STUDENT AND INSTRUCTOR EXPERIENCES WITH A COMPUTER-BASED TESTING FACILITY

C. Zilles, M. West, D. Mussulman, C. Sacris

University of Illinois at Urbana-Champaign (UNITED STATES)

Abstract

In this work, we explore how a large-scale introduction of computer-based testing has impacted students and instructors. The College of Engineering at the University of Illinois has been operating a Computer-Based Testing Facility (CBTF) for four years now, and the CBTF has matured into a fixture of our College. In Fall 2017, the CBTF served 21 courses from seven different departments and over 6,000 unique students. Over 52,000 exams were proctored, including 3,500 final exams.

This paper summarizes key findings from a collection of surveys completed by students and instructors. Most instructors report having positive experiences with the CBTF. A large majority report that, once the exam content is in place, they perceive reductions in the effort to run and grade exams and to handle exceptional situations. Instructors also like how the CBTF enables them to run small frequent tests, run second-chance exams, and test computational skills. The vast majority of the surveyed instructors plan to continue using the CBTF and think that it should be expanded.

Some highlights from our student surveys include: (1) students generally are more satisfied with CBTF exams relative to traditional paper exams (45% satisfied or very satisfied, vs. 17% dissatisfied or very dissatisfied), but this preference seems to vary by major, with computer science and electrical engineering majors even more strongly preferring the computerized exams, (2) students' favorite aspects of CBTF exams include the flexibility to schedule them at convenient times, that CBTF courses generally have more frequent, shorter tests, and the opportunities to take second chance exams, and (3) some students prefer the partial credit mechanism commonly used in traditional written exams, where credit is granted for work shown for incorrect answers.

Keywords: Computer, Exams, Testing, STEM, Engineering, Asynchronous, Proctored.

1 INTRODUCTION

Exams are a widely used method of summative assessment in college-level courses, especially introductory ones. The large size (e.g., 200+ students) of introductory courses at many universities presents many challenges to offering traditional pencil-and-paper exams, including requesting space, printing exams, proctoring, timely grading, and handling conflict exams [9], [10], [17]. These practical concerns often dominate pedagogical concerns in how assessment is performed in these classes.

This paper discusses experiences with a Computer-Based Testing Facility (CBTF, Figure 1(a)) that was developed at the University of Illinois as an alternative approach to handling exams for large classes. The goal of the CBTF is to make assessment with exams better for everyone involved—

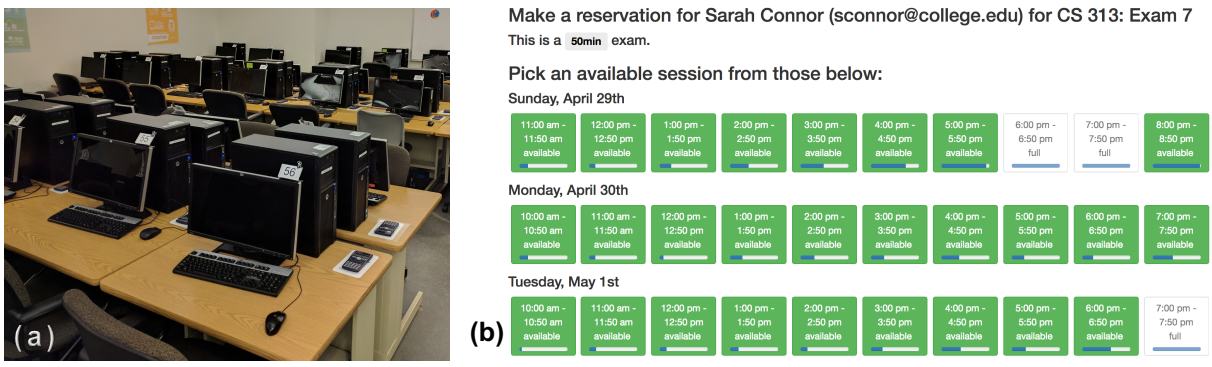

Figure 1. (a) The CBTF is a proctored 90-seat computer lab for taking exams. (b) Students make exam reservations using a web-based interface that lets them select from available times.

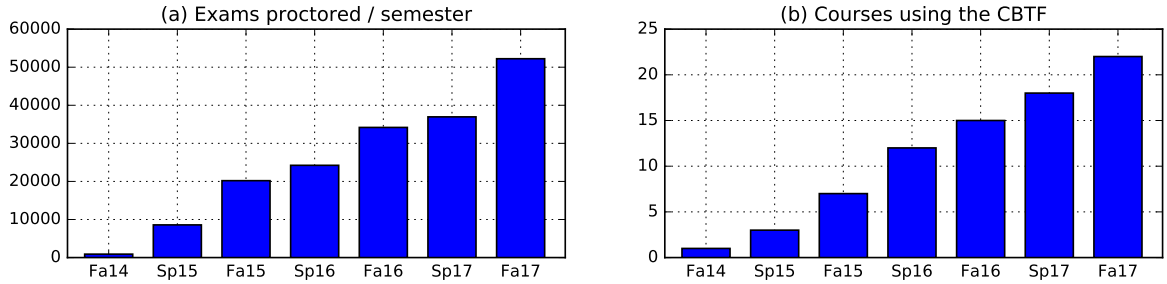

Figure 2. The exams proctored by the CBTF (a) and number of courses using the CBTF (b) have grown roughly linearly in the four years of its operation.

students, faculty, and course staff. Four concepts are key to achieving this goal. First, by running the exams on computers, we can write complex, authentic (e.g., numeric, symbolic, graphical, programming) questions that are auto-gradable, allowing us to test a broad set of learning objectives with minimal grading time and providing students with immediate feedback. Second, we write question generators that use randomness to produce a collection of problems, allowing us to give each student different questions and permitting the problem generators to be used semester after semester. Third, because each student has a unique exam, we allow students to schedule their exams at a time convenient to them within a specified day range (Figure 1(b)), providing students flexibility and avoiding the need to manage conflict exams. Finally, because exam scheduling and proctoring is completely handled by the CBTF, once faculty have their exam content, it is no more effort to run more-frequent, smaller exams which helps force students to keep up with the course and potentially reduces student anxiety. In addition, a number of courses see a reduction in fail rates by offering "second-chance" exams that allow students to do additional studying after a poor exam performance and take another equivalent exam for some form of grade replacement.

Our CBTF is now operating in its fourth year. During that period, we have scaled from less than a thousand exams in the first semester to over 52,000 exams in Fall 2017 (Figure 2(a)), which is a testament to the degree that instructors value the CBTF. There has been a corresponding growth in the number of courses using the CBTF (Figure 2(b)). Currently the CBTF hosts courses from Aerospace Engineering, Computer Science, Electrical and Computer Engineering, Mechanical Engineering, Material Science and Engineering, Physics, and Statistics.

The CBTF is sufficiently mature and successful that it is worth studying to see how it is perceived by instructors and students, which is the purpose of this paper. After a brief introduction to computer-based testing and our CBTF's operations in Section 2, we describe findings from one survey of instructors (Section 3) and three surveys of students (Section 4). In Section 5, we conclude.

2 COMPUTER-BASED TESTING AND THE CBTF'S OPERATION

Computer-based testing is not new, it dates back at least to the 1980s [1], [2]. Shacham has argued that exams are the most beneficial application of computers in engineering education [13]. Two of the major benefits of computer-based testing are that it greatly reduces the overhead of running exams and permits running exams asynchronously, allowing different students to take exams at different times [1], [6], [12], [13], [17]. In addition, the ability to provide students with immediate feedback about their errors has pedagogical value [13] and permits writing exams that allow re-trying until mastery is achieved. One major challenge of computer-based exams is generating the content [5], [8], [11].

What is still relatively rare, however, is universities developing college or campus-level resources for supporting computer-based testing for on-campus courses, although there are a few examples besides our CBTF. For more than ten years, the University of Helsinki has been running an electronic exam room where students can take their final exams at a time of their choosing in a computer lab [12]. Running less than 500 exams/year on average, they grade exams manually and run exams in a lab with 16 computers without proctors, instead relying on video recording the test taker, as well as the use of multiple versions of questions. Contemporary with our CBTF, the University of Central Florida developed a similar facility which also seeks to support a number of large-enrollment classes with a testing center much smaller than any one class, by running the exams asynchronously [6].

Our CBTF is a dedicated 90-seat computer lab that has been configured to serve as a secure, accommodating testing environment. It runs on an hourly schedule, offering 50-minute and 1-hour-and-50-minute exams based on the needs of the class. The CBTF is open all seven days of the week from 10am to 10pm. Students are provided a 3-4 day window in which to scheduler their exam (Figure 1(b)) and can freely re-schedule their exam slot up until the beginning of their reservation as space permits.

Students are seated for exams in the ten minutes before the exam begins. Proctors check university photo-ID cards, and our scheduling software assigns the student to sit at a particular computer. To minimize opportunities to cheat, students store all of their coats and bags (except a pencil or pen) on racks by the CBTF entrance. Blank scratch paper and handheld calculators are provided by the CBTF. Networking on the CBTF computers is controlled to prevent general internet access and communication. Individualized exams (using randomized problems) are generated when the student logs into the exam server. Students submit their answers and most questions are graded interactively; most exams are configured to permit multiple submissions for the same question with decreasing partial credit based on the number of attempts to reach a correct solution. During the period of the exam, there are no course staff present, only the proctors. As a result, the exam must be sufficiently self-explanatory to be interpreted by students without assistance.

During most weeks of the semester, the CBTF handles more than a dozen distinct exams each week. This means that the exam periods overlap and that students from six or more different classes may be taking an exam in the CBTF during the same time slot. The CBTF works with courses at the beginning of the semester to produce an exam schedule that meets the needs of courses without over-subscribing the available slots. We use a model to predict expected student usage for a given schedule [14] and typically aim for no more than an 85% expected utilization for any given day [18].

3 FACULTY FEEDBACK

A survey was conducted in May 2017 of all course instructors who had used the CBTF in the previous two semesters (Fall 2016 and Spring 2017). The response rate was 26/31 = 84%.

Instructors using the CBTF were generally very positive about the impact of the CBTF on their courses as shown in Figure 3. Large majorities of instructors reported reductions in the effort to run and grade exams and deal with student exceptions, as well as reductions in printing costs for their exams. Furthermore, the CBTF enabled improvements to the structures of exams, allowing instructors to offer smaller, more frequent tests and second-chance exam opportunities to students. In addition, our College of Engineering is working to emphasize the development of computational skills by our engineering students, and instructors found that the CBTF has benefits to achieving this goal. Finally, instructors find that the immediate feedback that students get in the CBTF to be positive and that the CBTF has positively impacted student learning in their classes. As one instructor commented:

> *"The CBTF has allowed us to move from a standard 3-midterm model to a weekly quiz model. As a result, students are staying on top of the material, which has made a substantial impact to their learning, but also feeds back into the lecture and lab components of our course. Students are more participatory in these sections because they have not fallen behind."*

Instructors also reported generally high satisfaction with the operational aspects of the CBTF as shown in Figure 4. Most instructors find getting started with the CBTF to be reasonable and the CBTF's support for handling exceptional circumstances (e.g., students missing exams, power outages, etc.) and for handling students with testing accommodations to be sufficient. Instructors have quite positive opinions of the custom learning management systems (PrairieLearn [15], [16] and RELATE [7]) that we use in the CBTF, their reliability, and their support for a broad range of question types. Instructors are generally satisfied with the anti-cheating measures that the CBTF provides in both question randomization to prevent collaborative cheating between students taking exams at different times [4] and the physical security provided by the proctors.

Instructors that have used the CBTF see it as an important resource for handling enrollment growth in our College of Engineering, as shown in Figure 5. About 85% of the surveyed instructors plan to continue using the CBTF. About half would even be willing to teach with a smaller course staff if they could continue to use the CBTF. Over 80% think that it should be expanded to enable additional

courses to take advantage of its benefits. When asked to provide advice to other instructors considering adopting the CBTF, some comments from survey participants include:

"You should have the materials in place before you attempt to adopt the CBTF. It is easy to underestimate how much effort it is to develop good question generators."

*"CBTF exams are *not* a drop-in replacement for traditional pencil-and-paper exams. They are different. Your exams (and policies) have to change."*

"This has revolutionized assessment in my course. It is much more systematic, the question quality is much improved, and my TA's and myself can focus on preparing questions (improving questions), rather than grading."

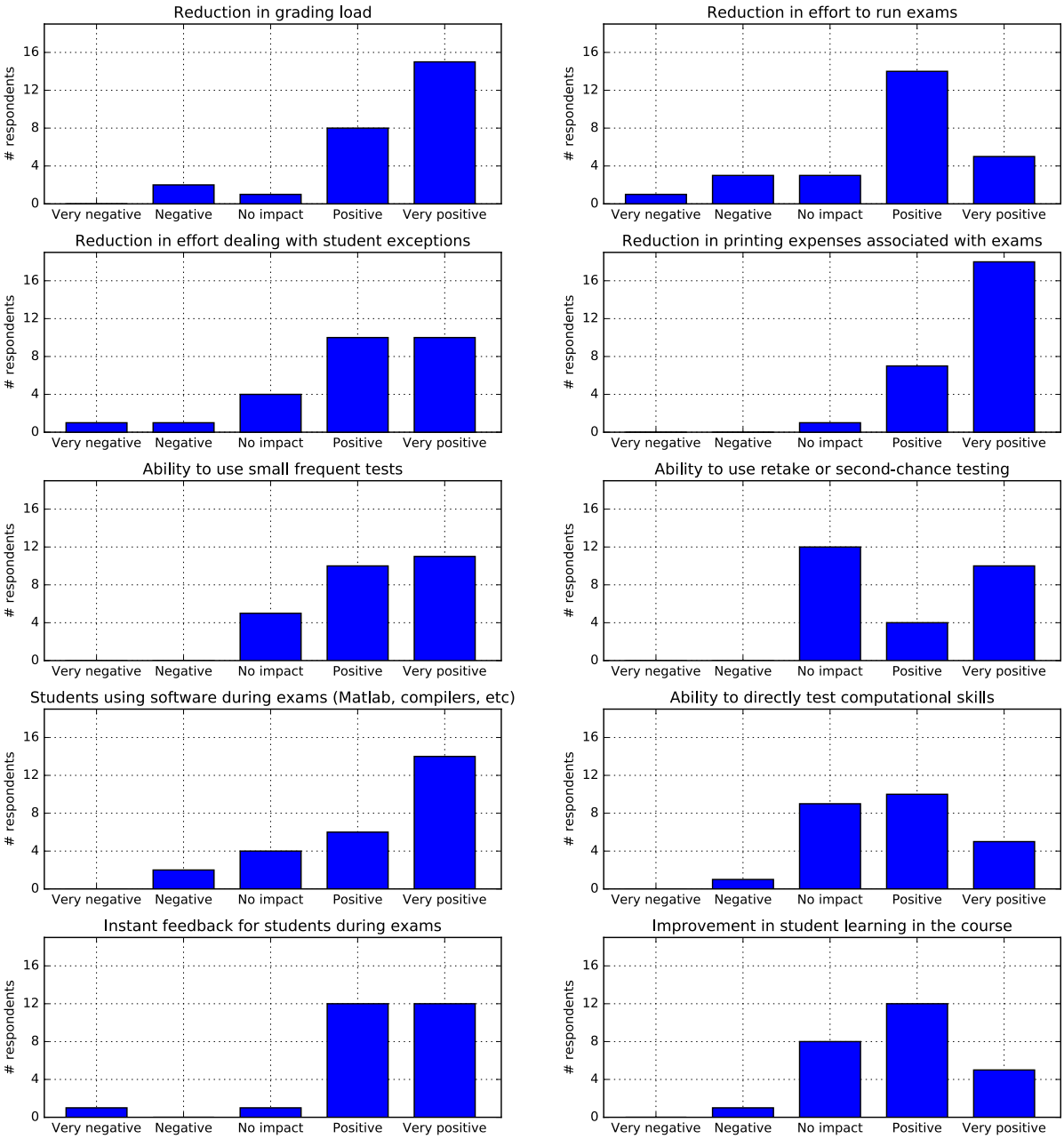

Figure 3. Instructor responses to questions asking, "How has the CBTF impacted your courses?" The majority of instructors find the CBTF to have a positive effect in all of the surveyed categories.

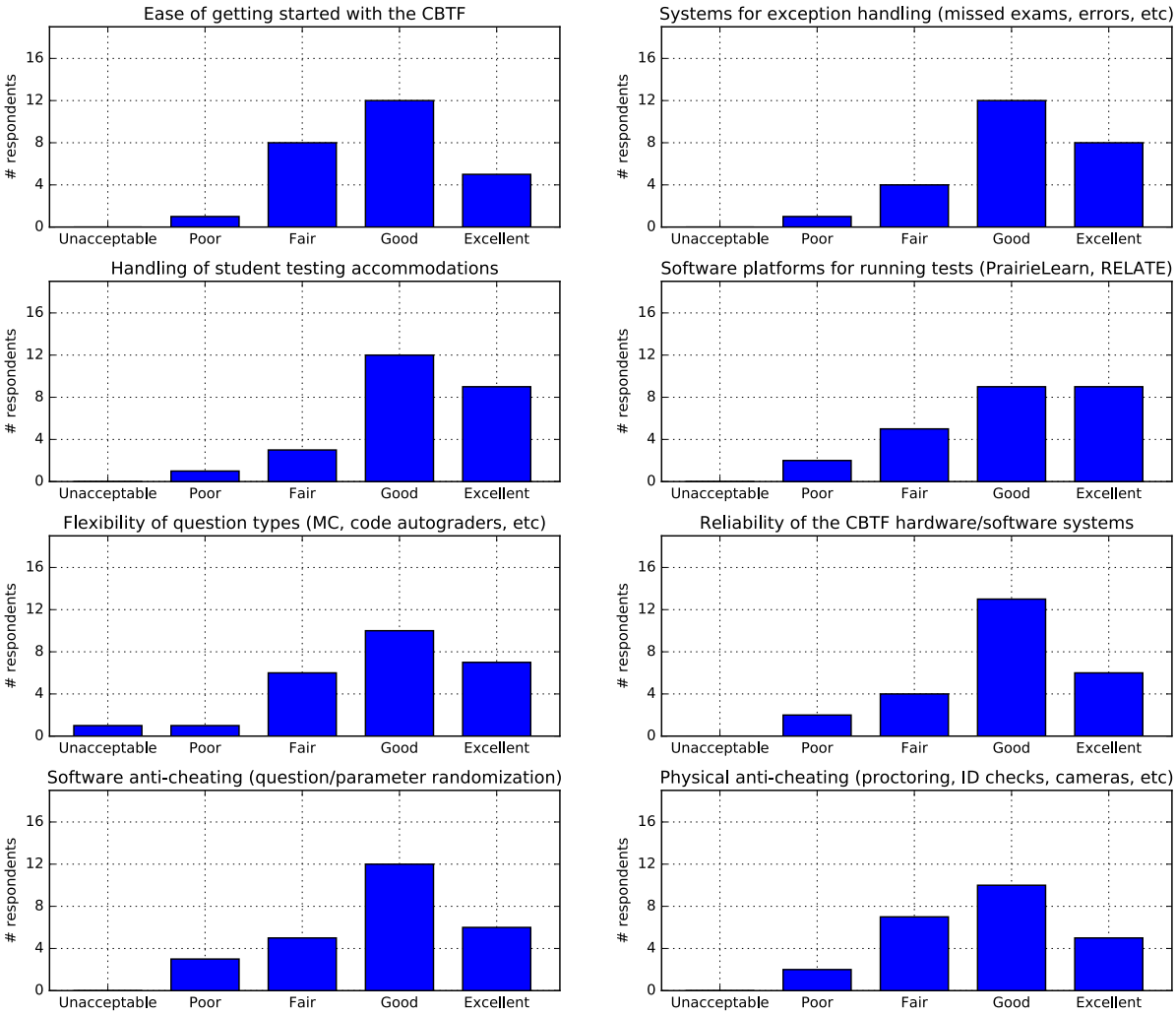

Figure 4. Instructor responses to questions asking "Please rate the following aspects of the CBTF". The majority of instructors rate all aspects of the CBTF as "Good" or "Excellent".

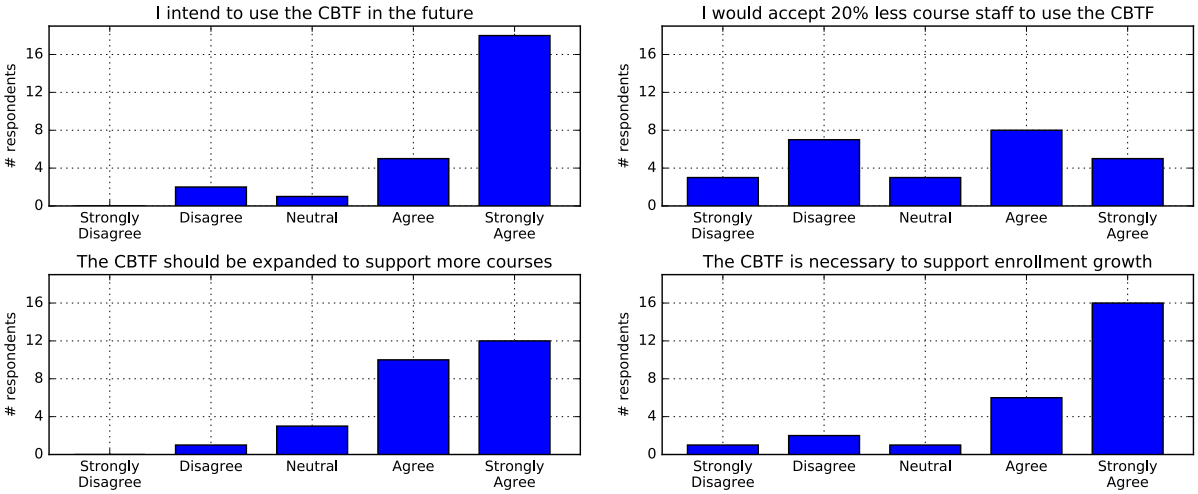

Figure 5. Instructor responses to questions asking "Please rate the following statements about the future of the CBTF". The majority of instructors plan to continue using the CBTF and believe that it should be expanded to support additional courses.

4 STUDENT FEEDBACK

We have solicited student feedback about the CBTF through a collection of large surveys over the past two years. In Fall 2016, a sophomore-level required Mechanical Engineering class surveyed its students and got 217 responses. In May of 2017, a survey about technology use by students in the College of Engineering included questions about the CBTF that received responses from 612 students. In May of 2018, the CBTF sent one of four surveys to each of its over 5,400 Spring 2018 users for a combined 872 responses (a 16% response rate).

Across all of these surveys we see rather consistent results, which we summarize here. We present our findings along three themes: overall satisfaction (Section 4.1), differences between computing and non-computing populations (Section 4.2), and cheating (Section 4.3).

4.1 Overall Satisfaction

In the 2017 survey, students were asked "how did your experience taking an exam in the CBTF compare with your experience taking traditional paper exams?" As shown in Figure 6, students on the whole reported being more satisfied than not when comparing the CBTF exam experience to traditional paper exams (45% satisfied or very satisfied, against 17% dissatisfied or very dissatisfied). Our 2018 survey sheds light on factors that contribute to this positive attitude. It shows that the vast majority of students find the CBTF policies to be reasonable and opportunities to take second-chance exams to be beneficial to their learning. Most students like getting immediate feedback about which problems they got right as they progress through the exam, but we know from open-ended feedback that test anxiety increases for some students when they grade a question in the middle of the exam and see that they've gotten it wrong. From the 2016 survey data (not shown) and open-ended feedback on the 2018 survey, we know that students are very fond of the flexible scheduling, even the ones that are not fond of other aspects of the CBTF.

Students are more mixed about whether they like more-frequent testing. Slightly more students are in favor of shorter, more-frequent tests (42% vs. 31%), but, at the same time, more students agree than disagree with the statement "I feel that my CBTF courses do too much testing," as shown in Figure 6.

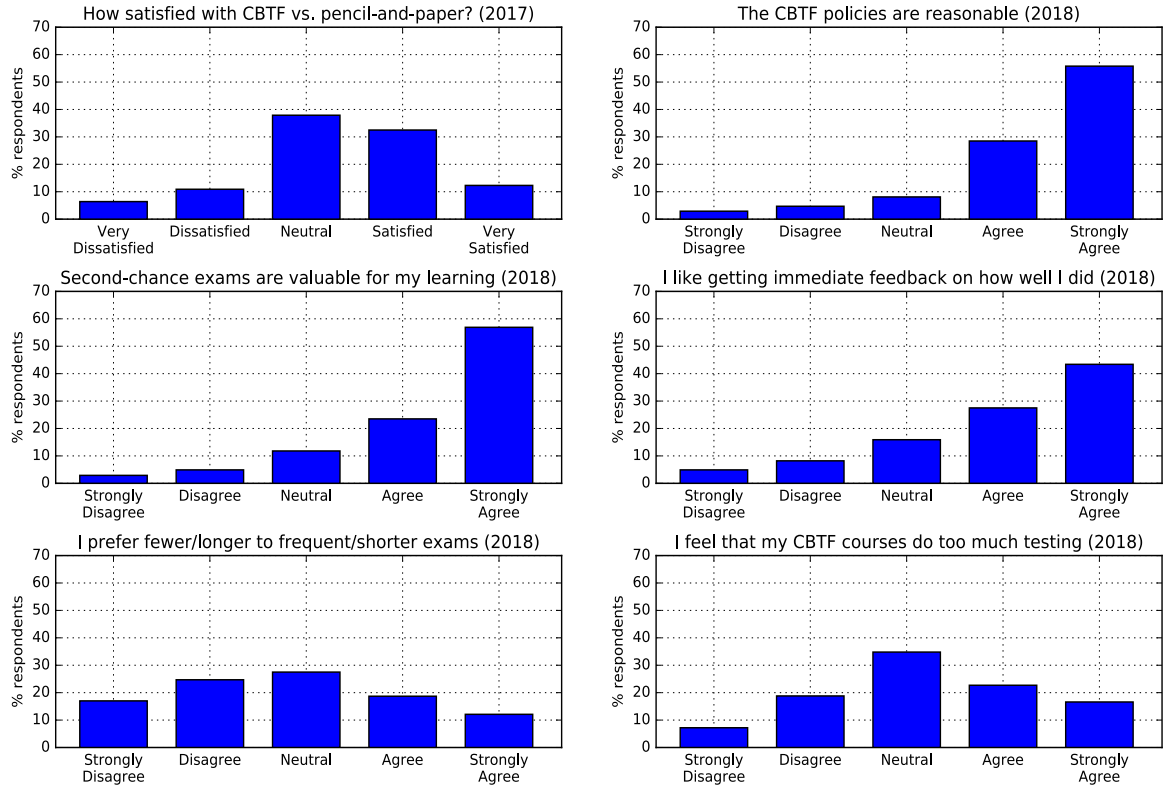

Figure 6. Student survey results relating to overall student satisfaction with the CBTF.

The two most common critiques of the CBTF relate to partial credit, which we discuss in the next subsection, and test anxiety. Topics relating to stress and anxiety show up in the open-ended responses from 34/872 (3.9%) of all survey responses. Another 28 students (3.2%) provided unspecific negative feedback (e.g., “it sucks”, “not my favorite exam environment”), but that is less than the 47 students (5.4%) that provided unspecific positive feedback (e.g., “more exams here would be great”, “satisfied so far!”). Many (27) students complained about the decrepit and unsanitary state of the keyboards and mice in the CBTF due to their heavy use; we plan to replace them all before next semester.

4.2 Computing vs. Non-Computing Students

The 2017 technology survey asked students to report their major, and we found that if we disaggregated the data by major, we had two main constituencies: “computing” majors (Computer Science and Electrical and Computer Engineering) and “non-computing” majors (Aerospace, Agricultural, Civil, Industrial, Mechanical, and Nuclear Engineering, Material Science, and Physics). The disaggregated responses show that the “non-computing” majors were somewhat more evenly split (only 36% satisfied with the CBTF versus 27% dissatisfied), while the “computing” majors were highly satisfied (56% satisfied versus 10% dissatisfied).

We see a similar divergence of opinions with the CBTF in the 2018 Survey, when we asked students how much they agreed with the statement, “I would prefer to take a computer-based exam to a pencil-and-paper exam of comparable length and difficulty.” The aggregated data results in the multi-modal distribution shown in Figure 7(a) with roughly equal numbers of students preferring CBTF exams, preferring pencil-and-paper exams, and having no preference. If we consider the subset of the data in which students identified one of the CBTF classes that they were taking and disaggregate the data based on whether the class was a “computing” or “non-computing” class, we see an explanation for this multi-modal distribution. In Figure 7(b) we see that “computing” students are more likely to prefer a CBTF exam to pencil-and-paper, but Figure 7(c) shows that almost no “non-computing” students prefer CBTF exams.

From the open-ended comments in both of the 2017 and 2018 surveys, we can see that partial credit plays a significant role in this difference of opinion. In the students’ pre-CBTF experiences,

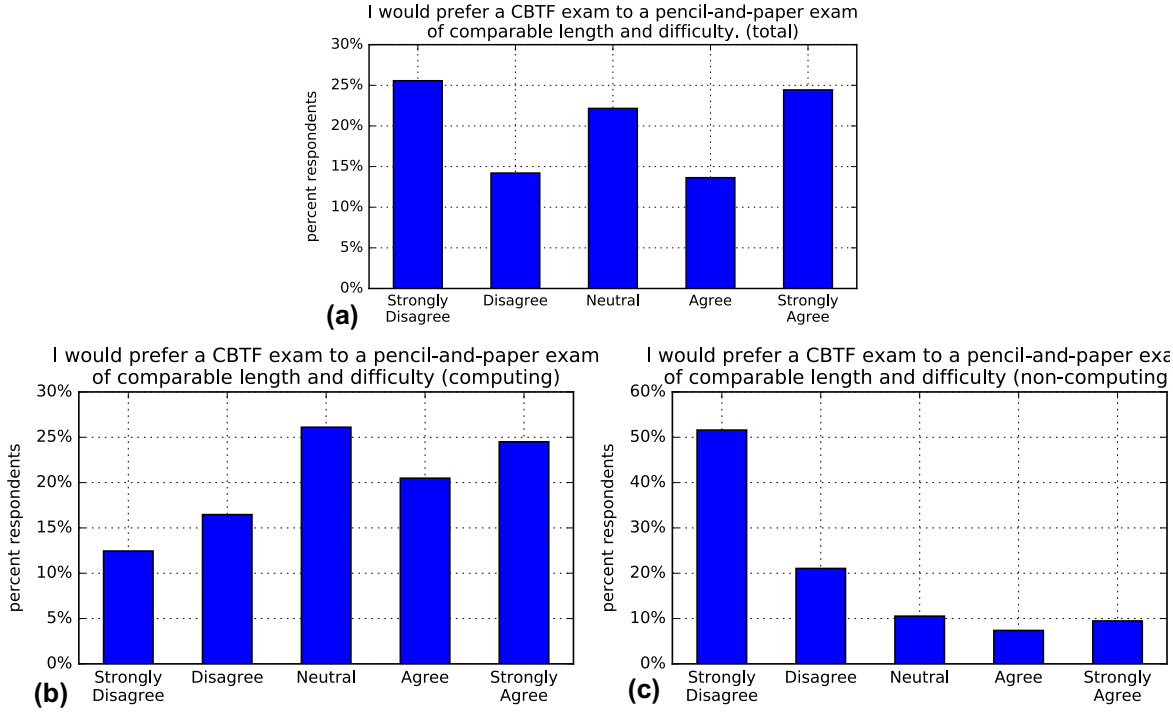

Figure 7. Multi-modal student preferences for CBTF vs. pencil-and-paper exams (a) results from an aggregation of “computing” students’ preference for CBTF exams (b) and “non-computing” students’ preference for pencil-and-paper exams (c).

mathematical and analytical classes (which dominate an engineering curriculum) often have written exams where students are expected to show their work on numeric problems and some partial credit is generally granted (based on the shown work) even if their answer is incorrect. Most classes in the CBTF don't have a mechanism for showing work and instead grant partial credit by allowing multiple submissions for a given problem and reducing the points earned as a function of the number of attempts taken to get the right answer. This policy is used not only because granting partial credit based on students' shown work would be difficult to automate, but also because instructors value the (eventual) arrival to the correct answer, to distinguish those that can solve problems from those that can only regurgitate relevant equations. Statements from students including "it (a CBTF exam) is not an accurate indication of learning as it does nothing to allow for partial credit or show thought process" show that there is a disconnect between students' beliefs about learning and those of instructors. Nevertheless, some students feel like they are giving up a natural right by taking an exam that doesn't meet their expectations for partial credit, and this concern was raised by a significantly higher fraction of non-computing majors (13.8%) versus computing majors (4.4%) in the 2017 survey.

In contrast, many computing students find that computer-based exams actually scaffold their test taking. Coding exams are commonplace in lower-level computing courses. Writing code on paper is notoriously error prone, and even the best students often find that they lose points from small syntax errors. In contrast, the CBTF permits students to compile, test, and debug their code before submitting it, allowing them to find and correct errors before points are lost. Furthermore, students that have done a lot of programming will have had more experience with interfaces that have precise expectations of correctness. Finally, computing students are more likely to be at home using a computer than a non-computing student.

Another, less common complaint, which is also unique to non-computing majors, is about the provided calculators. The CBTF provides Casio fx-300MS calculators (in addition to a software calculator and Matlab), because allowing students to use their own calculators would allow them to bring information into and out of the exam. Nevertheless, some students would prefer their own calculator for its additional power and/or their familiarity with it.

4.3 Perceptions of Cheating

Exam security is important to the CBTF; instructors won't use a form of assessment they don't perceive as secure. Interestingly, our survey of students also indicate that they put a high premium on exam security, as shown in Figure 8. Most students also perceive CBTF exams as at least as secure as pencil-and-paper exams. Student comments explain how the CBTF's physical and electronic security prevents common cheating strategies: "the system is made smartly enough that we can't access other websites, and we can't look at other screens", "people are seated randomly", "cheating is unlikely because people are taking different tests and you can't see anyone else's exams", "the CBTF generates different questions for each person who goes to take it", and "CBTF staff check for cheating more intensely than instructors in regular tests".

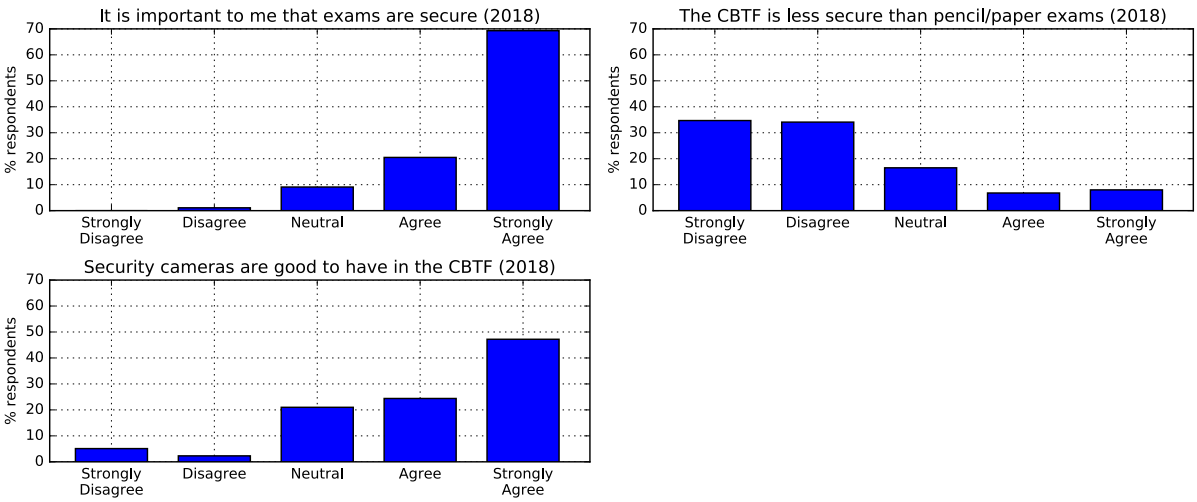

Figure 8. Student surveys show that students care strongly about exam security, that many feel that the CBTF is at least as secure as traditional exams, and that they don't object to security cameras.

There is also a minority of students that perceive CBTF exams to be much less secure than pencil-and-paper exams, entirely because they are offered asynchronously. Students explain: “as soon as people come out of the exam they tell the people they know waiting outside what they did. It’s easy to find out what’s going to be on the exam”, “students can talk about the exam after taking it before their friends take it. If there is not a large question pool, later exam takers have an advantage.” This is a well-known vulnerability and one that instructors are encouraged to address through building exams where different students get different problems, by picking problems randomly from pools. In other work, we’ve shown that cheaters have a significant (12%) advantage over non-cheaters when the whole class is given the same question and students can practice this question outside of the CBTF, even if the question is randomly parameterized [4]. This advantage becomes almost negligible (2%) if instead problems are drawn from a pool of 4 problems. Even when these precautions are ignored, we don’t believe cheating is widespread because of the almost universal trend of exam scores going down throughout the exam period [3].

A few students report that the physical security anti-cheating precautions contribute to making the CBTF into an unwelcoming environment. Students write: “Y’all gotta chill with that security. It feels like I’m entering Guantanamo Bay or something”, and “The workers shouldn’t stare at people taking tests, it’s intimidating”. Other students report that the privacy screens make it difficult to use the computers. Given the importance of maintaining physical security, we’ll need to explore ways to appear more welcoming without sacrificing effectiveness.

5 CONCLUSIONS

The Computer-Based Testing Facility (CBTF) has found rapid adoption in the College of Engineering at the University of Illinois, because it solves problems that instructors face when teaching large courses. We believe that there are many other universities that would similarly benefit from the deployment of a CBTF on their campus. To assist others in their adoption of a CBTF, in this paper we shared our findings from a collection of surveys of students and instructors that have used our CBTF.

While we’ve effectively addressed most of the instructor concerns, our surveys of students indicate that there are two main opportunities for improvement. First, the development of a strategy for granting partial credit that students perceive as fair and compassionate. Second, we need to do a better job of ensuring that instructors adopt best practices with respect to selecting test items randomly from pools, so that student perceive that the right way to study for an exam is to learn all of the material.

ACKNOWLEDGEMENTS

This work was supported by the College of Engineering at the University of Illinois under the Strategic Instructional Innovations Program (SIIP) and by grants NSF DUE 1347722 and NSF CMMI 1150490.

REFERENCES

[1] A. C. Bugbee and F. M. Bernt, “Testing by computer: Findings in six years of use 1982-1988,” *Journal of Research on Computing in Education*, vol. 23, no. 1, pp. 87–100, 1990.

[2] J. Carrasquel, D. R. Goldenson, and P. L. Miller, “Competency testing in introductory computer science: the mastery examination at Carnegie- Mellon University,” in *SIGCSE ’85*, March 1985.

[3] B. Chen, M. West, and C. Zilles, “Do performance trends suggest wide- spread collaborative cheating on asynchronous exams?” in *Learning at Scale*, 2017.

[4] B. Chen, M. West, and C. Zilles, “How much randomization is needed to deter collaborative cheating on asynchronous exams?” in *Learning at Scale*, 2018.

[5] R. F. DeMara, B. Chen, R. Hartshorne, and R. Thripp,“Elevating participation and outcomes with computer-based assessments: An immersive development workshop for engineering faculty,” *ASEE Computers in Education Journal*, vol. 8, no. 3, pp. 1–12, 2017.

[6] R. F. DeMara, N. Khoshavi, S. D. Pyle, J. Edison, R. Hartshorne, B. Chen, and M. Georgiopoulos, “Redesigning computer engineering gateway courses using a novel remediation hierarchy,” in *2016 ASEE Annual Conference & Exposition*, June 2016.

[7] A. Klöckner *et al.*, “RELATE,” https://relate.cs.illinois.edu/.

[8] M. Kuikka, M. Kitola, and M.-J. Laakso, “Challenges when introducing electronic exam,” *Research in Learning Technology*, vol. 22, no. 0, 2014.

[9] E. Lee, N. Garg, C. Bygrave, J. Mahar, and V. Mishra, “Can university exams be shortened? An alternative to problematic traditional methodological approaches.” Kidmore End: Academic Conferences International Limited, 06 2015, pp. 243–250.

[10] R. Muldoon, “Is it time to ditch the traditional university exam?” *Higher Education Research and Development*, vol. 31, no. 2, pp. 263–265, 2012.

[11] C. G. Parshall, *Practical considerations in computer-based testing*. Springer Science & Business Media, 2002.

[12] A. Rytknen and L. Myyry, “Student experiences on taking electronic exams at the University of Helsinki,” in World Conference on Educational Multimedia, Hypermedia and Telecommunications, 2014, pp. 2114– 2121.

[13] M. Shacham, “Computer-based exams in undergraduate engineering courses,” *Computer Applications in Engineering Education*, vol. 6, no. 3, pp. 201–209, 1998.

[14] M. West *et al.*, “PrairieLearn,” https://github.com/PrairieLearn/PrairieLearn.

[15] M. West, G. L. Herman, and C. Zilles, “PrairieLearn: Mastery-based online problem solving with adaptive scoring and recommendations driven by machine learning,” in *2015 ASEE Annual Conference & Exposition*, June 2015.

[16] M. West and C. Zilles, “Modeling student scheduling preferences in a computer-based testing facility,” in *Third Annual ACM Conference on Learning at Scale*, 2016, pp. 309–312.

[17] C. Zilles, R. T. Deloatch, J. Bailey, B. B. Khattar, W. Fagen, C. Heeren, D. Mussulman, and M. West, “Computerized testing: A vision and initial experiences,” in *2015 ASEE Annual Conference & Exposition*, June 2015.

[18] C. Zilles, M. West, and D. Mussulman, “Student behavior in selecting an exam time in a computer-based testing facility,” in *2016 ASEE Annual Conference & Exposition*, June 2016.