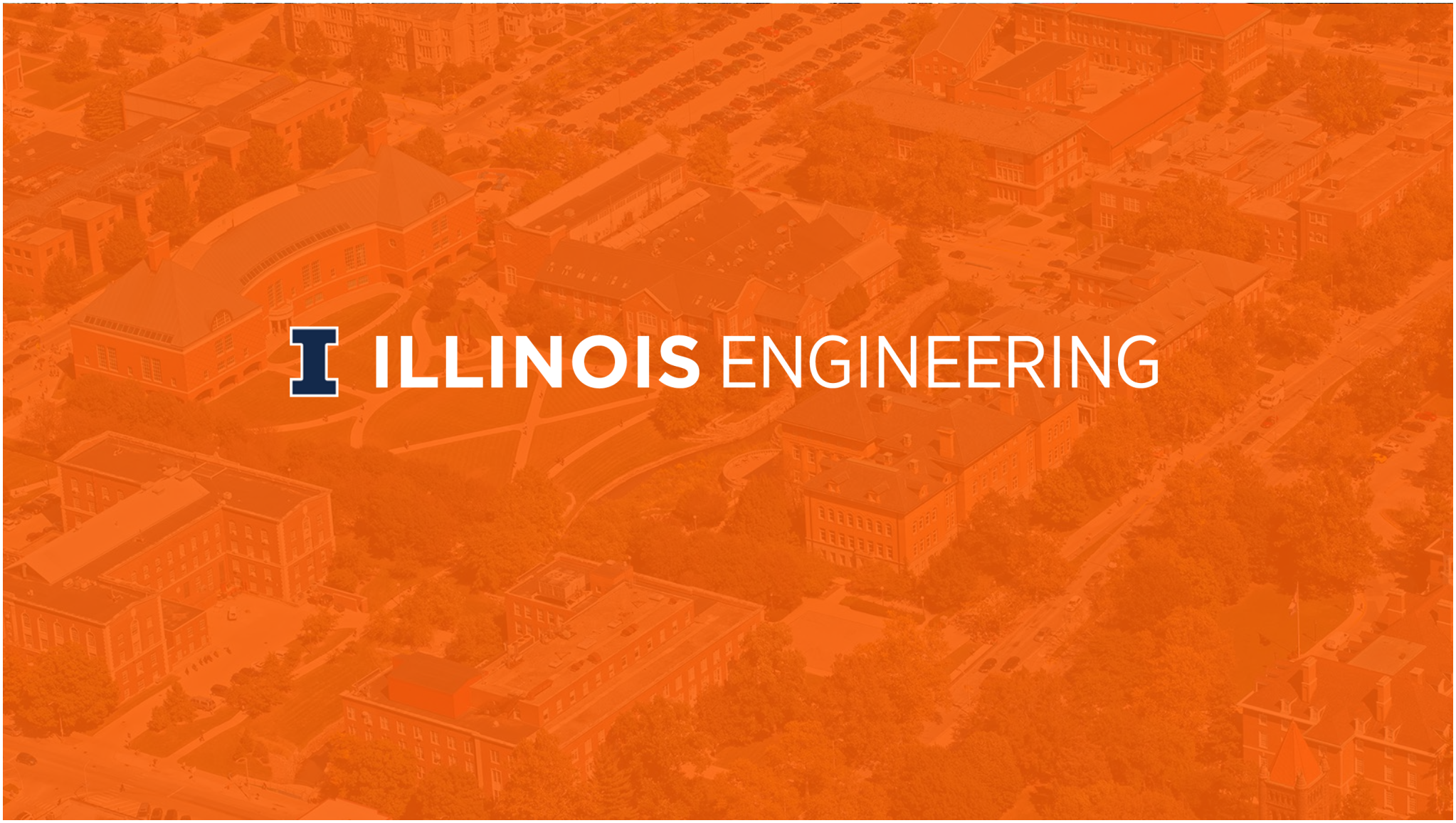
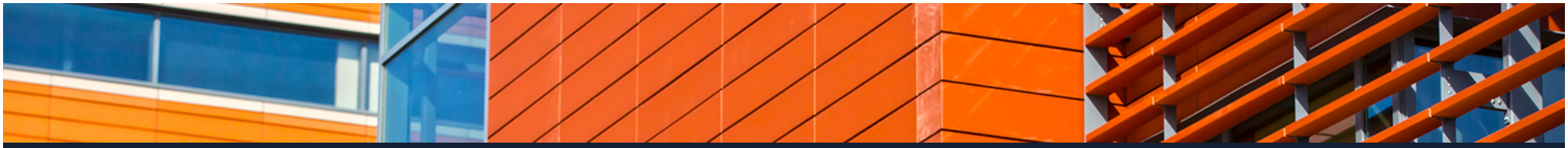

ILLINOIS ENGINEERING



Effective, Secure, and Efficient Summative Assessment using a Computer-based Testing Facility

Craig Zilles (Computer Science), zilles@illinois.edu

https://cbtf.engr.illinois.edu

October 7, 2019

ILLINOIS ENGINEERING

Outline

- **Motivation**
- **The Computer-Based Testing Facility (CBTF)**
- Operational details (security, testing accommodations, cost, etc.)
- What do exams look like? (the PrairieLearn LMS)
- Asynchronous exams, student behavior, and randomization
- Increased testing frequency and improved learning outcomes

Background & Motivation

Illinois is a large Engineering school

- 2nd largest in the U.S.
- 2,111 bachelor's degrees (in Engineering) graduated in 2016

Many foundational CS, Engineering classes are large

- ~1000 students/semester in **calculus-based mechanics** (Physics)
- 800 students/semester in **data structures** (Computer Science)
- 500+ students/semester in **introductory statics** (Mechanical Eng.)

Large classes -> Traditional exams are painful

- Printing and proctoring exams across many rooms
- Student conflicts, illness, testing accommodations
- Grading effort, inconsistency, slow feedback
  - (or heavy reliance on multiple choice)

- Sufficiently painful that most faculty do it very infrequently
  - Which is bad for student learning (testing effect, e.g., Roediger and Karpicke)

- Wanted assessment strategy that scaled **without** sacrificing quality

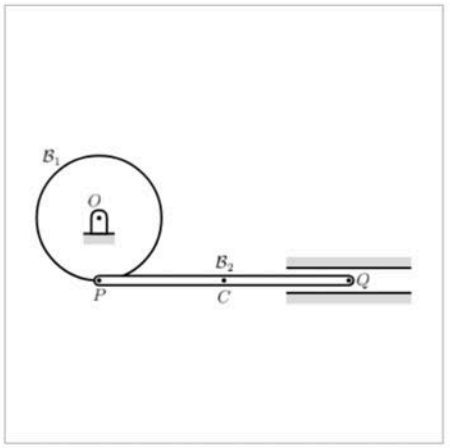
Our Solution: Computer-Based Testing Facility

Consists of two parts:

1. Sophisticated computerized problems
   - Randomized parameters, Auto-graded
   - Numeric, symbolic, code writing, etc.

PrairieLearn LMS
Open source: https://github.com/PrairieLearn

Body $\mathcal{B}_1$ is a uniform rigid disk that rotates about the fixed center O as shown, driven by a pure moment $\vec{M}$. Body $\mathcal{B}_2$ is a uniform thin rigid rod that connects pins P and Q (with center C), and point Q is constrained to move in a slot. Forces $\vec{F}_P$ and $\vec{F}_Q$ are exerted on the rod by pins P and Q, respectively, and there is no gravity or friction.



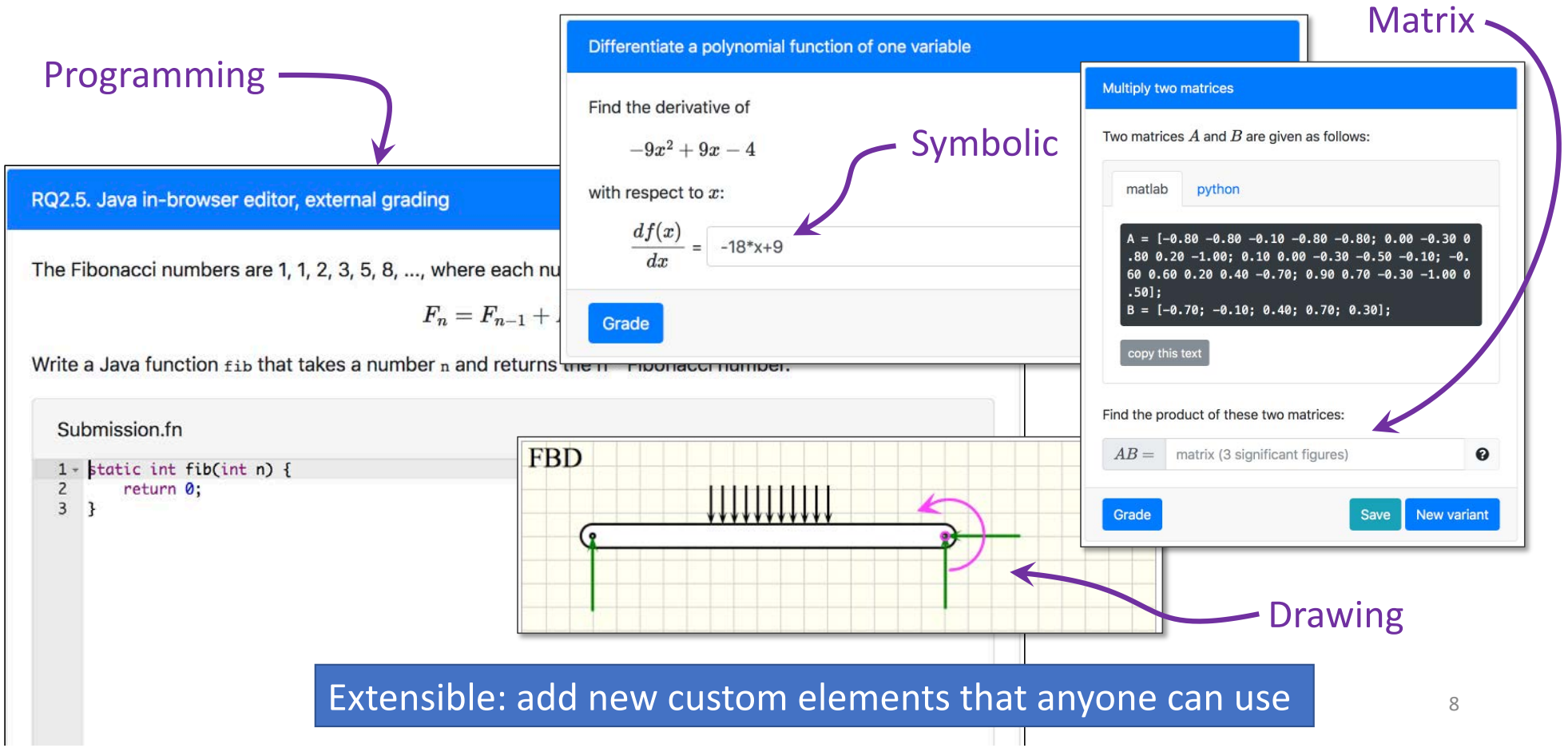
At the current instant we have:

$m_1 = 4$ kg $\quad m_2 = 6$ kg
$I^1_{O,\hat{k}} = 12$ kg m^2 $\quad I^2_{C,\hat{k}} = 32$ kg m^2
$\vec{\omega}_1 = -3\hat{k}$ rad/s $\quad \vec{\omega}_2 = 0$ rad/s

What is $\vec{a}_P$?

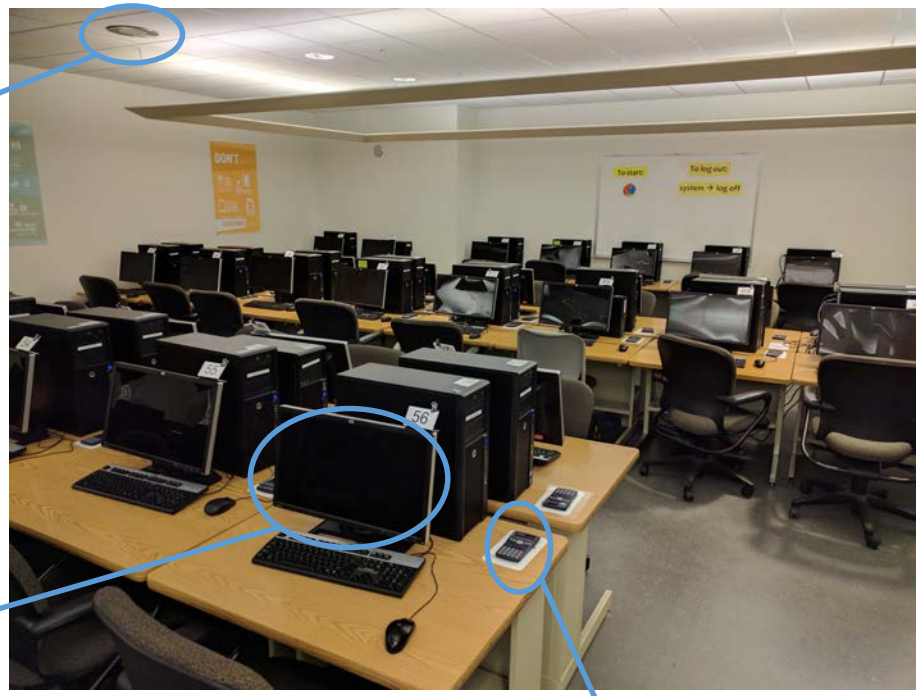
$\vec{a}_P =$ [ ] $\hat{\imath} +$ [ ] $\hat{\jmath}$ m/s^2

Rich set of possible problem components



Extensible: add new custom elements that anyone can use

Part 2: A secure computer lab



Security cameras

Professional proctors

Privacy screens

Calculators

Firewalled internet

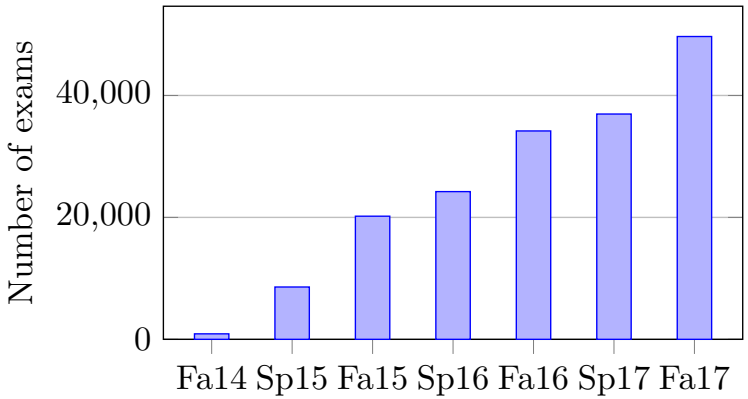
Full software platform (Python, RStudio, etc)

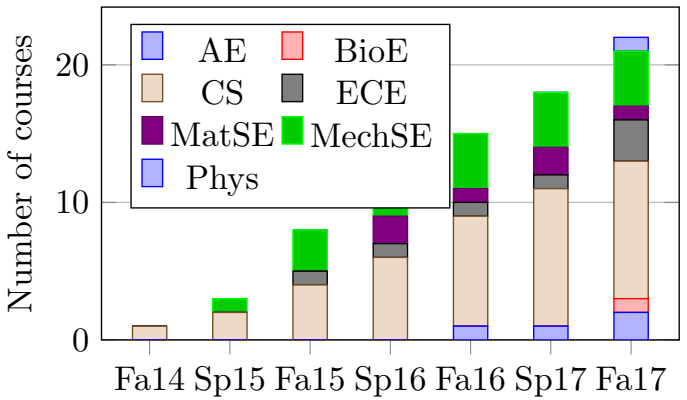
ID card swipe to check in

85-seat room
Open: 12 hours/day
7 days/week

Entering our 6th year this Fall

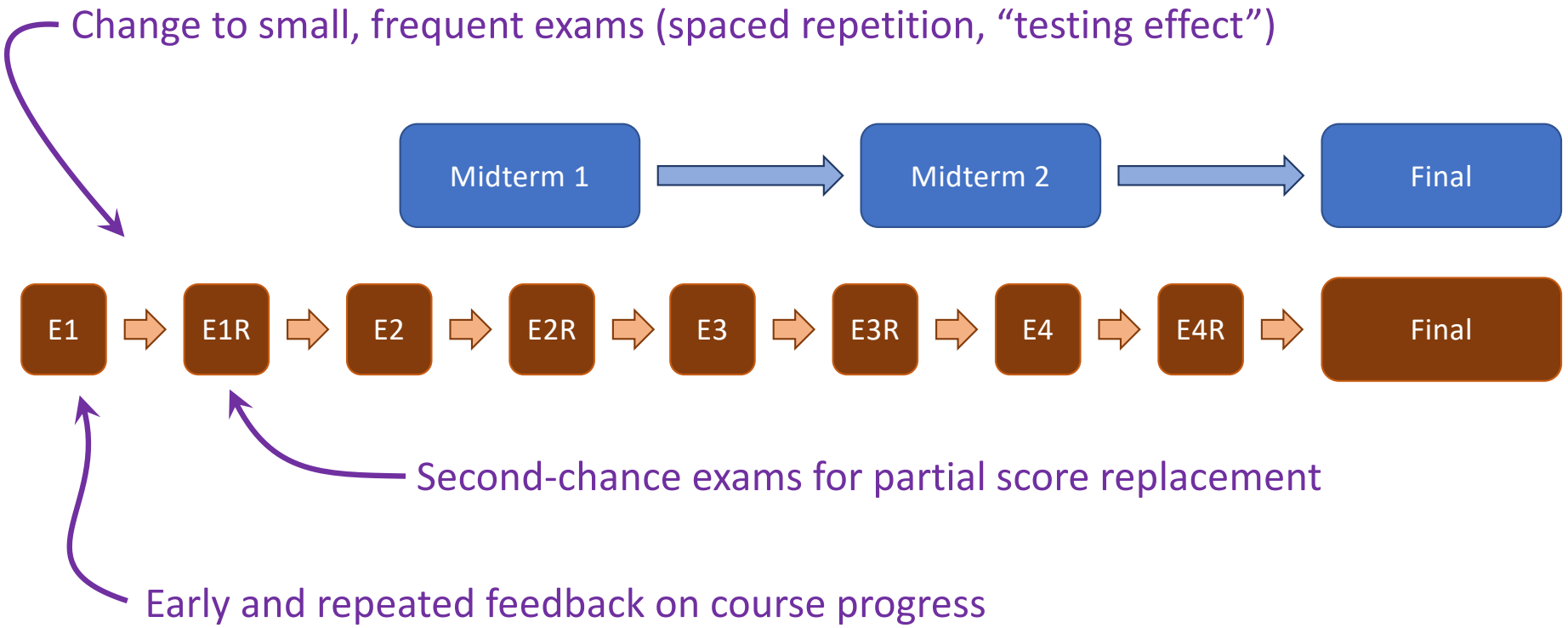
- Now used by many large Engineering/CS courses (29 courses in Fall 2018)





- Over 50,000 exams/semester and 6,000 unique students (last 5 semseters)

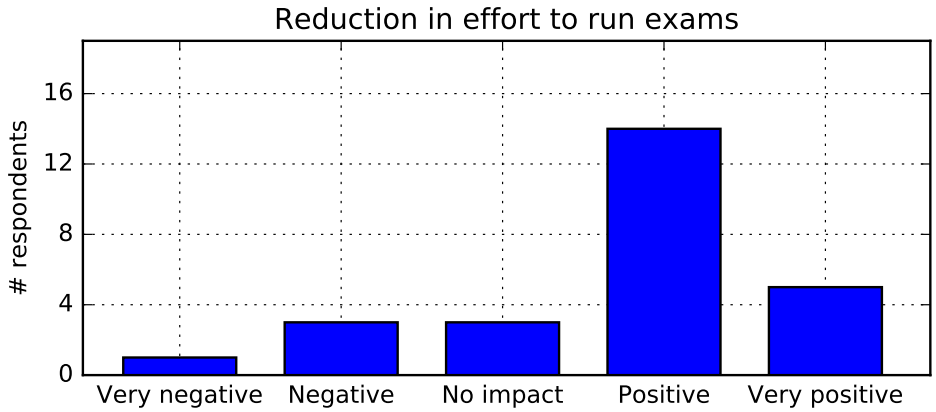
Changing the way we test

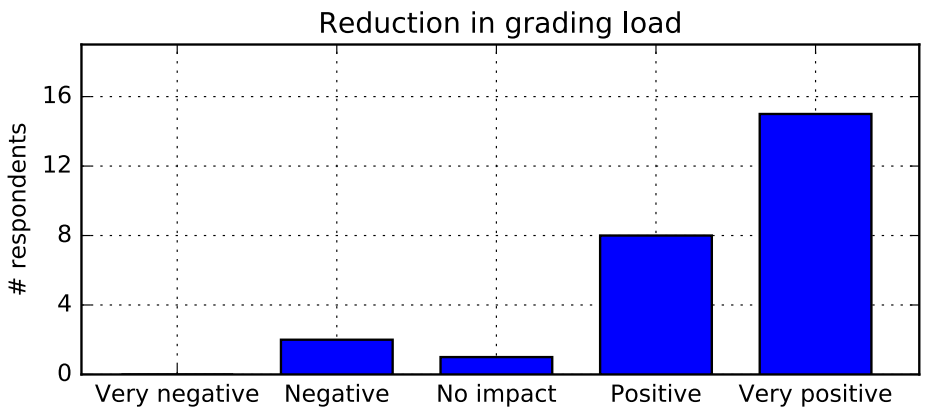


When assessment is cheaper…

Zilles, et al. EDULEARN 2018

Faculty Survey





*"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."*

Frees up staff time for higher value activities

- Less (recurring) time spent developing exams, proctoring, grading

- More office hours
- Re-introduction of written lab reports (MechSE) and term projects (Aero)
- Adopt more active learning pedagogies

- Key goal:

Automate what can be improved by automation, to free up time for the important things that are best not automated

Outline

- Motivation
- The Computer-Based Testing Facility (CBTF)
- **Operational details (+ security, testing accommodations, cost)**
- What do exams look like? (the PrairieLearn LMS)
- Asynchronous exams, student behavior, and randomization
- Increased testing frequency and improved learning outcomes

The Team

Founders

Craig Zilles
Computer Science

Matt West
Mechanical Sci/Eng.

Dave Musselman
Engineering IT

Carleen Sacris
CBTF Coordinator

Nathan Walters
CS (ex-ugrad)

Tim Bretl
Aerospace Eng.

Binglin Chen
CS (grad)

Mariana Silva
CS, MechSE

Geoffrey Herman
Computer Sci.

PrairieLearn Development

Analytics

Quasi-Experimental Studies

Illinois isn't the only one doing these things…

Ron DeMara, University of Central Florida

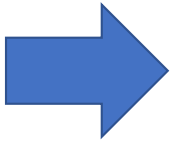
Evaluation and Proficiency Center

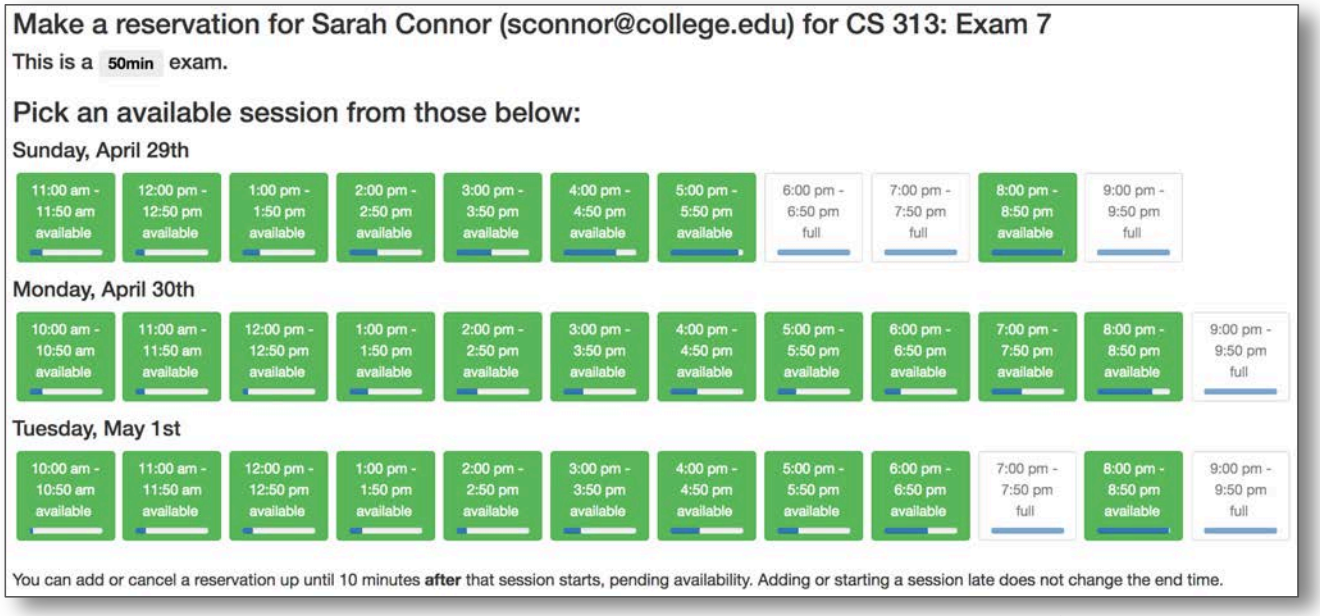
http://www.cecs.ucf.edu/EPCenter

Asynchronous exams: Students self-schedule

Each exam runs for 3-4 days, 10am to 10pm

Exam is available



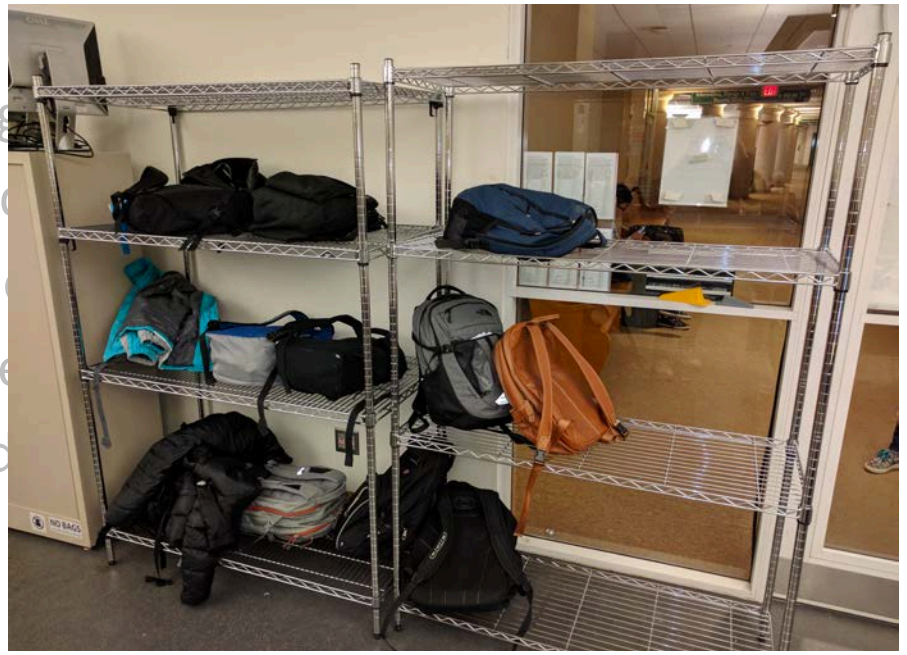


Unlimited rescheduling allowed before the scheduled timeslot

Proctored Exam Experience (1)

1. Students sent a reminder email the morning of their exam
2. 10 min. before exam: photo ID's checked/scanned; seat assigned

gs (except pencil) on racks by entrance



Proctored Exam Experience (2)

1. Students sent a reminder email the morning of their exam
2. 10 min. before exam: photo ID's checked/scanned; seat assigned
3. Students store their belongings (except pencil) on racks by entrance
4. Students pick up scratch paper
5. Student logs into computer, navigates to exam, waits for go signal
6. (Unique) exam generated for student
7. Students answer questions; graded interactively
8. Many exams allow additional attempts (if wrong) for partial credit
9. Proctors enforce time limits; students leave knowing their score

Security is paramount for the CBTF

Zilles, et al. FIE 2018

- Starts with physical security; CBTF is locked when not in use
  - Continually proctored by 2 proctors
- Proctors verify student identity using IDs, photos
- Seating randomized; avoids seating same exam students together
- Networking restricted; fresh local home directory

Biggest source of integrity violations: cell phones & scratch paper

- Hard (legal) boundary between storage and exam parts of the room
- Constantly changing color of scratch paper
- Proctors + students can report cheating; security cameras document

Handling Testing Accommodations

- CBTF handles upwards of 98% of students with accommodations
  - without additional effort from faculty/course staff
  - Students prefer CBTF to our DRES center (more convenient, better hours)
- Most common:
  - Extra time
  - “Reduced distraction environment”
  - Wheel-chair accessible
- Automatically handled in scheduler
  - Students bring letter each semester
  - Student’s record tagged by proctor

Communicate with faculty, Goldilocks-style

- Faculty don’t want to know about issues that CBTF can resolve itself
  - Students missing exams (unless recurring problem)
  - CBTF outages

- Faculty need corroboration from CBTF to handle student complaints
  - Students need to raise the issue before they leave the CBTF
  - Proctor produces an incident report, sent to faculty member
  - Leaving CBTF w/o an incident report is waiving your right to complain

Cost (more details in FIE 2018 paper)

- Biggest expense is staffing (proctors & coordinator)

- Cost per exam < $2. ($190k/year / 100k exams/year)
  - Includes scheduling, proctoring, paper, and grading

- Same order of magnitude as printing multi-page paper exams
- An order of magnitude cheaper than online “proctoring” services

Outline

- Motivation
- The Computer-Based Testing Facility (CBTF)
- Operational details (+ security, testing accommodations, cost)
- **What do exams look like? (the PrairieLearn LMS)**
- Asynchronous exams, student behavior, and randomization
- Increased testing frequency and improved learning outcomes

Q4: Rigid Body Kinematics

Total points: 0/70

0%

Assessment is **open** and you can answer questions.
Available credit: 100% (Instructor override)

For this quiz you can use the Dynamics reference pages.

Question	Best submission	Available points	Awarded points
Question 1	unanswered	10, 9, 7, 5, 2, 1	0 /10
Question 2	unanswered	10, 9, 7, 5, 2, 1	0 /10
Question 3	unanswered	10, 9, 7, 5, 2, 1	0 /10
Question 4	unanswered	10, 9, 7, 5, 2, 1	0 /10
Question 5	unanswered	10, 9, 7, 5, 2, 1	0 /10
Question 6	unanswered	10, 5, 3	0 /10
Question 7	unanswered	10, 9, 7, 5, 2, 1	0 /10

No saved answers to grade

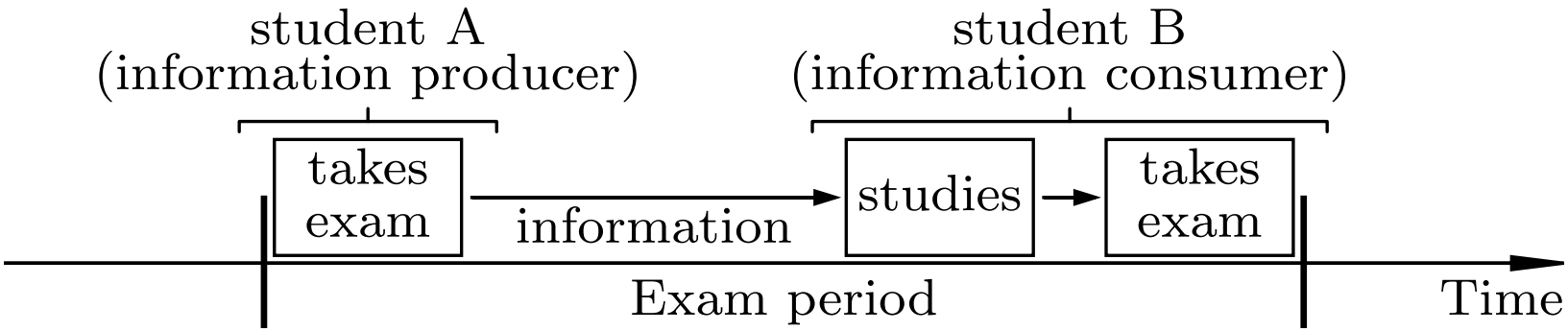
- Submit your answer to each question with the **Save & Grade** or **Save only** buttons on the question page.
- Look at **Best submission** to confirm that each question has been graded. Questions with **Available points** can be attempted again for more points. Attempting questions again will never reduce the points you already have.
- When you are done, please logout and close your browser; there is no need to do anything else. If you have any saved answers when you leave, they will be automatically graded before your final score is computed.

Outline

- Motivation
- The Computer-Based Testing Facility (CBTF)
- Operational details (+ security, testing accommodations, cost)
- What do exams look like? (the PrairieLearn LMS)
- **Asynchronous exams, student behavior, and randomization**
- Increased testing frequency and improved learning outcomes

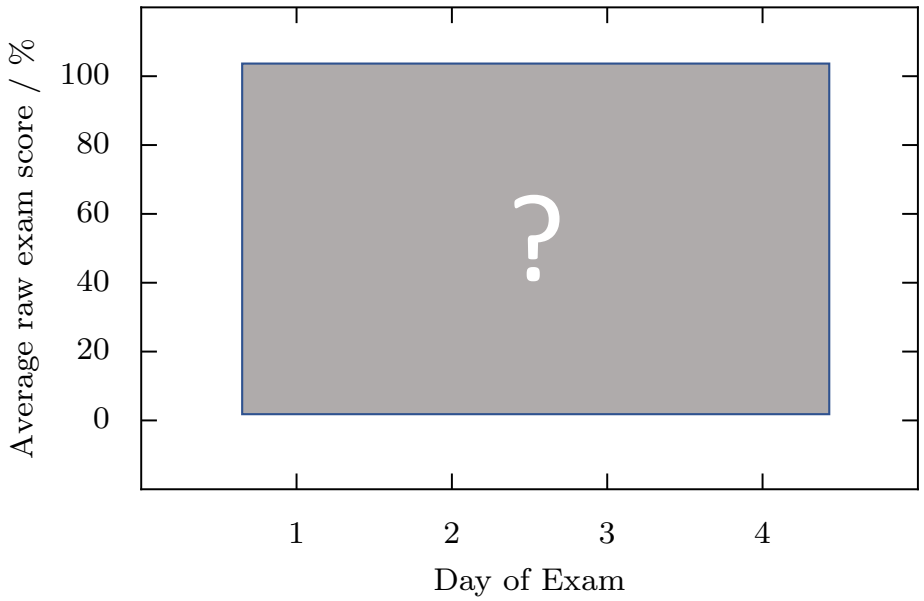
Asynchronous Exams

- Subject to ***collaborative "cheating"***



- Key solution: Exam Randomization

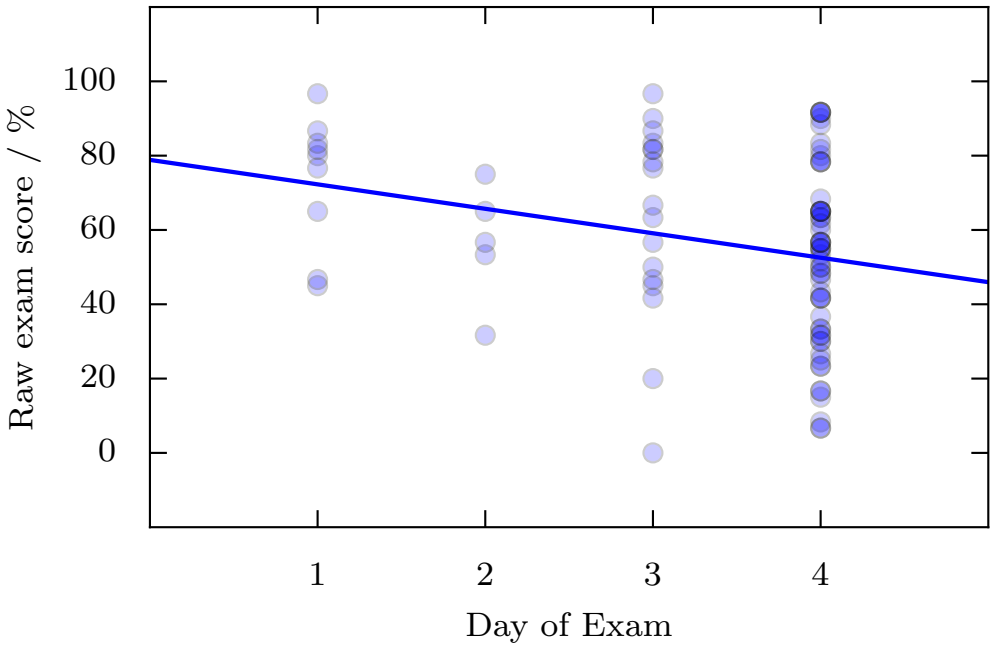
How do students behave?



- When do students choose to take the exam?
- Does performance vary over time? If so, how?

Example exam (clearly demonstrates trends)

Chen, West, Zilles, Learning@Scale 2017

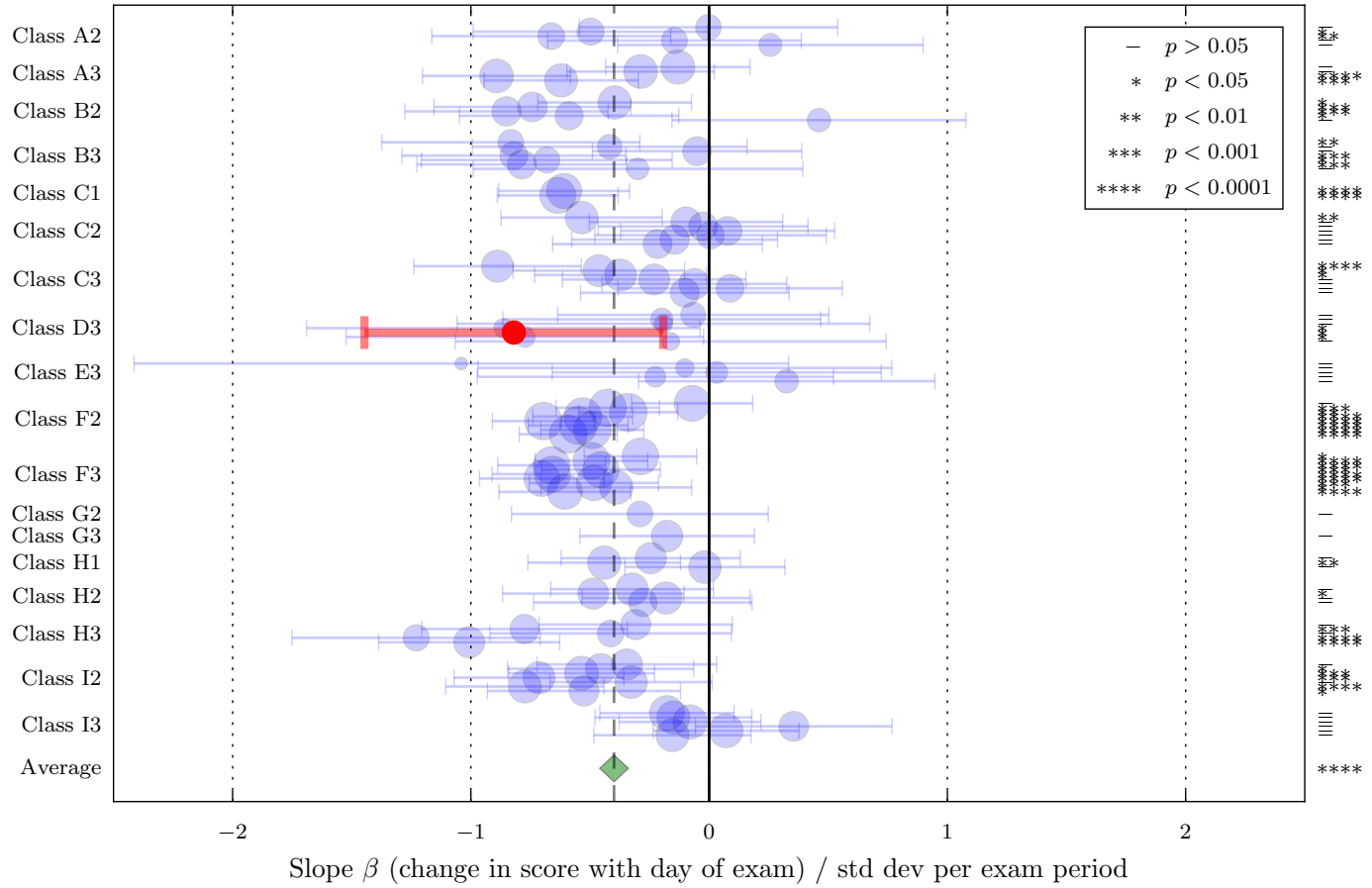


Students prefer slots on the last day

Students' scores drop over time

$\beta = -6.6$ percentage points per day

Meta-analysis: Forest plot

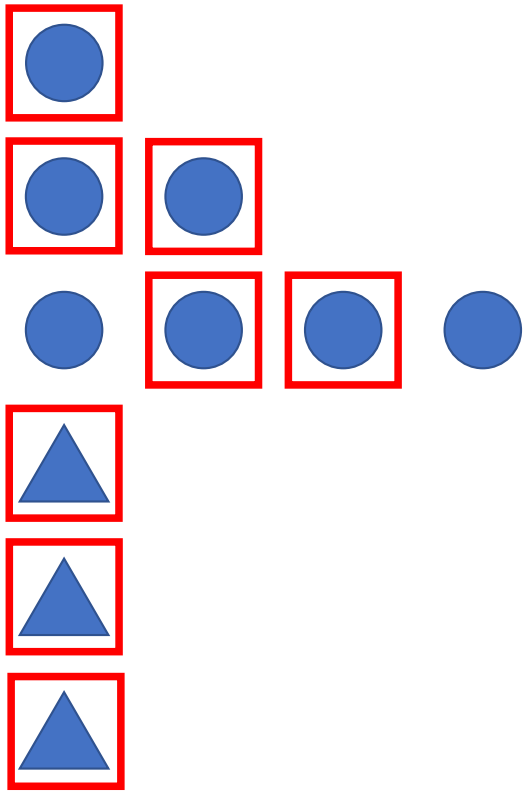


Students' performance decreases over exam period in general, $\beta = -0.4 \ \ (p < 0.0001)$

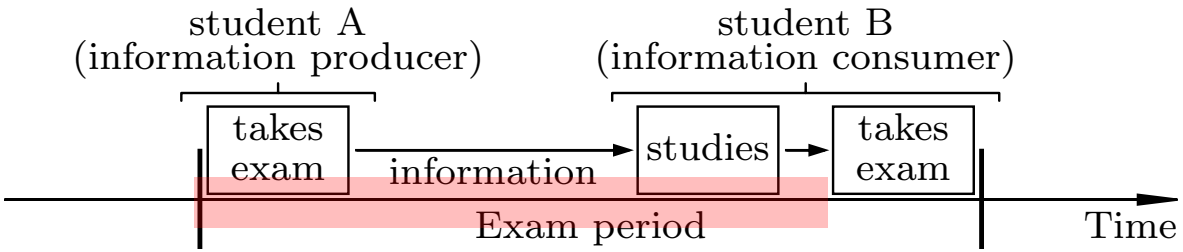
How much randomness is necessary in exams?

Chen, West, Zilles, Learning@Scale 2018

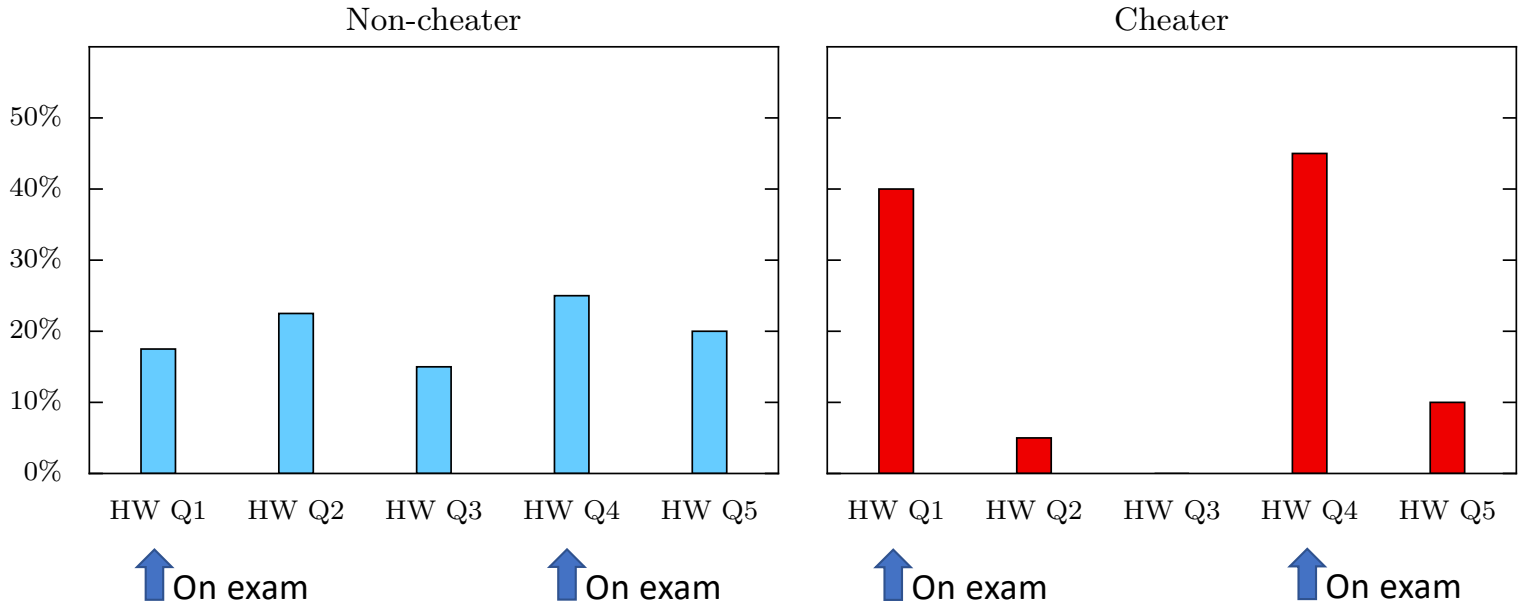
Exam Q1	Accessible pool size 1
Exam Q2	Accessible pool size 2
Exam Q3	Accessible pool size 4
Exam Q4	Inaccessible
Exam Q5	Inaccessible
Exam Q6	Inaccessible



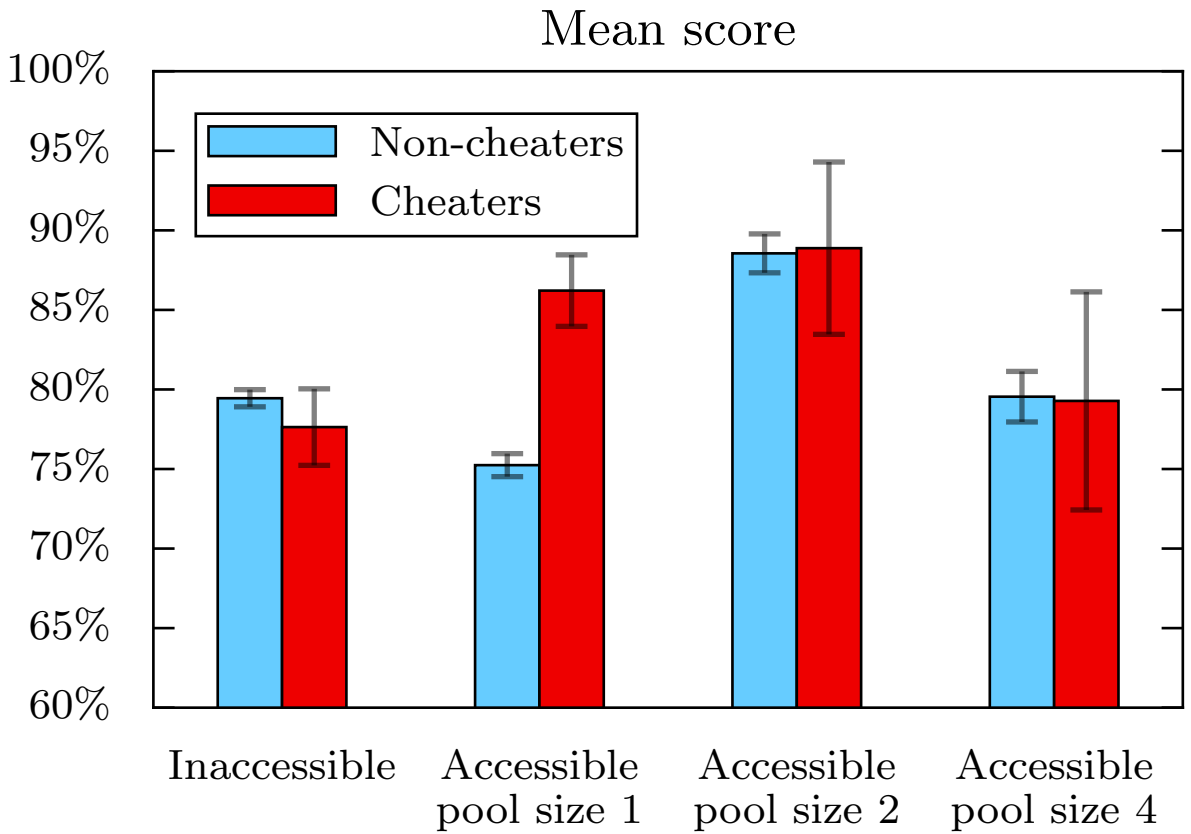
"Cheater" Identification



- Look at distribution of homework attempts during the student's potential cheating period



Mean score diffs: cheaters vs. non-cheaters

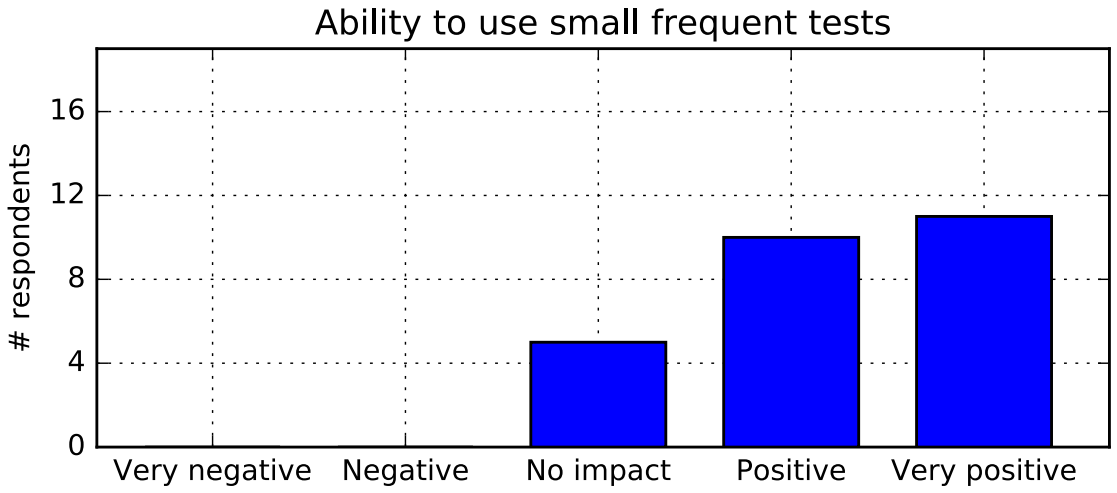


Outline

- Motivation
- The Computer-Based Testing Facility (CBTF)
- Operational details (+ security, testing accommodations, cost)
- What do exams look like? (the PrairieLearn LMS)
- Asynchronous exams, student behavior, and randomization
- **Increased testing frequency and improved learning outcomes**

Frequent Testing

Faculty Survey



"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****."*

Improved learning from frequent testing

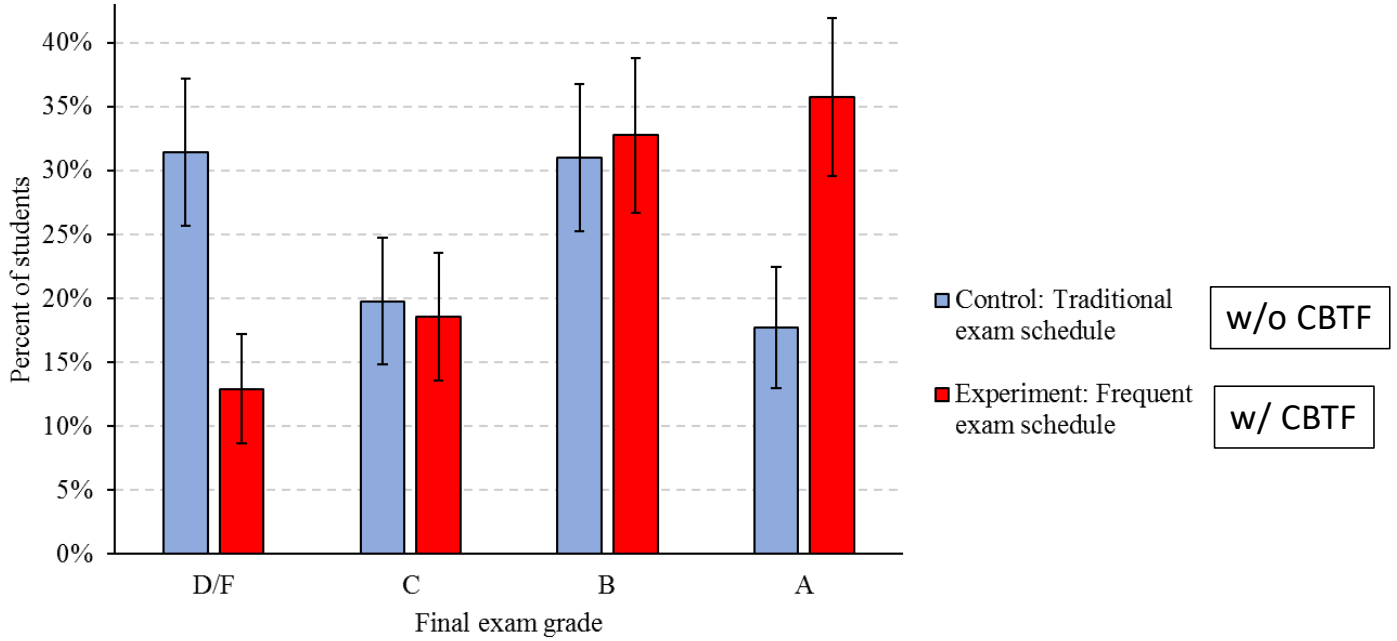
- Two pseudo-experimental studies:
  - Same class, same instructor, different semesters
  - Only change is shifting assessment activities to the CBTF
    - Enabling more frequent, reliable assessment

- Comparing scores on unchanged pencil & paper final exam
  - Better estimate of student knowledge than overall course grade

TAM 251: “Introductory Solid Mechanics”

Morphew, et al. In submission

- Replaced 2 mid-terms with 5 quizzes (+ 5 second chances)

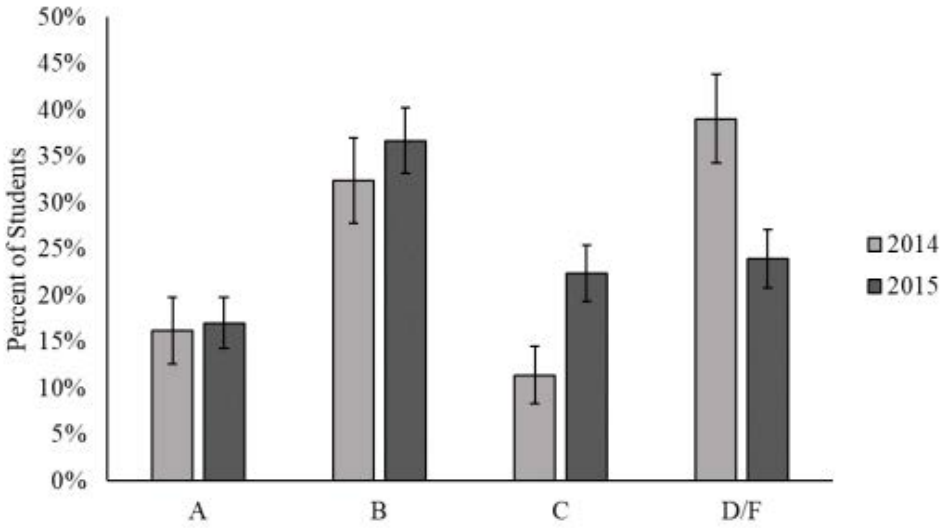


Significant reduction in failing grades, increases in A grades

CS 421: Programming Languages

Nip, et al. SIGCSE 2018

- Converted 2 mid-terms to CBTF
- 4 (of 11) machine problems submitted by re-writing random 1/5 at CBTF

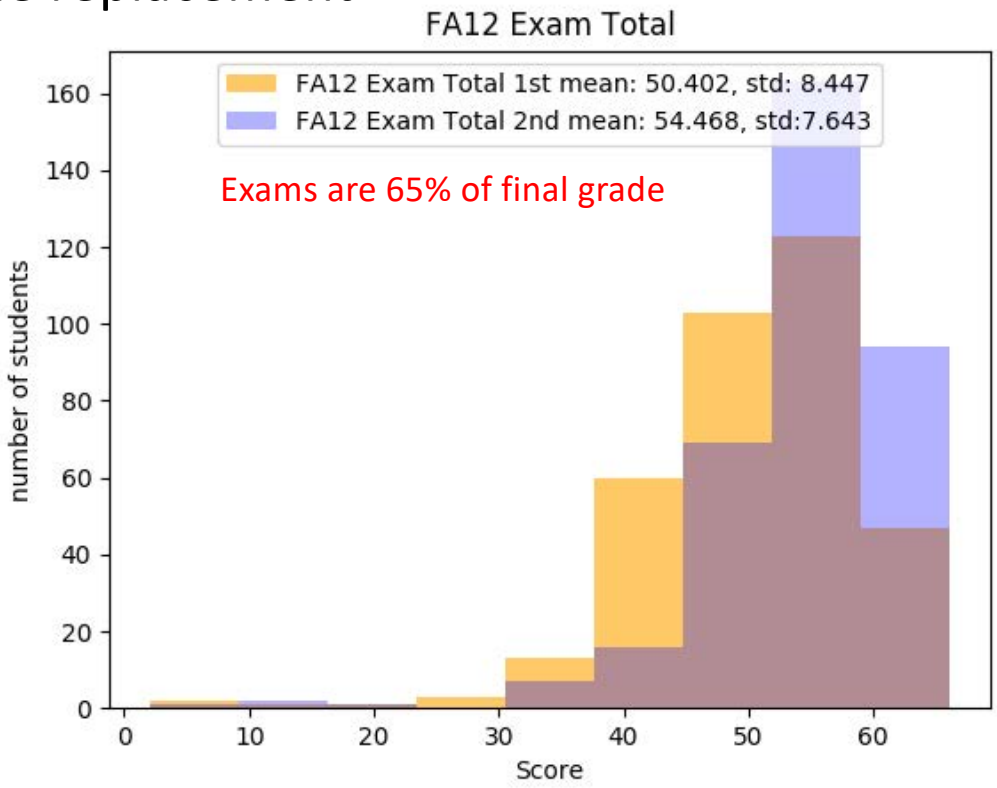


Fall 2014 mean (sd)	Fall 2015 mean (sd)	t-test	Effect size
72.5% (18.6%)	76.8% (15.6%)	p = 0.046	d = 0.25

Significant reduction in failing grades

Second-chance Testing

- Understanding expectations/preparation is hard
- Run exams twice; (partial) grade replacement
  - Improves mean score,
  - tightens the distribution



Summary

- The CBTF is a fixture of our College of Engineering
  - Making testing more efficient and more effective for learning
  - Mature operationally, secure, and running at production scale
- Well liked by faculty (hence its rapid adoption)
  - Does involve a per-course upfront investment in digitizing content
- We're happy to share our experiences and help you adopt a CBTF!
  - Looking for faculty willing to be first adopters on your campus

Keywords: CBTF, PrairieLearn, Zilles