

Scalable Exam Administration: The Proctoring Crisis Requires a SEA Change

Scalable Exam Administration

The Proctoring Crisis Requires a SEA Change

Armando Fox

University of California, Berkeley, fox@berkeley.edu

Craig Zilles

University of Illinois at Urbana-Champaign, zilles@illinois.edu

Two converging pressures are creating a “proctoring crisis” for high-stakes exams. First, the number of students qualifying for exam-related disability accommodations has grown explosively. At most US institutions, these students are proctored by campus disability support units — which we will call Student Accommodations Centers (SACs) — whose facilities and protocols were designed for a small population, not for the 40% of students now identifying as having a disability, as has been documented at Stanford¹. Second, the rise of generative AI has made it easier than ever for students to complete homework without actually learning the material. The natural institutional response — reducing homework weight and increasing exam weight — raises the stakes of exams and, perversely, intensifies the temptation to cheat on them. These conditions have created a crisis in how we can reliably assess what students have learned.

To address both problems, we propose four principles for Scalable Exam Administration (SEA), which replace the non-scalable mechanisms of traditional exam administration (TEA): Not Simultaneous; No Clarifying Questions; No Books, Cheat Sheets, or Notes; and Uniform-Length Exams.

We explain why each principle matters individually, and how together they enable computer-based exam administration that resolves substantially all of the scaling problems inherent in TEA.

CCS CONCEPTS • Insert your first CCS term here • Insert your second CCS term here • Insert your third CCS term here

Additional Keywords and Phrases: Insert comma delimited author-supplied keyword list, Keyword number 2, Keyword number 3, Keyword number 4

ACM Reference Format:

Armando Fox and Craig Zilles. 2026. Scalable Exam Administration: The Proctoring Crisis Requires a SEA Change.

¹ <https://fortune.com/article/rise-in-elite-students-seeking-accomodation-gen-z-phenomenon-find-success-in-competitive-job-market-stanford-university-skills-based-hiring/>

1 NOT SIMULTANEOUS

Most instructors design exams to be taken simultaneously to avoid "information leakage". This approach creates serious obstacles to scaling. A recent STEM course with 1,400 students required 18 classrooms and 18 proctors!² Beyond logistics, some students will inevitably have scheduling conflicts — known in advance or not — forcing the instructor to administer separate sittings while managing the risk of information passing between them.

The only scalable solution is to **decouple exam integrity from the requirement that everyone sit at the same time**. Some essay exam prompts in the humanities are customarily provided to all students in advance anyway, but in fields where this is not the norm, a more scalable-by-design alternative is to make it safe for students to take the exam at different times by **giving each student a different exam**. We can do this using one or both of the following approaches, which are cumbersome with paper exams but straightforward with computer-administered exams. First, draw each exam question from question pools, where each pool covers a single (or a few related) learning objective(s) at a particular difficulty. Second, create question generators—short bits of code that generate a range of parameterizations or configurations of a given problem in a controlled manner, such as generating a range of similar circuits or graphs for an analysis problem and randomizing the numeric parameters associated with the problem. Our research suggests [3] that having 3-4 question generators for each question “slot” on the exam is sufficient to make exams sufficiently hard to game that the rational strategy for a student is to just learn the material. Both approaches require investing time in a new exam-creation workflow, but the results amortize well, as exams can be more safely reused.

2 NO CLARIFYING QUESTIONS

Because proctors are traditionally course staff, students are accustomed to asking clarifying questions during exams, either because they are legitimately confused by a question or in hopes of gleaning hints. This practice already breaks down when students are proctored at a SAC where no course staff are present, and fails entirely under asynchronous administration. Instead, students should be able to flag issues during the exam — in computer-administered systems, this can be built directly into the software — so that instructors can review flags after the fact, or even in real time so that later exam-takers see clarified questions. Under this model, proctors need only be trained in general invigilation procedures, making them interchangeable across courses. Students taking exams from different courses can be proctored together in the same room at the same time, dramatically simplifying scheduling and lowering costs.

3 NO BOOKS, CHEAT SHEETS, OR NOTES

When proctors are not course-specific, the only scalable policy is to prohibit all physical materials: a proctor cannot be expected to know and enforce the cheat-sheet rules for every course. The rule becomes simple — no paper of any kind enters or leaves the exam room. Instead, both instructor-authored and student-authored reference materials are made available digitally, within the exam software, as appropriate for each course. Scratch paper is provided by proctors and collected at the end of the session; document cameras at each testing station allow students to photograph handwritten work and upload it for grading.

² Alexis Shusterman, UC Berkeley lecturer, personal communication, March 2026

4 UNIFORM-LENGTH EXAM SLOTS

Standardizing exam length is essential for scalable scheduling. Our campuses have standardized on one-hour exams during the semester and two-hour finals, on a fixed schedule. Uniform-length slots simplify the bin-packing problem associated with exam scheduling, enabling high utilization of the facility. They also minimize interruptions during an exam by aligning all exams' start and end times.

Many courses currently rely on one or two heavy midterms, but we argue this is largely a consequence of exams being painful to administer. When administration becomes easy, the better pedagogical choice becomes available: replacing a two-hour midterm with several shorter, lower-stakes exams, which research has shown produces equal or better learning outcomes and retention than massed studying, while also reducing exam anxiety. Also, courses running more frequent, short exams distributes the testing load more evenly over the semester, alleviating bursts from high-demand periods.

5 SEA AND COMPUTER-BASED TESTING

Computer-based assessment and SEA are mutually reinforcing. In addition to allowing unique exams per student, computer-based exams can be pedagogically richer than paper exams. Using a Web browser to present exam questions and collect student responses in digital form not only permits judicious use of auto-grading, but also enables much more complex and authentic question types: we have built question generators and auto-graders for symbolic math, circuits, finite-state machines, free-body diagrams, pipeline diagrams, and graphed functions.

The following four steps describe how SEA operates in Computer-Based Testing Facilities (CBTFs), which have been running at the authors' institutions and others for several years:

1. Each exam is open for a three-day window. Students reserve a time of their choosing within that window. Windows for different courses overlap, so a group of students arriving at any given time will generally be taking different exams.
2. Proctors check student IDs, ask students to put personal belongings and devices in a storage area at the front of the room, and direct each student to their assigned seat. Students taking the same exam are seated apart. Students requiring reduced-distraction environments are automatically assigned to appropriately-configured stations by the scheduling software at reservation time³. Students are offered scratch paper, which must be discarded on exit.
3. At the scheduled time, an exam is generated for each student, and they work for their individually allotted amount of time. Instructor-approved reference materials are provided digitally with the exam. Because the institutionally-managed computers restrict access only to allowed resources, proctors primarily monitor for suspicious behavior and use of disallowed materials, such as personal devices or paper cheat sheets. Two proctors can manage rooms of up to 100 students. If a proctor observes something suspicious, video camera footage can be reviewed to document the incident for the instructor.
4. Students leave as they finish. By the end of the hour, only students with extended-time accommodations remain, continuing into the next hour as their adjusted timer allows. The next group of students is waiting outside, and the cycle repeats from step 1.

UIUC operates several CBTF spaces this way 80 hours per week, including weekends; Berkeley is at 40 hours per week and planning to increase to 70-80.

³ CBTFs at the authors' institutions can handle nearly all exam-related accommodations without any effort by course staff or the SAC other than entering the accommodations into the scheduling software.

6 BENEFITS: FREED CLASSROOMS, BETTER OUTCOMES, LOWER COSTS

Hundreds of thousands of exams have been administered securely and economically to tens of thousands of students at an amortized cost of just a few US dollars per student-exam-hour [1]. Students with and without accommodations gain greater scheduling flexibility. Course staff are relieved of the administrative burden of running exams that offer weaker integrity guarantees than trained proctors can provide. Students are on notice: if they rely on AI to do their homework without learning the material, their exam performance will reflect it.

We note, though, that we began this work pre-AI with the goal of enabling *positive* pedagogy, not punishing the misuse of AI. Research confirms that frequent, short assessments, with the possibility of re-takes, are a key ingredient of mastery learning and enhance both retention [2] and equity [4]. The SEA approach has also been demonstrably better for mental wellness of both students [5] and instructors. CBTFs are a key SEA enabler that has helped us realize these benefits.

We emphasize that SEA is *not* simply an expanded version of SAC proctoring: It *replaces* the non-scalable aspects of exam administration at their root, centralizing the effort and resources required to run a secure exam rather than having each instructor repeat the process inefficiently for every sitting.

The proctoring crisis is real. But the answer is not a return to blue books. Faculty at our institutions who have adopted CBTF-based SEA have no desire to go back — and we are confident that similar results are achievable elsewhere.

REFERENCES

- [1] Jim Sosnowski, Julie M Baker, Olivia Arnold, Mariana Silva, David Mussulman, Craig Zilles, Matthew West . Reflections on 10 years of Operating a Computer-based Testing Facility: Lessons Learned, Best Practices. Proc. ASEE 2024.
- [2] Jason W Morpew, Mariana Silva, Geoffrey Herman, Matthew West. Frequent mastery testing with second-chance exams leads to enhanced student learning in undergraduate engineering. Applied Cognitive Psychology, Volume 34, Issue 1, Pages 168-181, 2020.
- [3] Binglin Chen, Matthew West, and Craig Zilles. How much randomization is needed to deter collaborative cheating on asynchronous exams? In Proceedings of the Fifth Annual ACM Conference on Learning at Scale (L@S 2018), New York, NY
- [4] Thomas R. Guskey. Closing achievement gaps: Revisiting Benjamin S. Bloom's "Learning For Mastery". Journal of Advanced Academics, 19(1):8–31, 2007.
- [5] Chinedu A. Emeka, Craig Zilles, Matthew West, Geoffrey L. Herman, Timothy Bretl. Second-Chance Testing as a Means of Reducing Students' Test Anxiety and Improving Outcomes. Proc. ASEE 2023.