressed in the literature are social factors, such as the co-learning that occurs when faculty discuss teaching with each other locally and at conferences. Another important factor is the social nature of learning environments, including student-student and student-faculty interaction.

Although we did not explicitly ask faculty how or whether students influence their teaching choices, every one of our 66 interviewees described decisions about what they do in the classroom based on different forms of student feedback. Student feedback is central to a professor's understanding of the relationship between what they do as a teacher and how students engage with and learn (or fail to learn) the course material. Faculty use student feedback as evidence for decision making, asking themselves, 'should I change an aspect of instruction or class content or not?' Understanding the many ways that students prompt faculty change is an underexplored area of faculty adoption of teaching practices. This paper takes a close look at how faculty describe student influence on their teaching practices in undergraduate computer science classes taken early in the curriculum. We chose early classes because we are most interested in faculty choosing to teach in ways that retain students. In early classes, students can still change to different fields of study with little effect on time to degree. Below we briefly review existing literature on student feedback, describe our methods, and discuss our findings. We close by describing two interviewees' methods of gathering effective feedback.

II. REVIEW OF LITERATURE ON STUDENT FEEDBACK

Research on student feedback regarding teaching practices has almost exclusively focused on one type of formal feedback mechanism, student evaluations of teaching (SETs). SETs usually are distributed to all students in class on paper or online and ask students to rate closed-response items about multiple dimensions of teaching. They generally also include space for students to add open-ended comments. The distributions and means of all closed-response items are compiled, then provided to faculty with the open-ended comments. Not all students respond to SETs. Response rates of questionnaires administered either in class or online are likely to represent the opinions of only 43 to 60 percent of the students respectively, and sometimes as few as 20 percent [20]. Therefore, depending on the class enrollment, the overall ratings instructors receive can be based on a non-representative sample. Response rates for online SETs are generally lower than in-class administration, though mean scores may not be different from paper administration [20]. This depends on whether enrollment is high enough for those who respond to be considered a large enough sample for inferences to be drawn.

Research on SETs is more often focused on exploring the reliability and validity of the closed answer responses to questionnaires. Researchers disagree on the reliability and validity of SETs for representing quality of instruction, with some arguing that they are good measures and others arguing they are not [21]–[23]. Some argue that student ratings of instruction are unrelated to whether an instructor has produced learning and that instructor personality is highly correlated with perceptions of having learned [24]. The quality of the ratings depends on many factors, but especially on construct validity, or the degree to which an item measures what it is intended to measure [22]. In a review of the literature, Kulik argues that the more students believe they have learned from an instructor, the more highly the instructor is rated [23]. Perceived learning is not the only indicator of quality on SETs, however. Others have argued for the importance of distinguishing between student satisfaction (usually translated as whether students enjoy a class) and student achievement or learning [25]. These are distinct aspects of what happens in a class and should be examined separately to determine whether a rating suggests changing teaching practices. For example, faculty may make changes to improve the social climate or to increase student engagement in classroom discussions without explicit understanding of whether the changes lead to greater learning. However, results from studies examining whether faculty make changes to their teaching based on student answers to closed-response portions of SETs are mixed [26], [27]. Closed-ended ratings may be too abstract to be helpful.

In contrast, faculty seem to find open-ended comments useful. A study at a private, mid-sized university found that among 436 faculty respondents to a survey, 43 percent reported that they make changes in classes based on students' comments [26]. Other studies support this finding [28]–[30]. Written comments are considered more detailed and informative and may offer ideas for specific

improvements to classes, suggesting that the open-ended comments have more likelihood of influencing teaching practices than scale items [28]. A major drawback to relying on written comments, however, is that fewer students contribute written comments in SETs than fill out the quantitative portion of the survey. Various studies have reported that as few as 10 percent and up to 70 percent of students supply open-ended comments [31]–[34]. Although our findings are consistent with these studies, our interviewees talked about several other sources of student feedback on which they based teaching changes.

III. SAMPLE PROFILE AND RESEARCH METHODS

This qualitative study employed a multi-method approach to studying faculty teaching practices and change within their particular institutional contexts. Qualitative studies are useful for developing an in-depth, nuanced understanding of the nature and motivations of human behavior and how these are influenced by situations. Multi-method approaches allow the collection of data from different viewpoints, in this case the self-reported practices of faculty in interviews and descriptions of learning environments and their student and faculty inhabitants by trained observers.

A. Participant Profile

The findings are based on interviews conducted with 66 computer science faculty members teaching early, “gatekeeper” courses in 36 postsecondary institutions across the United States. The institutions chosen for the study were deliberately diverse: research-intensive, women-only, U.S. minority-dominant, public, private, elite, two-year, four-year liberal arts, and technical institutes. They were also geographically diverse. Interviewees' job types were equally diverse, ranging from doctoral student teaching assistants, part-time and full-time teaching faculty, tenure-track faculty in teaching and research institutions, to heads of departments. This wide range of institution and faculty job types provided deep, contextualized understanding of how faculty think about and implement change in teaching.

B. Data Collection: Interviews and Observations

We used a semi-structured approach to both interviewing and observation. Semi-structured interviewing stimulates informants to talk about particular topics, but also allows them to express what is of primary importance to them. Based on empirical studies and a theoretical framework of adoption of teaching practices in higher education, the interview protocol asked faculty to describe their teaching experience and background, how long they had been teaching, class sizes, how they understood their role as teachers, and teaching practices they had used, past and present, and why. Interviews lasted between 43 and 107 minutes, averaging 71 minutes. Two interviews were conducted with focus groups of three faculty each to probe the issue of student feedback as discussed here.

Observations accompanied 51 of the 61 interviews, but are used here only for contextualization. The purpose of observations was to provide deeper understanding of what faculty mean when they say they use a teaching practice or

approach, a deeper understanding of student behavior in the classroom, technology available in classrooms, and the physical environment of the classroom that can influence teaching choices (e.g., chairs bolted to the floor make collaborative problem solving difficult).

C. Data Analysis

All but four of the 61 interviews were recorded and transcribed. Detailed, typed notes were taken of all interviews, including the four that were not recorded. Transcripts and observation notes were loaded into Dedoose qualitative research software. Each of the authors listened to and read the transcripts multiple times to identify patterns in the data. Patterns were developed into categories, and excerpts of transcripts were coded by these categories. As we listened to recordings and read transcripts, we were struck by the number of participants who talked about interactions with or observations of students as well as comments on student evaluations of teaching when talking about motivations to change teaching practices.

IV. FINDINGS

A. The Goals of Change: Learning and Engagement

Faculty change their teaching practices primarily for two interrelated reasons, improved learning and increased engagement. Faculty can make judgments about learning based on assessments, but judgments about student engagement are often based on faculty observation of student behavior in class. Faculty believe that increased engagement is important for motivating students to learn. For example, a professor who introduced robots into an introductory programming course said,

“The reason for using robots is not that I want them to become roboticists. Right? The whole reason why I am using a robot is that I see that it’s getting them excited. I see that it’s getting them into classes, getting women into our classes. And I see that they’re staying. Because they’ve seen what computer science could do or they’re excited about it, they want to do more things. And one very telltale sign of it is how many assignments they do that they weren’t asked to do. They will just sit there and voluntarily program more things because they want the robot to do this and that, even if I didn’t ask them to do that. So it’s, it’s a motivation thing.”

Faculty associate engagement with excitement and desire to learn and thus an important goal of changing teaching practices. In the analysis below, we describe student-prompted faculty change. First we describe sources of feedback, then discuss how those sources are problematic for evidence-based decision making.

B. Sources of Student Feedback

Four primary sources of student feedback were described by our interviewees. These are formalized evaluations; feedback inferred from performance; direct requests from students; and nonverbal behavior. We define these and give examples from interviews of how faculty used them in the following sections.

1) Formalized Evaluations

Formalized evaluations include student evaluations of teaching (SETs), but SETs are not the only formal method. This feedback is actively sought by faculty, and students may be anonymous or may be identifiable. Formalized evaluations are part of the structure of the class and intended to be broadly representative of student opinion. This may be the case when questions are standardized, students are anonymous, and the sample of total students is large and representative. But when a formal evaluation is conducted as a class discussion, faculty are less likely to hear students’ true opinions, either because of attitudes toward authority, because students like the instructor, or other reasons.

Faculty often cited student evaluations as evidence that a change was needed or had been successful or unsuccessful. Discussing why she decided to switch the mode of instruction in her large-lecture class, a professor said, “my evaluations went way down from when I taught during the summer with a small class to the big class. I wasn’t as happy. Students weren’t as happy.” Another professor noted that after changing to an inverted classroom approach [35], his evaluations improved, saying, “This is the first time when I got all fours in all the categories. Four being the highest. They loved the class. And what did they like about it? There was no intimidation anymore. ‘I’m going with my own pace.’” However, because overall ratings on evaluations are often a measure of whether students liked the class, rather than of learning, the use of student ratings may be a poor source of evidence for deciding to change. While it is much easier to make a change that students like than one that they don’t like it is important to distinguish student satisfaction from student learning.

Many faculty reported using standardized evaluations as an overall positive or negative reaction to instruction in the class that semester and as a reason to continue or discontinue the new practice. However, in a focus group at a small liberal arts college, consistent with literature reviewed earlier in this paper, interviewees told us that the questions on standardized evaluations were not specific enough to give helpful insights about potential positive changes that could be made to a course. “I want to know, not at some abstract level, but in ways that are specific enough to make a difference, ‘what should I do?’” To get better information, one interviewee created his own more specific evaluation form, asking questions contextualized to the class. Another described holding full-class discussions at the end of each semester to gather feedback.

With respect to open-ended comments, faculty tend to place more emphasis on those from a few students than on the data from the quantitative survey portion of evaluations, though the latter is more likely to be completed by all students. While written comments are valuable, they cannot be considered broadly representative. Faculty often referred to “students” when they talked about comments on SETs. When pressed to think about how many students had referred to that issue, they often replied that it was just one or as many as a few. For example, in response to the question, “How many students have to write that comment

for you to find it compelling?” a professor at a small liberal arts college reasoned,

“If you wanted me to pick a number it’d probably be four. And if, there’s some unexpected theme that’s mentioned, apparently independently on four different papers, it’s unlikely that this is a cabal that’s formed specifically for this purpose. It’s more likely they’ve seen something real.”

However, interviewees often automatically generalized from one to many, then implicitly used the generalization as evidence for making a teaching choice. Asked how particular comments would affect what he would teach in the next iteration of a class, an interviewee used “comment” in the singular:

“Well, I’ll make an effort to implement it. For example, I got this comment that I don’t solve enough examples in the class. And I think examples really help so, I want to take more time to do that.”

In this case, the request seems very reasonable for a computer science class, where modeling a problem solving approach is important. However, there is no way for the professor to know if this is the most important use of class time for the greatest number of students or if it is just one student’s need. A few interviewees described another way of formalizing the gathering of feedback. Some train their teaching assistants to routinely ask students how they are doing during lab, believing that students may be more likely to say what they think to near peers.

2) *Feedback Inferred from Performance*

Feedback inferred from performance is actively sought by faculty through assessments of student learning. Students can be identified and a measure of central tendency (usually the mean) of learning outcomes (scores or grades) can be identified. Together with distribution and range, a mean can be used to represent the entire class. This source of feedback is also formalized, with standard questions for all students. It is an involuntary form of feedback, since students must complete assessments in order to pass a class. Feedback concluded on the basis of student performance on assignments and tests has one major advantage over many of the other types of feedback, which is that it provides data from all of the students. Almost no other source of feedback (excepting SETs, discussed above) discussed by faculty in our interviews accomplishes this and clearly it is important to understand the needs of all students in a class before making a major teaching change that will affect all of them.

Interviewees often discussed grades on homework and tests as both impetus for change and evidence that a change had been successful or unsuccessful. For instance, one professor tried the inverted classroom after hearing about it at a conference. She made the decision to try the new practice because of poor grades on a test: “I had given a test and I was just destroyed over the grades I got. So I said, okay, I’m going to try this idea that I heard about.” In this case, a professor was already aware of a new teaching practice and was spurred to implement it because of worse than expected student performance on a test. The main

drawback to this source of feedback is that while it can tell faculty what topics students are having difficulty with, assignments and tests may not provide good information on *why* students are having trouble with those topics – what their misconceptions are. Understanding students’ misconceptions requires more direct interaction with students, who may or may not be able to articulate those misconceptions.

3) *Direct Requests from Students*

Direct, voluntary requests from students are offered by students, but not actively sought by faculty. Students may think that they are taking a risk, given that they are identifiable and may be perceived as challenging the instructor’s authority. Such requests, made by email, questions in class, or informal conversations, are rarely representative of an entire class. Although sometimes these requests were for small changes, such as changing the requirements for robot programming from fighting to dancing, several instructors remembered these types of interactions with students as catalysts for a large change in teaching practices or style. This was often a shift from a lecture-style class to approaches that increase interaction such as the inverted classroom, peer-led team learning, group problem solving, or a mix of such practices in addition to some lecture. A professor at a large public research university describes her reaction to students asking her to stop lecturing in class, but instead to depend on students to read the material and use the class for working together. She said,

“I was just like, huh?? At that time all of my classes were kind of formal lecture classes where I stood at the front of the class and delivered material. But that’s not what they wanted. It turned out that what they wanted was more appropriate for that class than the approach I originally took. And so the class as I taught it over the years had evolved into more of a practicum/laboratory sort of thing.”

A professor at a mid-size state university serving a relatively low-income student body talked about some students who had been awarded a \$10,000 scholarship, who wanted to leave the major, which would mean losing the precious funds. In his class, he lectured and answered student questions for a few minutes, then walked around helping students work out “homework” problems. Later, he explained that this teaching approach was based on a direct request:

“I’m pretty easy going on my students. Right? They know they can talk to me. And she came up and suggested: ‘You need to do more to get involved with the class.’ She and another person told me that they almost dropped out of CS1 in their first semester and if hadn’t been for the scholarship they had received, in about the third or fourth week they would’ve just run like bunnies. And I took that seriously. ‘What do I need to change it?’ I asked them, and what they told me was ‘not enough hands-on’. ‘We were too lost at the beginning of your course.’ They, they didn’t have any background. I

realized, boy, you know, I thought I was doing a good job at bringing in people who were not comfortable yet with computer science. I started to really rethink my teaching. About what I do.”

A similar request was made of an assistant professor at a small, private liberal arts college. He was new to his position, having recently graduated from a large, public research university where lecturing was the norm. He describes his decision to stop lecturing based on requests from multiple students:

“And I think it took about a year for me to realize that it just wasn’t working anymore. My prep time was not being spent well at all. I wasn’t spending the time mastering the material, and thinking about how these students can best learn it. And you know I’ve heard that issue from several of my colleagues that it took them about a year to get out of that mentality, and I think it’s because a lot of us do come from bigger schools.”

This professor interpreted his lack of success as a kind of cultural conflict in that his students at the small school where he was teaching expected to interact more directly with their instructors than the lecture-style teaching he was used to. He successfully addressed both his students’ complaints and his concerns about learning outcomes by shifting his teaching style from solely traditional lecture to include other, more interactive practices such as group problem solving. However, some direct requests may serve only to avoid problems. This professor described how he stopped using pair programming because of a complaint by a female student that her male programming partner wanted more than a co-learning relationship:

“There was one woman that emailed me privately that had had an uncomfortable experience with the male that I matched her with. So I kind of discreetly broke them up but it was, it was a situation nonetheless. So I haven’t done that since.”

In this case, the professor chose to entirely discontinue a teaching practice, rather than pre-empting a repeat occurrence by setting some ground rules with students.

4) *Nonverbal Feedback*

Students communicate with more than words. We found that faculty were quick to interpret nonverbal feedback that students may or may not have been aware they were “giving off” [36]. Nonverbal feedback occurs through body language, facial expression, or even lack of any communication. Of course, a big problem is that interpretation of nonverbal behavior may or may not be accurate. Faculty often referred in passing to interpreting students’ behavior and demeanor in the classroom in order to judge whether students were engaged and paying attention. They also discussed making changes based on these judgments (or making changes with these judgments as partial motivations for the changes). These might be as simple as noticing differing levels of attendance across different classes. One professor felt that attendance was an

indicator of engagement, comparing an introductory course where attendance was low to higher level courses:

“Some of my upper division classes that I taught put together graphics and game programming. And I think students that sign up for those classes are really interested and the attendance numbers are better in those classes. But I generally don’t require attendance so, I think either maybe the material’s too easy or they have their own issues going on.”

Facial expression and demeanor were other sources of nonverbal feedback that faculty used. We heard both positive and negative examples of classroom behavior for implementing change or judging change. For example, these two professors evaluated the outcome of teaching change, saying “...okay, these students’ faces look happy and I’m feeling more connected to them, and I’m better able to manage what I’m doing because I know what’s going on with them” contrasted with “I get this feeling that, I look at their faces and I get this feeling I wasn’t really effective.”

Making judgments about students from their nonverbal feedback may also provide faculty with a reason not to make a change. One professor, who perceived his students as quiet and unengaged, was worried that the inverted classroom would not work in his case:

“I can just see Monday morning rolling around and after ten minutes they’re out of ideas, uh, questions. Or half of them haven’t done the reading and they’re just lost. Or maybe don’t come to class because they hadn’t done the reading. I hate to do it.”

Two other examples show how instructors’ perceptions of students’ engagement are not always accurate. For instance, appropriate classroom behavior can be very different across cultures. One professor at a American Indian tribal college found that,

“(American Indians in this tribe) are taught not to look at, it’s disrespectful to look at somebody in the eye that is an elder or a superior. And so when it looks like they’re not paying attention they’re actually giving you respect.”

This professor initially interpreted lack of eye contact from students as disengagement, but he found those students to be engaged and intelligent when he talked to them one on one. When he asked the department chair about disengaged in-class demeanor, the chair pointed out that the professor was probably unaware of the cultural difference at play in the situation.

In a second case, a professor at a small liberal arts college was concerned because she felt that only a few students were responding to questions and that these students may have been intimidating others. She combined two strategies to overcome lack of talkativeness in her students. She required that students complete a questionnaire about studying data structures before the first day of class. One question asked about students’ comfort with speaking up and she asked them to try to come out of their comfort zones. She backed this up throughout the

semester by using turn taking techniques to call on all students equally (see [37]). This method led her to the revelation that her interpretations of student engagement were sometimes wrong. Having observed a student lying across his desk, she called on him anyway because it was his turn. She was surprised:

“...So... I asked <student> the question, and what I actually found is that at least some of the time <student's> head comes up, and he answers the question. He was following along just fine. He was totally on the same page with everyone else and so on. And so that was really revealing to me in that you may think, from your observations of the class, that you know what's going on and what the dynamics are. And obviously it's not always wrong, but sometimes it is wrong. And that's a really useful reality check.”

She also noted that one can assume that students are smarter than they are because they are willing to speak up:

“...there are also people who are kind of vocal. And from their pose and demeanor and the fact that they'd ask you questions on some things, you would think that they were pretty smart and on the ball. But when you follow this kind of scheme, and you pick on them with an arbitrary question, well, actually, they aren't quite as good as you might have thought they were. So, you know, it goes the other way as well.”

In that case, she felt students were willing to speak up about concepts they knew, giving off the perception that they knew a lot. As we discuss below, student feedback can be problematic in many ways.

C. *The Problematic Characteristics of Feedback aka How Do You Know What to Believe?*

Student feedback has several characteristics that make it a risky kind of evidence for decision making. It may be actively sought by faculty or inferred from student behavior; students may be anonymous or identified; feedback may be broadly representative or a minority opinion; topics can be standardized for all students or generated by students' individual concerns; and feedback may be volunteered by students or simply “given off” unintentionally.

Some feedback is intentionally sought, either because faculty are required to seek it (e.g., student evaluations of teaching) or because they want to. In contrast, other feedback is not actively sought, but is based on observing student behaviors. Both elicited and inferred feedback can be problematic as sources of evidence. When faculty go out of their way to elicit feedback, as in a class discussion held just for that purpose, students' opinions may change toward the instructor, since students may take that as a sign of concern for their learning and welfare. In contrast, inferred behavior is a problem because as we described above, nonverbal communication is difficult to interpret accurately.

Students either give feedback with the expectation that they are anonymous (as in the case of SETs) or knowing that they can be identified by their professor (as in the case

of emails or direct requests). Students may be less likely to contribute critical feedback in situations where they know that the feedback can be connected to them. They may also do “facework” [38] for faculty, not wanting to disappoint nor criticize, because although they may not like the instruction, they may like or respect the instructor. Effective facework is transparent, so faculty may believe only the positive comments. On the other hand, however, when students know they cannot be identified, they sometimes use the opportunity to give unwarranted criticism.

In our findings, relatively few of the sources of feedback on which faculty relied did in fact broadly represent the entire class and feedback that represented very limited numbers of students sometimes had an outsized influence on decisions about teaching change. Every person who has ever taken statistics or studied reasoning understands that in order to generalize the data from a sample to a larger group, the sample must be large enough that a finding is not due to chance and representative of the population it is meant to describe. SETs are intended to represent an entire class and if the subset of students who completes the form is large enough, one could infer generalizability with a certain degree of confidence. Comments on SETs, however, cannot be generalized. Even if two or more persons wrote about the same issue, it is not clear that their opinion can be quantified and generalized. The only way to generalize is to ask all students about that issue. It is also hard to know whether the sample of students who do respond might be biased either in favor of the instruction or against it, as is the case with voluntary reviews on commercial web sites.

The content of comments or ratings can either come from a standard set of questions that all students are asked as a part of activities that are structured into the course (such as SETs), or can be based on the interests or needs of a particular student. This distinction is important because standardized questions reduce the information to a point that it is not useful for decision making, whereas comments that are not limited by a standardized form may bring to light issues that the professor would not have considered otherwise. Certainly, comments are more closely *about* a particular class, its contents, and instructor. On the other hand, one might problematize even the closed-ended items, since given their level of abstraction, it's hard to know exactly what details were brought to mind when a student rated the item.

Feedback may be intentionally volunteered by students or may be communicated unintentionally. Both can have an effect on teaching decisions. For example, a student reading their Facebook newsfeed through one's lecture in class is unintentionally (hopefully) communicating a lack of interest in the lecture. The potential danger of relying on this source of feedback alone is that judgments of facial expression, body language, and (mis)behavior can be inaccurate, without adding another method of determining students' engagement or learning outcomes. Faculty often take some types of voluntary feedback particularly seriously, such as one-on-one conversations or direct requests from students. However, they often also take unintended feedback seriously, drawing conclusions that can be inaccurate.

Faculty piece together all of these characteristics of feedback in context and make decisions about what students need and how and whether those needs can be provided.

V. SOLUTIONS OFFERED BY INTERVIEWEES

Our findings are consistent with those of Fossati & Guzdial, who, in interviews with 14 computer science instructors, found that teachers tended to not use empirical evidence [39]. Instead, faculty in their study relied on intuition and anecdotal evidence. We argue above that our intuitions can be wrong and anecdotes are not generalizable, and the reasons that we hear them may be biased. One solution to the problem of inaccurate interpretations of feedback is to increase the frequency and structure of feedback mechanisms, making sure that it comes from all of the students, not just those who sit in the front row or the ones who are willing to take the time to come to office hours. We conclude by describing two feedback mechanisms used by interviewees, one who taught in a large lecture format, and the other in a small, lab-based setting.

Gathering feedback from all students is particularly advantageous, since feedback that occurs during classroom discussions generally does not have this characteristic. One professor at a small liberal arts college routinely distributed a three page questionnaire that asked quite detailed questions about the course. Questions include whether the particular instructional approach contributed to understanding the topic, if the student would have preferred a different approach, and whether the approach was engaging. Also requested are comments on pace, required laboratory exercises, and class activities. Finally, students are asked to rate the learning resources (e.g., textbook). Although they did not have to put their names on the questionnaire, students were required to turn it in before they could take the final exam. These questionnaires clearly provide a wealth of information to support evidence-based decision making. The practice, however, is untenable with large student enrollments.

At a large, public state research university, one professor gathered frequent feedback from students in large classes (roughly 80-100 students) in two ways. First, she gathered immediate feedback on topic comprehension during class by having students “vote” on the right answer using “clickers.” This gave her a general idea of whether or not most students understood a topic. Before telling them the correct answer, she had students break out into small groups to discuss the question while wandering around the classroom and listening in on several of the groups. This allowed her to gather feedback from individual students that was invaluable for her understanding of their misconceptions. She described the value of this feedback for day-to-day teaching choices:

“So the individual vote comes in and I think whoa, I thought this was an easy question but, you know, forty percent of the class got it wrong. ‘How is it even possible to not get this right? You know you just think this is so easy! I don’t even understand how, what can go wrong here.’ (slight laugh) And so when I go listen to what the students are saying then,

and I hear what they’re confused about, I say ‘Oh, that? Ok.’ And sometimes it’ll remind me of, you know, those memories, now long buried (slight laugh), of what I used to find confusing about that. And then sometimes it’s things that never ever would have occurred to me to be confused about but, um, just a different way of thinking about it. So that gives me a chance to quickly study what do I need to talk about when I debrief that question.”

The professor believed that this practice also had the additional effect of making students comfortable speaking up in class:

“And, you know, every faculty member knows the feeling of saying ‘Are there any questions?’ and no hands go up. And then, you know, everyone bombs the midterm or whatever. But I find that that doesn’t happen as much anymore because it’s kind of like we’ve broken that taboo against speaking up in class. So now they’re used to talking, and they’re much more forthcoming with questions or answering my questions even when we’re structurally in a more traditional mode because they’re just, they’re just more... I frankly think they’re just also more awake. There’s more blood flowing (laughs).”

Throughout our interviews with computer science faculty, the idea of student feedback as an important factor or motivator in decisions about changing teaching practices came up over and over again. It may be possible to inform students about effective ways to communicate their desires with faculty. Understanding the process by which faculty decide to adopt new teaching practices is central to supporting adoption of effective teaching practices.

VI. ACKNOWLEDGMENTS

This research was sponsored by the National Science Foundation under Division of Undergraduate Education #1123332. Any opinions, findings, conclusions, or recommendations expressed are those of the authors and do not reflect those of the National Science Foundation. We thank Kathleen Adrian, Rick Barker, Jesus Contreras, Jonatan Contreras, Evan Hofmockel, Christopher Hovey, Russell Mehring, Kamar Nassor, and Ann Seago for their interview, logistics, and transcription support. Finally, we thank the computer science faculty members who committed to be interviewed and allow us in their classes.

VII. REFERENCES

- [1] L. Perna W., A. Ruby, R. Boruch, N. Wang, J. Scull, C. Evans, and S. Ahmad, “The life cycle of a million MOOC users,” presented at the MOOC Research Initiative Conference, Arlington, Texas., 2013.
- [2] L. J. Barker, C. McDowell, and K. Kalahar, “Exploring factors that influence computer science introductory course students to persist in the major,” in *Proceedings of the 40th ACM Technical Symposium on Computer Science Education*, Chattanooga, TN, USA, 2009, pp. 153–157.
- [3] J. M. Cohoon, “Toward improving female retention in the computer science major,” *Assoc. Comput. Mach.*, vol. 44, no. 5, pp. 108–114, 2001.
- [4] J. M. Cohoon and W. Aspray, “A critical review of the research on women’s participation in postsecondary computing education,” in

- Women and Information Technology: Research on Underrepresentation*, Cambridge, MA: MIT Press, 2006, pp. 137–180.
- [5] D. D. Davis, D. A. Major, J. V. Sanchez-Hucles, S. J. DeLoatch, K. A. Selgrade, and S. K. Meert, “Organization climate, satisfaction, commitment, and career change in computer science,” presented at the 115th Annual Convention of the American Psychological Association, San Francisco, CA, 2007.
- [6] J. Margolis and A. Fisher, *Unlocking the Clubhouse: Women in Computing*. Cambridge, MA: MIT Press, 2002.
- [7] C. McDowell, L. Werner, H. E. Bullock, and J. Fernald, “The impact of pair programming on student performance, perception and persistence,” in *International Conference on Software Engineering*, Portland, Oregon, 2003, pp. 602–607.
- [8] L. Werner, “Female computer science students who pair persist,” p. Article 4, 2004.
- [9] C. McDowell, L. Werner, H. E. Bullock, and J. Fernald, “Pair programming improves student retention, confidence, and program quality,” *Commun. ACM*, vol. 49, no. 8, pp. 90–95, 2006.
- [10] W. M. Waite, M. H. Jackson, and A. Diwan, “Assessing the conversational classroom,” University of Colorado at Boulder, 2006.
- [11] N. M. Webb and A. S. Palinscar, “Group processes in the classroom,” in *Handbook of Educational Psychology*, New York: Macmillan, 1996, pp. 841–873.
- [12] L. A. Williams, “Strengthening the case for pair programming,” *IEEE Softw.*, vol. 17, no. 4, pp. 17–25, 2000.
- [13] B. C. Wilson, “A study of factors promoting success in computer science including gender differences,” *Comput. Sci. Educ.*, vol. 12, no. 1–2, pp. 141–164, 2002.
- [14] M. A. Hogg and D. J. Terry, “Social identity and self-categorization processes in organizational contexts,” *Acad. Manage. Rev.*, vol. 25, no. 1, pp. 121–140, 2000.
- [15] M. Taylor Huber and P. Hutchings, “Building the teaching commons,” *Change*, vol. 38, no. 3, 2006.
- [16] J. R. Kimberly, “The organizational context of technological innovation,” in *Managing Technological Innovation*, San Francisco: Jossey-Bass, 1986, pp. 23–43.
- [17] K. J. Klein and J. S. Sorra, “The challenge of innovation implementation,” *Acad. Manage. Rev.*, vol. 21, no. 4, pp. 1055–1080, 1996.
- [18] D. L. Rubin and K. L. Greene, “Effects of biological and psychological gender, age cohort, and interviewer gender on attitudes toward gender-inclusive/exclusive language,” *Sex Roles*, vol. 24, no. 7, pp. 391–412, 1991.
- [19] L. G. Tornatzky, J. D. Eveland, M. G. Boylan, W. A. Hetzner, E. C. Johnson, D. Roitman, and J. Schneider, *The Process of Technological Innovation: Reviewing the Literature*. Washington, D.C.: National Science Foundation, 1983.
- [20] C. J. Dommeyer *, P. Baum, R. W. Hanna, and K. S. Chapman, “Gathering faculty teaching evaluations by in-class and online surveys: their effects on response rates and evaluations,” *Assess. Eval. High. Educ.*, vol. 29, no. 5, pp. 611–623, 2004.
- [21] S. d’ Apollonia and P. Abrami, “Navigating Student Ratings of Instruction,” *Am. Psychol.*, vol. 52, no. 11, pp. 1198–1208.
- [22] H. W. Marsh and L. A. Roche, “Making students’ evaluations of teaching effectiveness effective: The critical issues of validity, bias, and utility,” *Am. Psychol.*, vol. 52, no. 11, pp. 1187–1197, Nov. 1997.
- [23] J. A. Kulik, “Student Ratings: Validity, Utility, and Controversy,” *New Dir. Institutional Res.*, vol. 2001, no. 109, pp. 9–25, 2001.
- [24] J. Jones, “Students’ ratings of teacher personality and teaching competence,” *High. Educ.*, vol. 18, no. 5, pp. 551–558, Sep. 1989.
- [25] C. C. Lo, “How Student Satisfaction Factors Affect Perceived Learning,” *J. Scholarsh. Teach. Learn.*, vol. 10, no. 1, pp. 47–54, Jan. 2010.
- [26] T. Beran, C. Violato, D. Kline, and J. Frideres, “The Utility of Student Ratings of Instruction for Students, Faculty, and Administrators: A ‘Consequential Validity’ Study,” *Can. J. High. Educ.*, vol. XXXV, no. 2, pp. 49–70, 2005.
- [27] L. P. Schmelkin, K. J. Spencer, and E. S. Gellman, “Faculty Perspectives on Course and Teacher Evaluations,” *Res. High. Educ.*, vol. 38, no. 5, pp. 575–592, Oct. 1997.
- [28] F. N.-A. Alhija and B. Fresko, “Student evaluation of instruction: What can be learned from students’ written comments?,” *Stud. Educ. Eval.*, vol. 35, no. 1, pp. 37–44, Mar. 2009.
- [29] M. D. Svinicki, “Encouraging Your Students to Give Feedback,” *New Dir. Teach. Learn.*, vol. 2001, no. 87, pp. 17–24, 2001.
- [30] J. C. Ory and L. A. Braskamp, “Faculty perceptions of the quality and usefulness of three types of evaluative information,” *Res. High. Educ.*, vol. 15, no. 3, pp. 271–282, Sep. 1981.
- [31] M. Theall and A. Franklin, “Using student ratings for teaching improvement,” *New Dir. Teach. Learn.*, vol. 48, pp. 83–96, 1991.
- [32] N. Hardy, “Online Ratings: Fact and Fiction,” *New Dir. Teach. Learn.*, vol. 2003, no. 96, pp. 31–38, Winter 2003.
- [33] D. M. Zimmaro, C. S. Gaede, E. J. Heikes, M. P. Shim, and K. G. Lewis, “A study of students’ written course evaluation comments at a public university,” The University of Texas at Austin, 2006.
- [34] B. Oliver, B. Tucker, and Pedgen, “An investigation into student comment behaviors: Who comments, what do they say, and do anonymous students behave badly?” Paper presented at the Australian Universities Quality Forum, Hobart, 11-Jul-2007.
- [35] G. C. Gannod, J. E. Burge, and M. T. Helmick, “Using the Inverted Classroom to Teach Software Engineering,” in *Proceedings of the 30th International Conference on Software Engineering*, New York, NY, USA, 2008, pp. 777–786.
- [36] E. Goffman, *The Presentation of Self in Everyday Life*, 1 edition. Garden City, NY: Anchor, 1959.
- [37] L. J. Barker, M. O’Neill, and N. Kazim, “Framing Classroom Climate for Student Learning and Retention in Computer Science,” in *Proceedings of the 45th ACM Technical Symposium on Computer Science Education*, New York, NY, USA, 2014, pp. 319–324.
- [38] K. Tracy, “The many faces of facework,” in *Handbook of language and social psychology*, H. Giles and W. P. Robinson, Eds. Oxford, England: John Wiley & Sons, 1990, pp. 209–226.
- [39] D. Fossati and M. Guzdial, “The Use of Evidence in the Change Making Process of Computer Science Educators,” in *Proceedings of the 42nd ACM Technical Symposium on Computer Science Education*, New York, NY, USA, 2011, pp. 685–690.