Striving for Uniformity and Efficiency in Math Assessment

Matt Cecil & Rebecca Noonan-Heale

Department of Mathematics

University of Utah, USA

April 11, 2019

State of Assessment in Math

- ELOs are assessed on the final exams of most 1000-level courses and some 2000-level courses.
- In Fall 2018, there were almost 5500 students enrolled in our 1000-level courses distributed across 18 courses and 97 sections.
- There is variance among these courses in terms of instructors, course structure, etc.
- We're interested in improving
  - Efficiency
  - Uniformity (across courses but also within individual courses)

Course	# Stud/Sem	# Sect.	Class sz.	Instructors	Com. Fin.	Amt. Assessed	Graded by
1010	500	9	$40-100$	adj	yes	all	inst, und
1030	500	9	$40-90$	cl, grad, adj	yes	half	inst, und
1050	800	7	$20-250$	cl,grad	yes		inst, und
1060	400	7	$20-250$	cl,grad	yes		inst, und
1070	275	5	60	grad, adj	no	3 q's	inst
1080	225	4	50	grad, adj	yes	all	inst, und
1090	300	6	$15-75$	grad	yes	all	inst, und
1210	700	9	$60-140$	grad, cl, tt	no	4 q's	inst
1220	500	6	$30-180$	grad, cl, tt	no	4 q's	inst
1310	250	5	$30-120$	grad	no	$4-6$ q's	inst
1320	100	2	50	grad	no	$4-6$ q's	inst
2210	400	6	$50-100$	grad, cl, tt	no	4 q's	inst
2250	350	5	$40-120$	grad	no	$4-6$ q's	inst

Key: adj=adjuncts, grad=grad students, cl=career-line, tt=tenure-line, inst=instructors, und=undergraduate graders

Calculus & Engineering Calculus Assessment Procedure

- Problems are determined to reflect on certain ELOs, split between ‘computational’ problems and ‘theoretical’ problems.

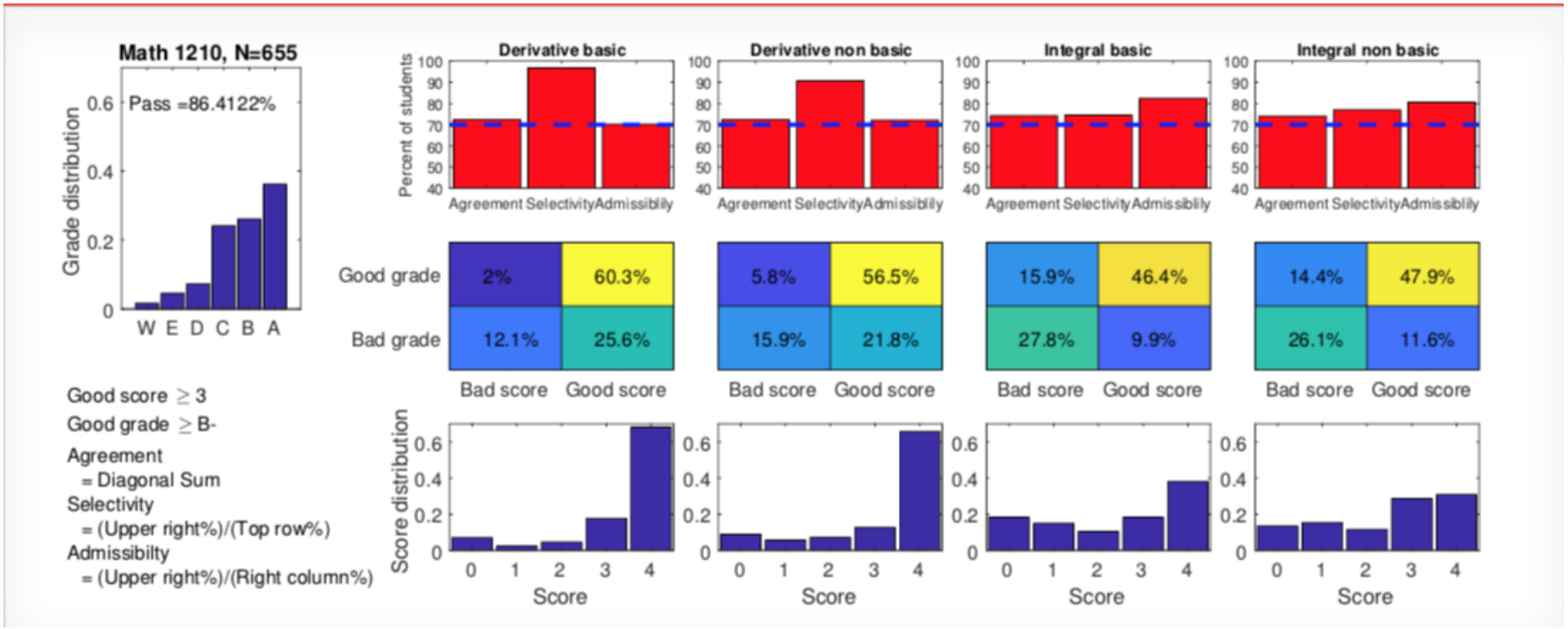
- The questions are discussed with the instructors early in the semester so that everyone is aware of their responsibilities and the problems content. The instructors can make the rest of their final exams as they see fit.

- The instructors can chose to grade the problems out of however many points they chose, however they need to assign the students assessment scores on the problems from 0-4. There is discussion before grading so that everyone has a rubric for what constitutes the various scores.

- Instructors grade the exams and report scores in spreadsheets which are then combined with other student data (grade, major, year, etc).

Calculus & Engineering Calculus Score Reporting

- Classify scores on a given problem as *good* (3 or 4) or *bad* (0-2).
- Classify student grades as *good* (B- and above) or *bad* (C+ and below).
- For each assessment problem we get a 4×4 grid.



Calculus & Engineering Calculus Score Reporting

Analysis is used to identify course content

- that students have particular difficulty with
- that instructors should emphasize (and de-emphasize) in lecture
- that is a good candidate for active learning (lab) exercises.

In theory, we could track our progress using subsequent assessments, however, scores are not so stable from semester to semester so this will require more data and more time.

gradescope

Online Grading Tool

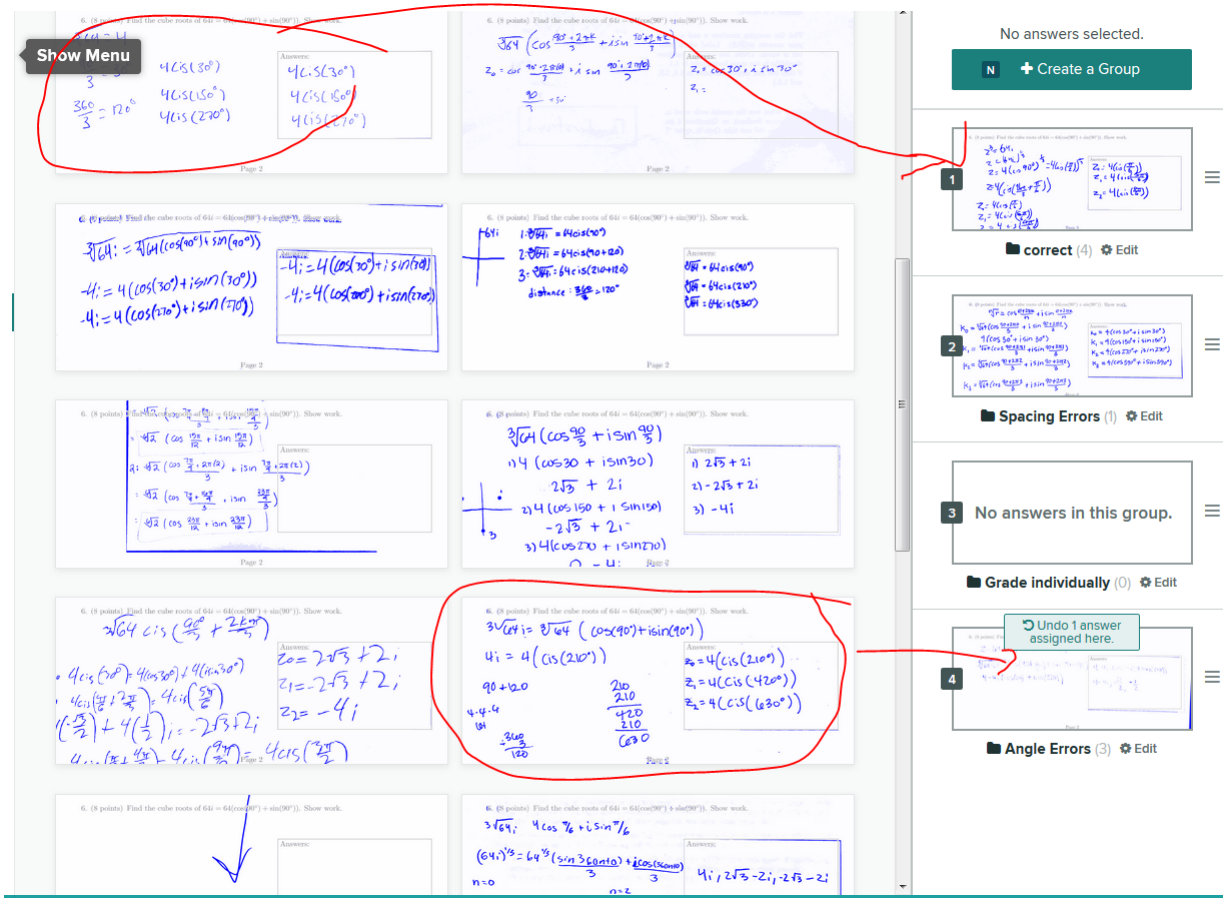
- Individual Instructors in Math have been using Free Trial
  - Precalculus and some online courses, 2017-2019
  - Some Calculus and Engineering Calculus Sections, 2018-2019
  - Free-trial runs out for most instructors this summer
- Computer Science piloting 2018-2019

How Gradescope Works

1. Sync Gradescope course with Canvas
2. Upload assignment to grade (done by student or instructor); Assignments matched with AI and manually to students
3. Sort answers
4. Grade and Give Feedback to Batches of Answers
5. Post Grades to Canvas
6. Students view feedback through link to Gradescope (complies with FERPA)
7. Download feedback-level or problem-level data about assessment

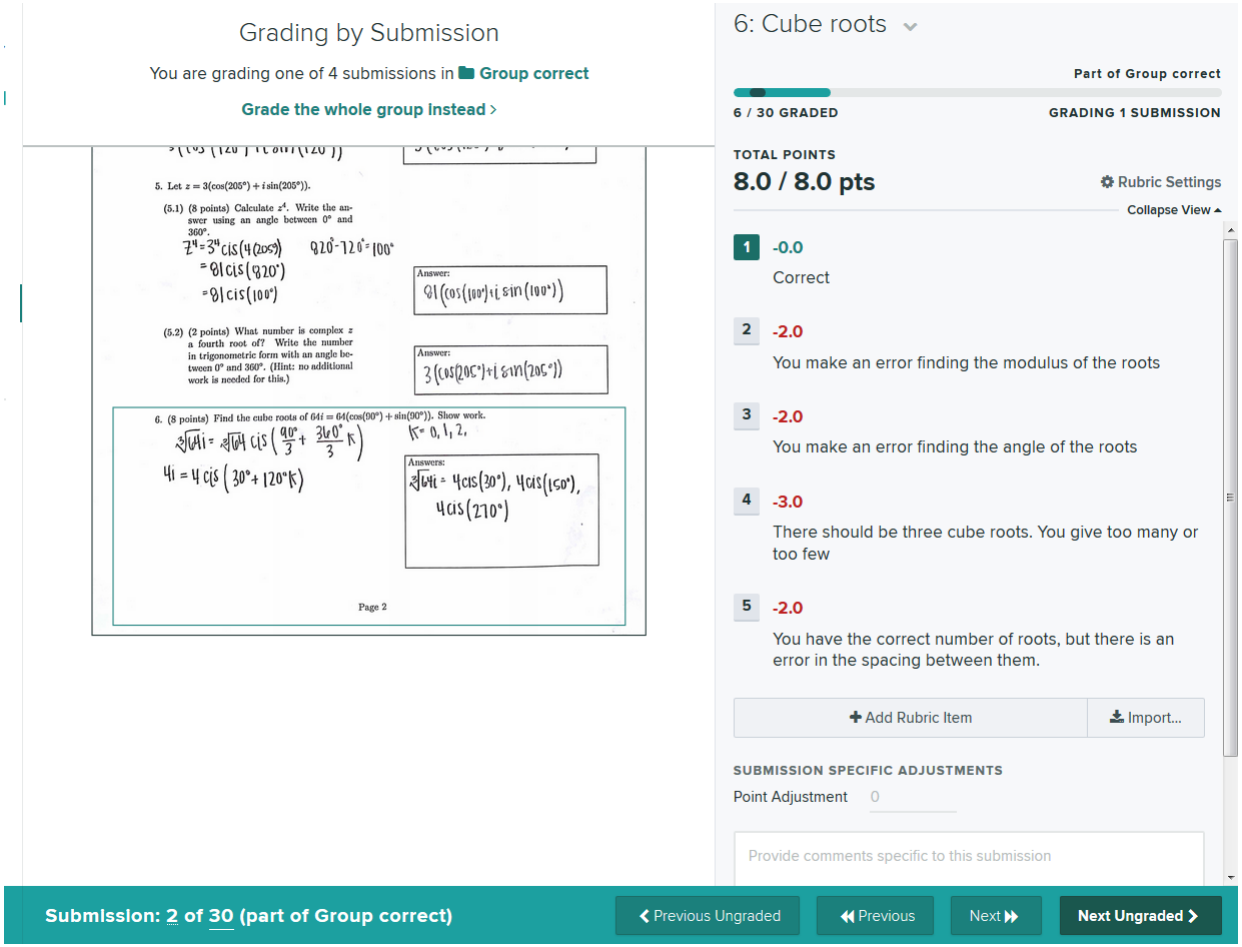
Sorting

- Math expressions, one-word answers, multiple choice problems graded with Artificial Intelligence (Check-manually)
- Sort Other types of answers Manually



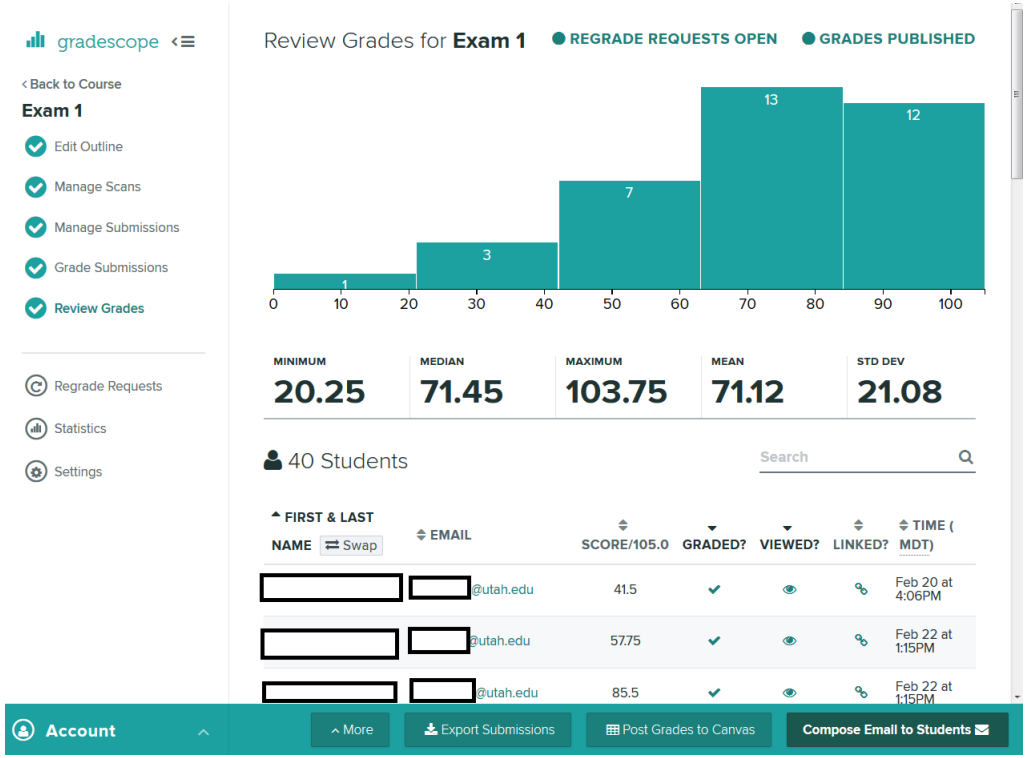
Grade Batches of Answers

- Grade sorted batches of answers
- Write feedback and assign to batches or individual students
- Easily adjust point values retroactively
- Easily compare students getting same or different feedback



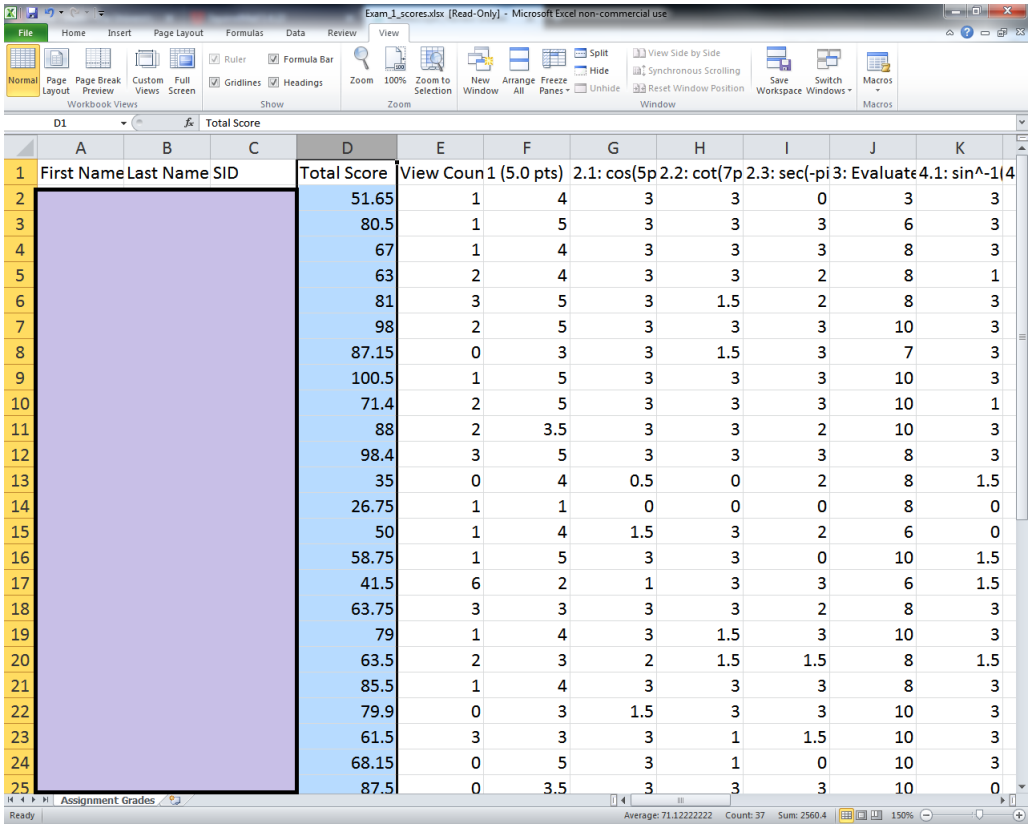
Reviewing Grades

- Post grades to Canvas
- Send students e-mail with link for them to look at problem scores and feedback
- Students can request regrading of problems
- Contrast this to passing back assignments in class/lab/etc.



Download Information about Student Results

- Problem-level information



	A	B	C	D	E	F	G	H	I	J	K
1	First Name	Last Name	SID	Total Score	View Coun	1 (5.0 pts)	2.1: cos(5p	2.2: cot(7p	2.3: sec(-pi	3: Evaluate	4.1: sin^-1(4
2				51.65	1	4	3	3	0	3	3
3				80.5	1	5	3	3	3	6	3
4				67	1	4	3	3	3	8	3
5				63	2	4	3	3	2	8	1
6				81	3	5	3	1.5	2	8	3
7				98	2	5	3	3	3	10	3
8				87.15	0	3	3	1.5	3	7	3
9				100.5	1	5	3	3	3	10	3
10				71.4	2	5	3	3	3	10	1
11				88	2	3.5	3	3	2	10	3
12				98.4	3	5	3	3	3	8	3
13				35	0	4	0.5	0	2	8	1.5
14				26.75	1	1	0	0	0	8	0
15				50	1	4	1.5	3	2	6	0
16				58.75	1	5	3	3	0	10	1.5
17				41.5	6	2	1	3	3	6	1.5
18				63.75	3	3	3	3	2	8	3
19				79	1	4	3	1.5	3	10	3
20				63.5	2	3	2	1.5	1.5	8	1.5
21				85.5	1	4	3	3	3	8	3
22				79.9	0	3	1.5	3	3	10	3
23				61.5	3	3	3	1	1.5	10	3
24				68.15	0	5	3	1	0	10	3
25				87.5	0	3.5	3	3	3	10	0

Advantages in Assessing Assignments: Uniformity in Grading

- The exams and the graders don't need to be in the same place
  - Large classes traditionally have grading sessions with 5-25 people, need to find space, coordinate schedules, communicate everything, often give up evenings or weekends
- Multiple people can grade the same exam at the same time
- Easier for one person to grade a single problem OR review other grader's grading and make changes if needed
- Students also help spot mistakes in grading
- Mistakes in Grading are easy to address

Advantages in Assessing Assignments: Collecting Data

- Current System: someone has to prepare database and enter 4-30 scores per student
  - 500 students $\times$ 2 min/student = 1000 min OR almost 17 hours
  - Currently 5500 students in 1000-level classes each semester
  - If instructors are responsible for data, coordinators often need to chase it down, and sometimes do not get it at all
  - If undergrad assistants enter data, this delays getting exams to instructors
- Gradescope eliminates data-entry step AND data entry errors
- Data is automatically adjusted if problems are regraded
  (compare to either not fixing data-set or needing to find individual data points)
- Allows collecting of data at mass scale

STEM Currently Lacks Assessment Tools

- Compare to humanities: students submit projects/papers in Canvas; aggregated into an e-portfolio (SLCC) that shows student work over time (Can be accessed by students and by institution)
- On humanities side, Turn-it-in Feedback Studio is software comparable to Gradescope; U has license for these
- Nothing comparable in STEM fields
- Biggest road-block in gathering assessment data is transforming stacks of paper exams into useful, digital information

Big Picture

- The University requires that we perform assessment on student performance; departments decide how to implement this.
- Assessing without a tool limits what we do and ask to the bare minimum
- Assessing with a tool like Gradescope allows us to ask and answer relevant questions
- Applies to MANY areas of University
- Gradescope free-trial runs out in Summer 2019. We need a tool like it!
- Investigated Gradescope Math Department License: $10,500/ year (not doable - this is about total software budget for Maple, Mathematica, etc.)
- Investigated Gradescope University license: $104,000 based on enrollment 27,000 (and we think this could be negotiated)