Scalable Delivery and Remediation of Engineering Assessments using Computer-Based Assessment

Ronald F. DeMara, Tian Tian, Soheil Salehi, Navid Khoshavi, and Steven D. Pyle
ronald.demara@ucf.edu, Tian.Tian@ucf.edu, soheil.salehi@knights.ucf.edu, nkhoshavinajafabadi@floridapoly.edu, steven.pyle@ucf.edu

Abstract - Inspired by the success of Computer-Based Assessment (CBA) prevalent in disciplines outside of Engineering, a large-scale multiple-year project has been undertaken at the University of Central Florida (UCF) to develop authentic CBA assessments. To-date 29 faculty at UCF have converted their assessments for delivery in a vacated computer lab which was repurposed to create an Evaluation and Proficiency Center (EPC), which realizes proctored testing, autograding, and scanning of scratch sheets for formative and summative assessments. The scrap sheets are utilized to self-motivate learners to revisit the EPC to remediate their missed problems with the assistance of the Graduate Teaching Assistant tutors who were freed up via reduction in their grading loads. Meanwhile, the learners build their soft-skills, as well as their confidence, to convey and explain solutions to engineering problems. Using this EPC-based delivery model, the scalability of the assessment infrastructure is measured and evaluated across seven Engineering degree programs and seen to be capable of harvesting over $400,000 of instructional value annually.

Index Terms - Computer-Based Assessment (CBA), assessment digitization, scalable delivery, exam integrity, tutoring.

EVALUATION AND PROFICIENCY CENTER (EPC) MISSION

Extensions of Computer-Based Assessment (CBA) have begun to be explored for Engineering disciplines [1]. The pedagogy of the Evaluation and Proficiency Center (EPC) approach which realizes proctored testing, autograding, and scanning of scrap sheets for formative and summative assessments has been identified in [2]. Students complete their computer-based quizzes/exams in the EPC during the testing window. The EPC enables students to review their submissions after the testing window closes. Students can receive feedback on their test from an onsite Graduate Scholar Assistant (GSA). If students need additional explanations, they can schedule an appointment to visit their instructor with specific questions. This remedial process aids students to have a preliminary discussion with the GSA, and a targeted discussion with instructor that results in maximized learning and teaching efficiency. The UCF College of Engineering and Computer Science (CECS) EPC is depicted in Figure I. The EPC targets value-based instructional harvesting using a novel cost-saving educational infrastructure for both students and faculty. It recasts GTA and faculty roles of labor-intensive tasks towards high-gain learning activities such as:

• self-paced exam preparation and secure exam delivery,
• digital resources, auto-grading, and score clarification,
• GTA-guided content tutoring, and
• post-test remediation including scanned-in scratch sheets.

Thus, the well-cited “Testing Effect” engages learners with retrieval practice through closed-book proctored quizzes interwoven with rapid tutored remediation. It pools together instructional and human resources (GTAs) from 31 courses across 7 Engineering degree programs to achieve higher learning impact at reduced cost, via rapid student feedback and detailed statistics for instructors to tune their delivery. Trends of increasing enrollment, reduction in cost of PCs, and the success of computerized testing in other disciplines have been motivating its recent research in Engineering [1-8]. A comparison among some digitization approaches within STEM programs is provided in Table I.

FIGURE I
(A) 120-SEAT TESTING AREA AND (B) 30-SEAT AND UNTETHERED TUTORING AREA.
The last row in Table I lists a recent advance in Testing Centers (TCs) which address the needs to support design skills, partial credit, scanned scratch sheets, and remedial tutoring for problem solving within engineering. A comprehensive list of services provided by the EPC infrastructure is presented in Table II. These range from student services such as appointment scheduling, authentication, and stowage of unauthorized materials, pre/post-tutoring, and self-paced solution review. Faculty services provided by the EPC span turnkey original and make-up exam delivery, proctoring, scratch paper scanning, auto-grading, gradebook entry, and video/attendance recording.

Challenges to authenticity of submission and academic integrity within online courses is pervasive in the literature. As listed in Table II, EPC services to fortify the integrity of assessment delivery center around the Test Proctors who validate campus ID cards that are swiped upon check-in to the testing room, and also the tutoring room to track participation. The check-in/out logs are timestamped and electronically filed to record attendance. Access for testing is only granted for those courses which a student is enrolled. Locker use is compulsory for stowage of cell phones, books, and notes. Only non-programmable calculators are allowed and are provided for student use during testing and tutoring activities. The Respondus lockdown browser is used to deliver testing material, which restricts internet access including search engines, email, and messaging. Test questions and solutions are under IP-address restriction, which confines access to quiz questions within the testing room, as well as their solutions. Finally, a dozen point-tilt-zoom cameras provide deterrents and the video footage of suspected violations available to Office of Student Conduct.

The result is that the EPC approach provides a uniform testing environment that realizes quiet, spacious, clean, and consistent assessment conditions. Dedicated PCs in the EPC provide stable testing machines which preclude the presence of keyloggers and malware, and trained GTAs monitor activities and camera feeds, while resolving the test-takers’ concerns. Scratch paper is scanned in after students compose their free-hand design work on the blank sheets that are provided, collected, scanned-in electronically and then shredded. To enable asynchronous multiday testing windows, crosstalk between students is mitigated by disbursing questions randomly from question groups, while instantiating the values within questions randomly across each instance of exam delivery.

**STUDENT AND FACULTY IMPACT**

Starting with a pilot offering in an Electrical Engineering course using a decommissioned computer laboratory, subsequently over the last twelve semesters, 28 additional faculty in other departments were trained to adapt their courses to utilize the EPC’s pedagogical techniques. The EPC expansion project is proposed to create a high-engagement, high-impact, and technology-enabled learning facility. The proposed project would convert an existing open computer lab into an integrated testing and tutoring center in the CECS, beyond the successful two-year 30-seat pilot facility, as shown in Figure 1(B), whose current and future demands greatly exceed its capacity. The function of the EPC is divided into an Evaluation component, to deliver testing, and a Proficiency component, to deliver tutoring. It enhances student experiences by quiet and convenient delivery of exams while ensuring academic integrity in the results obtained. This provides instructional technology resources to increase the effectiveness of in-person GTA tutoring sessions and increase the availability of currently

### TABLE I

INFRASTRUCTURES TO DELIVER DIGITIZED ASSESSMENTS WITHIN ENGINEERING AND SCIENCE CURRICULA [9].

<table>
<thead>
<tr>
<th>Approach</th>
<th>Type</th>
<th>STEM Programs</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>UIUC CBTF [1]</td>
<td>Digitized Assessments Tools &amp; Testing Center</td>
<td>Computer Science, Mechanical Engineering</td>
<td>• Interactive graphical response tool for assessing STEM content</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Unlimited retakes and keeping the highest score</td>
</tr>
<tr>
<td>Univ. of Utah TC [6]</td>
<td>Testing Center</td>
<td>Undergraduate Chemistry</td>
<td>• Difficult topics in general chemistry identified with Item Response Theory</td>
</tr>
<tr>
<td>BYU TC [10]</td>
<td>Testing Center</td>
<td>Physics and Astronomy, Mechanical Engineering</td>
<td>• Symbolic exam problems enable partial credit</td>
</tr>
<tr>
<td>EPC (utilized herein)</td>
<td>Integrated Testing and Tutoring Center</td>
<td>Civil Engineering, Electrical Engineering, Computer Engineering, Computer Science, Industrial Engineering, Mechanical Engineering, Information Technology</td>
<td>• Diverse STEM question formats including Design-by-Selection, Code Completion, Declarative Statement in Multiple Answer Format, Cloning Strategies, etc.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Scanned scratch sheets used for score clarification services and review sessions.</td>
</tr>
</tbody>
</table>
TABLE II
SERVICES PROVIDED TO STUDENTS AND INSTRUCTORS BY THE TESTING CENTER.

<table>
<thead>
<tr>
<th>Service</th>
<th>Methods and Policies for Effective Delivery of Computerized Assessments in Engineering</th>
</tr>
</thead>
</table>
| Appointment Scheduling        | Test scheduling; student-facing website for one-click scheduling of test appointment & email confirmation  
                                | Login credentials; hassle-free login from existing campus computing sign-on page  
                                | Appointment confirmation; emailed to student and cancellation/rescheduling are supported  
                                | Faculty configuration webpage: supports specification and adjustment of multi-day testing windows |
| Student Authentication        | Campus ID cards: swiped upon check-in to testing room & tutoring room to track participation  
                                | University database photo: displayed on PC for verification by attendant  
                                | Check-in/out logs: timestamped and electronically filed to record attendance  
                                | Access: only tests within courses which a student is enrolled become enabled for delivery to that student  
                                | Staffed without cost: check-in services by federal govt-sponsored work-study undergraduate assistants |
| Control of Aides and Materials | Lockers: locker use is compulsory for stowage of cell phones, books, and notes  
                                | Non-programmable calculators and pens: provided for student use during testing and tutoring activities  
                                | Formula sheets: displayed electronically with the Learning Management System (LMS)  
                                | Lockdown browser: used to deliver quiz: no internet access, no search engines, no messaging possible  
                                | IP address restriction: confines access to quiz questions within testing room & solutions within tutoring room  
                                | Cameras: provide deterrent; video footage of suspected violations available to Office of Student Conduct |
| Assessment Delivery and Proctoring | Uniform testing environment: realizes quiet, spacious, clean, and consistent assessment conditions  
                                | Dedicated PCs: provide stable testing machines which avoid keyloggers and malware  
                                | Proctoring: trained graduate assistants monitor activities and camera feeds, and resolve concerns  
                                | Individual work enforced: no talking or collaboration are allowed  
                                | Contingencies handled: exceptions, late arrivals, and external bathroom escort available when needed  
                                | Scratch paper: blank sheets are provided, collected, scanned-electronically and then shredded  
                                | Crossstalk reduction: randomized question groups & instantiated values available to between test deliveries |
| Scoring and Statistics        | Paperless logistics: no photocopies and no miscoded / lost sheets helps to reduce disputes  
                                | Immediate grading: populates course gradebook, which is under mutable control by instructor  
                                | Makeup exams: effortless anytime delivery and Student Disability Support  
                                | Automated regrading: supported for selected question formats  
                                | Statistics provided: test duration of each student and discrimination indices for several question formats  
                                | Abridged grading loads: reallocated to tutoring and proctoring roles for cost neutral operation |
| Review and Remediation        | Self-paced review: solutions under IP restriction within testing facility to restrict dissemination of content  
                                | GTA-assisted review: Socratic discussion of scratch sheet work for partial credit  
                                | Students self-motivated to participate: remediation affords an opportunity to recoup points based on Socratic discussion with GTAs  
                                | Sequencing for efficiency: routine issues resolved in EPC prior to visits to faculty during their office hours. Office visits are refocused on eliminating the students’ knowledge gaps and actively assisting their learning |

staffed GTA’s by providing online tutoring sessions in conjunction with a Microsoft Surface tablet screen-sharing using electronic pen annotations. Figure I(B) depicts learners conducting secure self-paced review of their formative quizzes (in foreground of image), while raising hands to summon tutors (in background) who engage Socratic questioning [11] to gain partial credit based on scanned scratch sheets.

**Increasing Access, Humanity Benefits, and Furthering Wealth of Learner Diversity**

UCF’s innovative EPC infrastructure is realizing a novel high-efficiency teaching pedagogy to accommodate enrollment growth. It is also enhancing learning quality including an individual’s and community-wide well-being in the UCF experience. While mitigating cost, the EPC is significantly alleviating grading workloads and errors, thus substantially elevating faculty and GTA effectiveness, morale, happiness, and job satisfaction. It is also equipping graduate assistants to become more thoroughly trained faculty of the future via conducting tutoring in lieu of rote grading. The EPC remains vigilant to accommodating SAS/SDS needs through its website-based scheduling system. For instance, even disposable earplugs for distraction-free test-taking are provided free via a personal funds expenditure of the EPC Faculty Director, as well as calculators, pens, and lockers at no charge.

Finally, the EPC leverages an ensemble of GTAs pooled together across a variety of degree programs to cultivate diversity of thought and delivery. Their collective instruction strengthens understanding throughout UCF’s diverse student...
population. Expansion is underway for faculty participation further into STEM interdisciplinary programs for Math and Science faculty at UCF. Another exciting option could be to organize the First Annual International Workshop on Scalable Learning Technologies, which could be convened at UCF. Table III indicates that all of these benefits are attained while harvesting $444,969 of instructional value annually, which can increase further as the EPC continues to be scaled-up.

As shown in Figure IV, the cost savings that increase with enrollment can occur in three levels. In the Self-Paced level, the extensive resources related to the course content are available in the flexible time windows for the students. Students can solve additional problems in this step to gain more knowledge about the concepts of the covered topics in the class. In the Graduate Scholar Assistants (GSAs) level, students can get help from the onsite GSA to solve examples of Lecture & Study Sets in EPC. The GSA provides first-responder support in a learning hierarchy for grading concerns. The GSA clones prototypes and composes solutions. In addition, the GSA holds office hours in EPC to deliver engineering design guidance and debugging. In the Instructors level, the class-wide learning effectiveness, the EPC tutors, and the EPC quizzes are managed by the instructor and the EPC manager. In addition, the instructor refines the materials for open & closed quiz formats.

**Case Study: EML4142-Heat Transfer Course**

The flow depicted in Figure V shows how students progressed through Knowledge Acquisition, Assessment, and Knowledge Refinement activities in the EML4142-Heat Transfer course at UCF. Prior to lectures, students conducted online reading assignments and unproctored practice problems via McGraw-Hill LearnSmart® which guided students to read the textbook through adaptive practice questions. Homework assignments were also delivered online through McGraw-Hill Connect. Proctored formative assessment including four 40-mins quizzes and two 110-mins midterms were then conducted, followed by a summative final exam. The histogram of all quiz submissions and Midterm Exam-2 submissions from both Paper-Based Assessment (PBA) and CBA cohorts are depicted in Figure VI(A) and Figure VI(B), respectively. The score distributions between the two cohorts were generally comparable and the CBA distribution was close to a normal distribution. This dissuades the view that CBA may result in only bimodal score distributions. It instills confidence that CBAs for STEM content can be relied on to perform similar to expectations of PBAs for STEM content.

Among enrolled 118 students, 55 students in each cohort, 110 students in total, attended the two Midterms in both delivery modes. Treating each individual student as his or her own control, Table IV summaries that 46 (41%) students scored A or B in both locations, 72 (64%) students
scored C or above C in both locations, and 8 (7%) students scored D or below D in both locations. In total, 73% (80 out of 113) of the students scored consistently in both PBA and CBA. Satisfactory consistency was obtained with PBA and CBA despite inadvertent grade inflation within the CBA delivered Midterm Exam-1, which occurred due to the lack of sufficient distracting factors to dissuade the impact of guessing.

Examining student performance in Cohort A, it was observed that 28 (51%) students achieved higher scores in the paper-based Midterm Exam-1 and 27 (49%) students outperformed in the computer-based delivery format during Midterm Exam 2. This symmetry further suggested that well-constructed computer-based exams can effectively assess student achievement in viable means which are comparable to paper-based exams. In contrast, 36 (62%) students in Cohort B scored higher in the computer-based delivery of Midterm Exam-1, while 18 (31%) students in Cohort B scored higher in the paper-based Midterm Exam-2 with four (7%) students scored the same. This difference in delivery modes likewise reflects the slight grade inflation inherent in the CBA delivery of Midterm Exam-1 prior to distractor tuning.

1. Workload Analysis

Table V listed the estimated GTA hour allocation for both cohorts. In this intervention semester specifically, there were 58 students in each of the PBA and CBA cohorts. In total, 51 Graduate Teaching Assistant (GTA) hours were expended for grading the submissions of the PBA cohort.

---

**Figure V**
Assessment Summary. Percentages indicate weight of each assignment towards course grade.

---

**Figure VI**
Grade distribution (A) All PBA and CBA quiz submissions (B) Mid-term 2 PBA and CBA submissions.
Therefore, the CBA cohort could potentially have received more GTA support for higher gain tasks such as score clarification and tutoring. In the authors’ university, typical annual enrollment of this course is 450 students per year. By delivering all quizzes and midterms via CBA, 395 GTA hours could be harvested each year, which would be equivalent to $7,900’s worth of tutoring value when considering GTA salary is roughly $20 per hour. This amount is stipend alone, exclusive of tuition costs for the graduate assistant, and other fees incurred.

II. Instructor Evaluation

Instructor reflection on time logs and narrative perceptions was positive. The instructor of EML4142 begins by noting her initial apprehension that PBAs are inherently amenable to partial credit due to manual grading. The instructor however became increasingly aware and ultimately convinced, that carefully designed step-wise computerized questions can remain authentic to the characteristics of the technical content and conduct an accurate assessment of the learners' skills. Based on CBA delivery, the authors gleaned valuable insights hidden within exam submissions which were not previously examined due to high grading loads. The rapid statistics facilitated by CBA increased opportunities to recognize repeated misconceptions, analyze student learning, and adapt teaching accordingly while concepts were still fresh in the learners’ minds.

Automation was measured to reduce the burden of manual grading as quantified above, while also noted to increase scoring accuracy and consistency. The demands of trying to interpret the final answer and how to grade it were reduced significantly. CBA was also valued by the instructor as a new means to facilitate rapid remediation and reduce logistic overheads that crosscuts the engineering instructional ecosystem. The EPC-based delivery is seen to realign educational and human resources without a net personnel increase by reallocating low-gain grading tasks to high-gain activities such as tutoring, lab assistance, and remediation. The instructor noted additional effort is required during digitization period of roughly one hour spent per question in total. However, a substantial advantage obtained is that the CBA questions can stay locked from future students via IP address restriction, including during solution review within the EPC. Thus, questions can become more fully-refined each term, instead of an endless treadmill of constantly creating similar-but-distinct questions. This yields another valuable time-saving aspect for faculty. It also improves the student experience with a growing bank of validated, vetted, and unpublished questions.

CONCLUSION

This paper explores the use of Computer-Based Assessments (CBA) within engineering courses in comparison with Paper-Based Assessments (PBA) in a pilot study. According to our study, use of CBA will result in an effective and authentic evaluation relative to traditional PBA. However, in order to fully implement and utilize CBAs, instructors are required to effectively reconstruct engineering problem statements to be appropriate for CBA approach. In addition to CBA’s auto-grading, in order to achieve the advantages of PBA method, which is retaining flexibility for partial credit and self-incentivizing learners to defend their work, use of novel hybrid machine/human-based two-phase grading methods, such as use of scanned scratch papers, as well as using a hierarchy of the freed GTAs’ and instructor’s expertise can be advantageous. Furthermore, there is supportive evidence that significant Graduate Teaching Assistant hours can be harvested for higher-gain learning tasks. In conclusion, the approach is promising and is being further scaled to other disciplines beyond Engineering by expanding the 6-week faculty development course used to train new faculty how to digitize their assessments [12]. These training materials are available to interested readers by contacting the authors to request them.

<table>
<thead>
<tr>
<th>TABLE IV</th>
<th>TABLE V</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBA AND PBA SCORE UNIFORMITY. SUBMISSIONS SCORING ACROSS MIDTERMS 1 AND 2.</td>
<td>GRADING WORKLOADS AND EXTENSION OF OFFICE HOURS</td>
</tr>
<tr>
<td>Achievement Range</td>
<td>Week</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of students who scored A or B in both CBA and PBA settings</td>
<td>2, 6, 10, 11</td>
</tr>
<tr>
<td>N=19 (34.5%)</td>
<td>4, 8</td>
</tr>
<tr>
<td>N=32 (58.2%)</td>
<td>3, 5, 7, 9</td>
</tr>
<tr>
<td>Number of students who scored D or F in both Exams</td>
<td>12</td>
</tr>
<tr>
<td>N=6 (10.9%)</td>
<td>Total</td>
</tr>
<tr>
<td>N=2 (3.4%)</td>
<td></td>
</tr>
</tbody>
</table>

| Number of students who scored A or B | N=55 | N=55 |
| N=19 (34.5%) | N=27(46.6%) |
| N=32 (58.2%) | N=40 (69.0) |
| N=6 (10.9%) | N=2 (3.4%) |
REFERENCES


AUTHOR INFORMATION

Dr. Ronald F. DeMara is a Professor of Electrical and Computer Engineering at the University of Central Florida where he has been a faculty member since 1992 and is Digital Learning Faculty Fellow. His educational research interests focus on classroom and laboratory instructional technology, digitally-mediated team learning, and the digitization of STEM assessments. He has completed over 260 technical and educational publications, 47 funded projects as PI/Co-PI, and established two research laboratories. He serves as the founding Director of the Evaluation and Proficiency Center (EPC) at UCF and is the recipient of UCF’s university-level Scholarship of Teaching and Learning Award, Teaching Initiative Program Award, Research Initiative Award, Excellence in Undergraduate Teaching Award, Advisor of the Year, Distinguished Research Lecturer, Marchioli Collective Impact Award, and is an iSTEM Fellow. He received the Joseph M. Bidenbach Outstanding Engineering Educator Award from IEEE in 2008 and Effective Practice Award from the Online Learning Consortium in 2018.

Dr. Tian Tian is an Associate Lecturer of Mechanical and Aerospace Engineering at the University of Central Florida, which she joined in 2013. She has been frequently teaching undergraduate lecture and laboratory components of Heat Transfer, Thermodynamics and Fluid Mechanics. Her educational research interests focus on project-based learning, online learning, and the digitization of STEM assessments. She received the Teaching Incentive Award, Excellence in Undergraduate Teaching Award, the Dean’s Advisory Board Faculty Fellow, Professor of the Year Award and Advisor of the Year Award.

Soheil Salehi received his M.Sc. degree in Computer Engineering in 2016 from Department of Electrical and Computer Engineering of the University of Central Florida, Florida (UCF), USA. He is a Graduate Research Assistant (GRA) and currently working toward the Ph.D. degree in Computer Engineering at the University of Central Florida, Orlando, Florida. His research interests include: Reconfigurable and Adaptive Computer Architectures, Spintronic-Based Computing Architectures, and Low Power and Reliability-Aware VLSI circuits. He has also been a Graduate Teaching Assistant (GTA) for Department of Electrical Engineering and Computer Science of UCF from 2014 to 2018. His educational interests are innovations and laboratory-based instructions, technology-enabled learning, and feedback driven grading approaches. He is the recipient of the Award of Excellence by a GTA for the academic year of 2015-2016 at UCF.

Dr. Navid Khoshavi received his Ph.D. degree in Computer Engineering from the University of Central Florida in 2017. He joined Florida Polytechnic University in Fall 2017. His research interests are design of Fault-tolerant Hardware Architectures and Domain-Specific Architecture for Machine Learning Applications. He has received multiple awards recognizing his research initiative and academic success.

Steven D. Pyle received the M.Sc. degree in Electrical Engineering at the University of Central Florida and is currently a Ph.D. candidate in Computer Engineering at the University of Central Florida. His research interest lies in bridging the multidisciplinary fields of neuromorphic hardware with emerging devices, machine learning, computational neuroscience, and neurophysiology to realize brain-inspired computational hardware at ultra-low-power.