



Peer-Review Easily with “Confidence”: A Look at Replication and Randomization as tools for Statistically-Significant Assessment via Platypus

Dr. Surya P. N. Singh

spns@uq.edu.au – <http://robotics.uq.edu.au/spns>

(Part of The Platypus Team – S.Singh, A. Dekker, K. Seiler, H. Kurniawati, M. Kearney, I. Jahn, *et al.*)

ITaLI Teaching Masterclass Series

September 1, 2017

<http://robotics.uq.edu.au>



PLATYPUS

Platypus
Question-based peer review

Replication & randomization of reviewing supporting faster, richer, and more diverse feedback

[Demo](#) [UQ Version](#) [Github](#)

[Get Started](#)

Peer Review for the Digital Age
Platypus provides a workflow optimised for learning

- Submit**
Using the platypus interface students submit answers to an assignment question
- Distribute Broadly**
The Platypus system will automatically distribute anonymized questions randomly to students
- Review & Reflect**
Students and tutors mark the assigned answers, providing feedback and insight
- Collate & Analyse**
Use the review results to provide insight into class performance and understanding

PLATYPUS

How It Works Features What You Get

Features and Tools

Platypus provides focused features to help enable learning

- Random Shuffling**
Anonymized question based shuffle ensures accurate marking and effective feedback
- Powerful Analytics**
Automatically collect a wealth of data. Quickly assess the state of your class and keep your students on track.
- Peer Feedback**
Students receive wider feedback, quicker than ever before, helping them take control of their learning
- Latex Support**
Built in support for [LaTeX](#) provides students with the tools to answer any question their way
- Any Size**
With support for per question allocation, platypus is suitable for classes of any size

Richer Feedback

Provide students with richer feedback, faster.

By integrating feedback from multiple perspectives, including course staff and multiple peers, as well as reducing turnaround times, Platypus provides students [detailed feedback](#) and the support required to fully develop content understanding.

Question

The expression $\frac{1}{x^2}$ is the same as x^{-2} . It is true that the derivative of x^{-2} is $-2x^{-3}$. Therefore, the derivative of $\frac{1}{x^2}$ is $-2x^{-3}$. This can be written as $-\frac{2}{x^3}$. Therefore, the derivative of $\frac{1}{x^2}$ is $-\frac{2}{x^3}$.

Mark: 85%

Question

Let $f(x) = x^2 + 3x - 5$. Find $f'(x)$. The derivative of x^2 is $2x$. The derivative of $3x$ is 3 . The derivative of -5 is 0 . Therefore, the derivative of $f(x)$ is $2x + 3$.

Mark: 95%

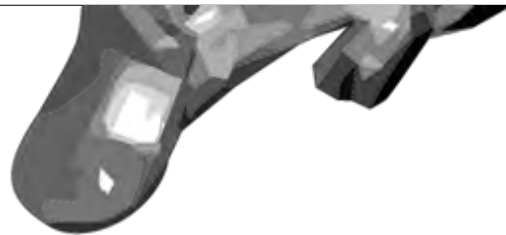
Question

Let $f(x) = x^2 + 3x - 5$. Find $f'(x)$. The derivative of x^2 is $2x$. The derivative of $3x$ is 3 . The derivative of -5 is 0 . Therefore, the derivative of $f(x)$ is $2x + 3$.

Mark: 95%

What is Platypus?

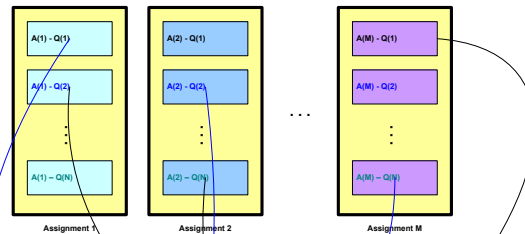
*Many Things
Put Together!*



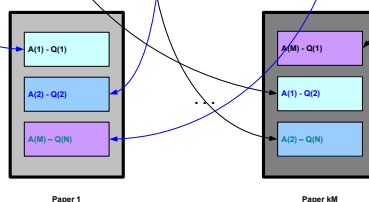
- *(Yet another) Electronic Submission System*
- *A Question-based Online Peer Assessment Tool*
- *An “Anonymous” Review + Student Progress Tool*
- *A Tool With a Nuanced (Data-centric) View of Assessment*

Question-Based Peer Review: Workflow Model

I. Collect Assignments (one per student)



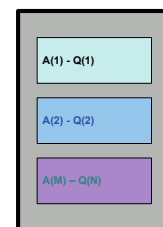
II. Randomly Shuffle Questions between assignments to create kM "Papers"
(Where k is the peer review factor, or the number of papers a student needs to review, eg 3)



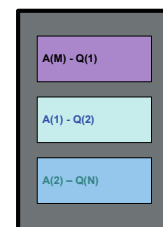
III. Each student grades k papers (may be assigned randomly or via an alternative metric)

Why a Question-Based Shuffle?

- Replication & Randomization
- Assessment is an Observer
 \therefore We would like $N(0, \sigma)$
- Less Numbers Needed for Analysis
 - A Project Class is not Super Large
- Summative Assessment by parts
- Faster feedback

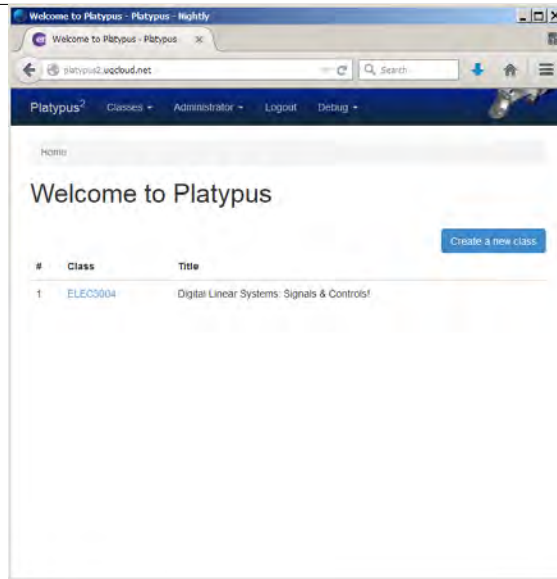


Paper 1

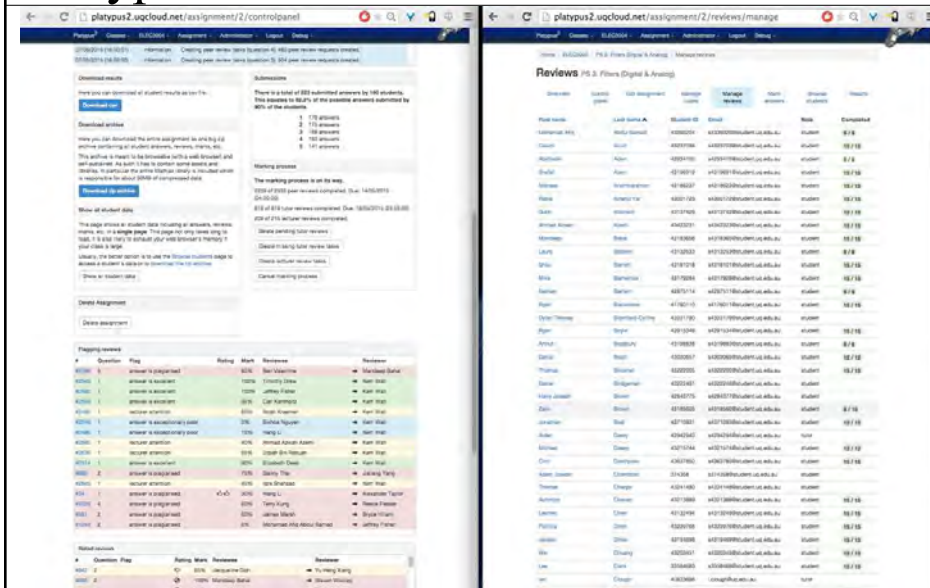


Paper kM

Platypus



Platypus



It's Open...

```

5  * This file is an application-wide controller file. You can put all
6  * application-wide controller-related methods here.
7
8  * PHP 5
9
10 * CakePHP(tm) : Rapid Development Framework (http://cakephp.org)
11 * Copyright 2005-2012, Cake Software Foundation, Inc. (http://cakefoundation.org)
12 *
13 * Licensed under the MIT License
14 * Redistributions of files must retain the above copyright notice.
15 *
16 * @copyright Copyright 2005-2012, Cake Software Foundation, Inc. (http://cakefoundation.org)
17 * @link http://cakephp.org CakePHP(tm) Project
18 * @package app.Controller
19 * @since CakePHP(tm) v 0.2.9
20 * @license MIT License (http://www.opensource.org/licenses/mit-license.php)
21 */
22
23 App::uses('Controller', 'Controller');
24
25 /**
26  * Application Controller
27  *
28  * Add your application-wide methods in the class below, your controllers
29  * will inherit them.
30  *
31  * @package app.Controller
32  * @link http://book.cakephp.org/2.0/en/controllers.html#the-app-controller
33  */
34 class AppController extends Controller {
35
36     public $baseUrl = '/platypus';
37     public $pageTitle = 'NO PAGE TITLE';
38     public $navState = 'notloggedin';
39     public $userdetails = array();
40     public $loggedin = false;
41     public $isadmin = false;
42     public $memberaccesslevels = array('contributor' => 'Contributor', 'administrator' => 'Administrator');
43     public $breadcrumbs = array();
44     public $fullurl = 'http://robotics.ites.uq.edu.au/platypus';
45 }

```

UQ Robotics

9

SQL + Laravel

The screenshot shows a MySQL 5.5.19 database interface with a table named 'answers'. The table has columns: id, user_id, question_id, content, times_answered, and estimated_grade. The data is as follows:

id	user_id	question_id	content	times_answered	estimated_grade
4033	142	41	<p>a</p>From source cod...	4	20
4034	142	42	<p>The output of both transform...	1	10
4035	142	43	<p>a</p>A valid rotation...	7	40
4036	34	41	<p>(a)</p><p>a</p><p>its the same</p><p></p>...	2	10
4050	225	43	<p>R2 is not a valid rotation matr...	1	40
4051	85	41	<p>Take roll to be about the x-a...	5	20
4052	85	42	<p>Using aerospace convention (...)	8	10
4053	85	43	<p>Using aerospace convention (...)	2	40
4054	211	41	<p>a) rotate matrix by the followi...	3	20
4055	211	42	<p>The first transformation matr...	1	10
4056	211	43	<p>a. Valid rotation matrices hav...	3	40
4057	217	41	<p><p><p>Question 1a.</p>...	2	20
4058	217	42	<p>The output of both those tran...	1	10
4059	217	43	<p>a. All but R2 are valid becaus...	2	40
4060	215	41	<p>(a) (a) <img src="http://latex.co...	2	40

UQ Robotics

10

SQL: Sometimes it's handy

SELECT users.first_name, users.last_name, users.student_id, users.id FROM ((users INNER JOIN answers ON users.id = answers.user_id) INNER JOIN text_blocks ON text_blocks.id = answers.answer_text_id) WHERE text_blocks.text LIKE '%fbcdn.net%'

first_name	last_name	student_id	id
Phoen	hadrigh	42469712	2092
Liam	Williams	43324205	818
Wei Feng	Tye	43318140	803
Wei Feng	Tye	43318140	803
Wei Feng	Tye	43318140	803
Wei Feng	Tye	43318140	803
Housten	Stater	43928361	5299
Housten	Stater	43928361	5299
Housten	Stater	43928361	5299
Housten	Stater	43928361	5299
Housten	Stater	43928361	5299
Wei Feng	Tye	43318140	803
Wei Feng	Tye	43318140	803
Housten	Stater	43928361	5299
Housten	Stater	43928361	5299
Housten	Stater	43928361	5299
Housten	Stater	43928361	5299

Benefit: Focus on Learning Outcomes



Peer Feedback Research

Tuned Models of Peer Assessment in MOOCs

Peer Assessment Learning Sessions (PALS): an innovative feedback technique for large engineering classes

Division of Civil Engineering

This paper reports the results of a study in Civil Engineering that facilitates a student within a "traditional" out against worked reviewed by the lecturer provide very clear a below required class. Students obtain feedback within a their review of class of anonymous survey assessment for large quality formative feedback.

Keywords: Peer assessment, Learning, Self-reflection

1. Introduction

As Ramsten (1988) states

Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments

Justin Kruger and David Dunning
Cornell University

People tend to hold overly favorable views of their abilities in many social and intellectual domains. The authors suggest that this overestimation occurs, in part, because people who are unskilled in these domains suffer a dual burden: Not only do these people reach erroneous conclusions and make unfortunate choices, but their incompetence robs them of the metacognitive ability to realize it. Across 4 studies, the authors found that participants scoring in the bottom quartile on tests of grammar, logic, and logic grossly overestimated their test performance and ability. Although their test scores put them in the 12th percentile, they estimated themselves to be in the 62nd. Several analyses linked this miscalibration to deficits in metacognitive skill, or the capacity to distinguish accuracy from error. Paradoxically, improving the skills of participants, and thus increasing their metacognitive competence, helped them recognize the limitations of their abilities.

It is one of the essential features of such incompetence that the person so afflicted is incapable of knowing that he is incompetent. To have such knowledge would already be to remedy a good portion of the offense. (Miller, 1993, p. 4)

In 1995, McArthur Wheeler walked into two Pittsburgh banks and robbed them in broad daylight, with no visible attempt at disguise. He was arrested later that night, less than an hour after videotapes of him taken from surveillance cameras were broadcast on the 11 o'clock news. When police later showed him the surveillance tapes, Mr. Wheeler stared in incredulity. "But I wore the juice," he mumbled. Apparently, Mr. Wheeler was under the impression that rubbing one's face with lemon juice rendered it invisible to videotape cameras (Fraccon, 1996).

We bring up the unfortunate affairs of Mr. Wheeler to make three points. The first two are noncontroversial. First, in many domains in life, success and satisfaction depend on knowledge, wisdom, or savvy in knowing which rules to follow and which strategies to pursue. This is true not only for committing crimes, but also for many tasks in the social and intellectual domains, such

as promoting effective leadership, raising children, constructing a solid logical argument, or designing a rigorous psychological study. Second, people differ widely in the knowledge and strategies they apply in these domains (Dunning, Meyerowitz, & Holzberg, 1989; Dunning, Petre, & Story, 1991; Story & Dunning, 1988), with varying levels of success. Some of the knowledge and theories that people apply to their actions are sound and meet with favorable results. Others, like the lemon juice hypothesis of McArthur Wheeler, are imperfect at best and wrong-headed, incompetent, or dysfunctional at worst.

Perhaps more controversial is the third point, the one that is the focus of this article. We argue that when people are incompetent in the strategies they adopt to achieve success and satisfaction, they suffer a dual burden: Not only do they reach erroneous conclusions and make unfortunate choices, but their incompetence robs them of the ability to realize it. Instead, like Mr. Wheeler, they are left with the mistaken impression that they are doing just fine. As Miller (1993) perceptively observed in the quote that opens this article, and as Charles Darwin (1871) sagely noted over a century ago, "ignorance more frequently begets confidence than does knowledge" (p. 3).

In essence, we argue that the skills that engender competence in a particular domain are often the very same skills necessary to evaluate competence in that domain—one's own or anyone else's.

Justin Kruger and David Dunning, Department of Psychology, Cornell University
We thank Bruce Brown, Mark Scholander, and Boris Yermakov for their

Question-Based Peer Review: Where it Works & Doesn't

Student review

A very intuitive approach towards solving the issue. Some points you may want to consider though:

- * Other towers may obstruct the target tower (in terms of straight line motion)
- * Calculating a number of points would require multiple rotational instructions which, as the Lego motors are not the most accurate, would potentially compound a larger offset from the desired target and return position.

Your alternative solution is also very well thought out. In terms of calculating the effective velocity, a scaling of arm distance and rotational speed (motor speed settings) would need to be accounted for. In addition, this may require a lot of trial-and-error calibrations to acquire an appropriate scaling for effective rotation; motor speed; time taken (delay settings post-command).

Mark: **100%**

Reviewers: Terry Kung
Email: t4294475@student.uq.edu.au
Student ID: 42844754

review #11310

Student review

~|

Mark: **100%**

Reviewers: Felix Ng
Email: s4323950@student.uq.edu.au
Student ID: 43239508

review #11317

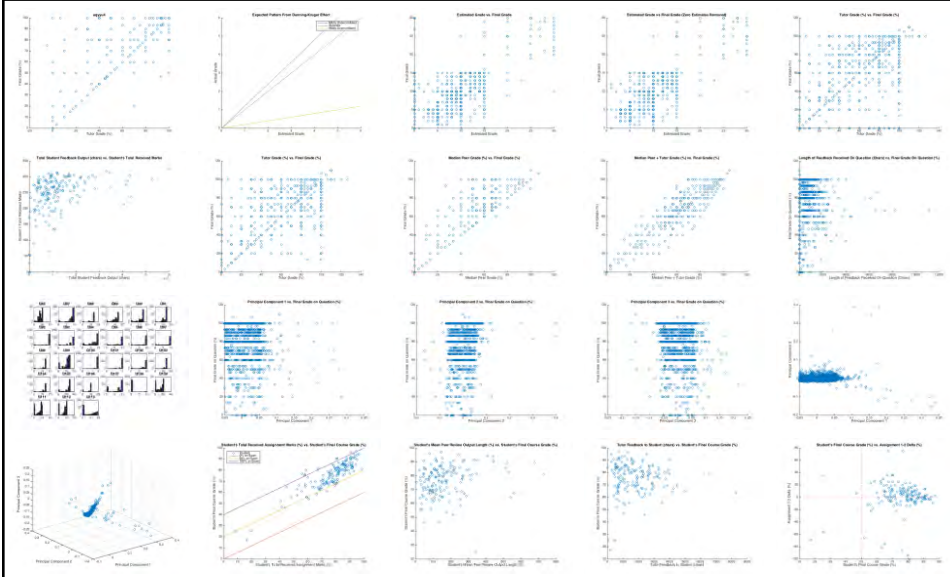
Student review

A good solution for resolving this issue in the future.

Mark: **100%**

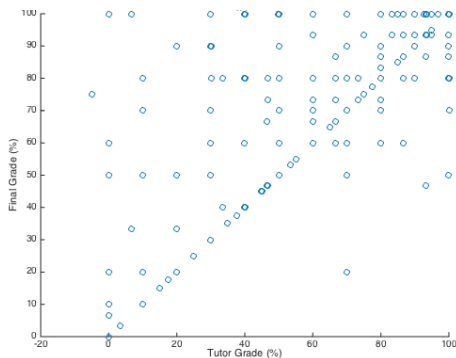
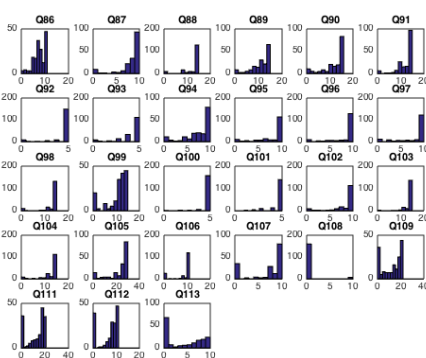
Reviewers: Zachary Morris
Email: s4392987@student.uq.edu.au
Student ID: 43929875

“Statistics”: Tools to “Assess” Learning?



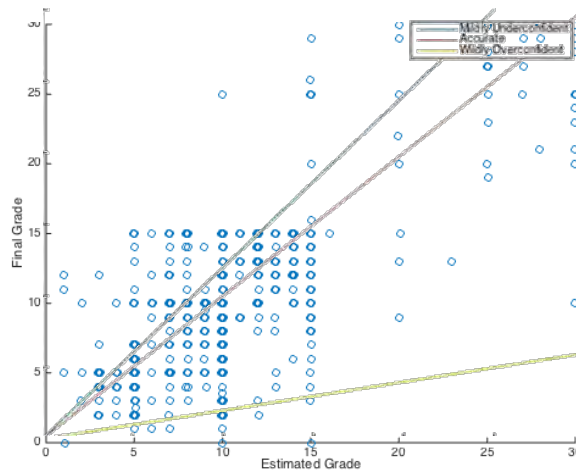
Basic Counts...

- Scores Histograms...
- Grade Correlations...

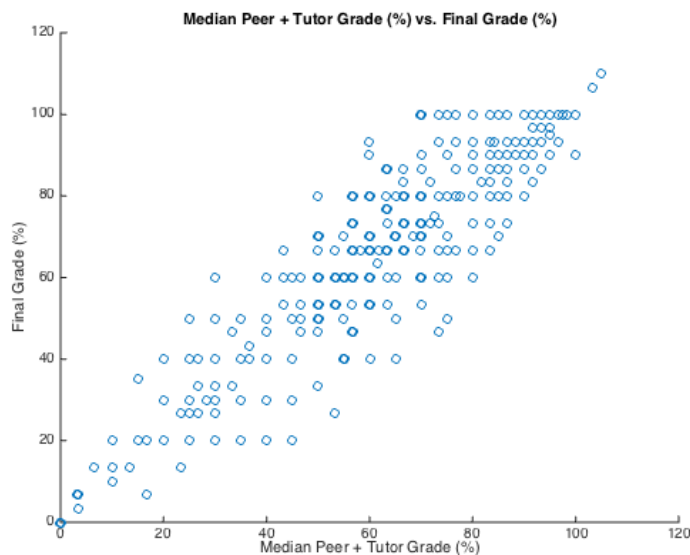


Interesting Finds: Dunning-Kruger Effect?

- Correlation Between Students' Actual & Estimated & Final Grades

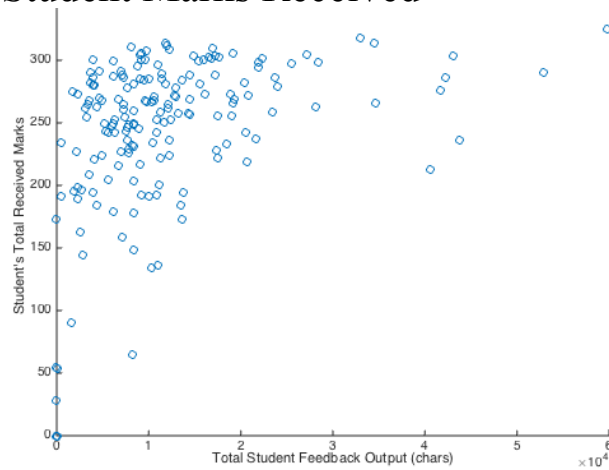


Trustworthiness of Peer & Tutor Marking



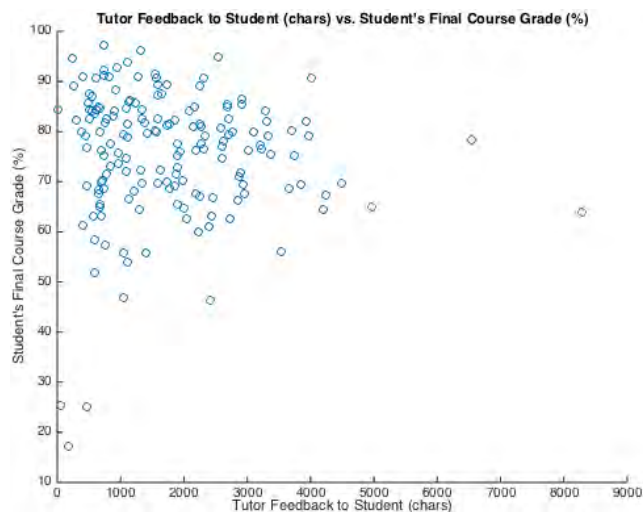
Longer answers better?

- Total Student Feedback Output Characters vs. Total Student Marks Received

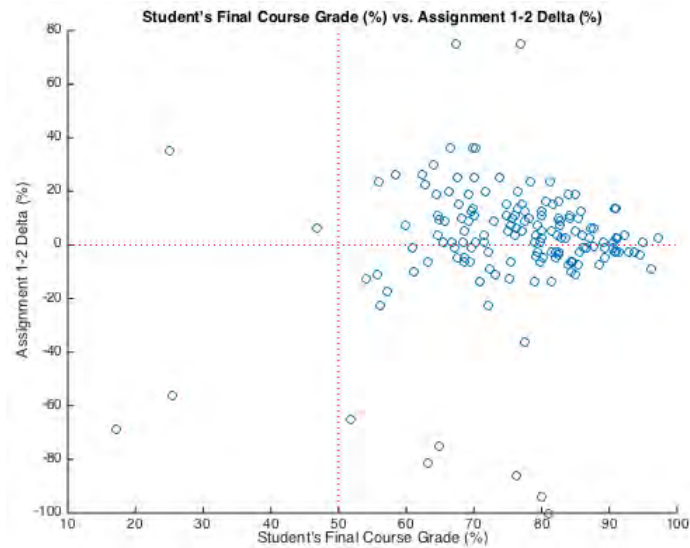


Longer feedback better?

- Tutor Feedback Length and Final Course Grade



Positive Δ : Positive Outcome?

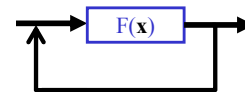
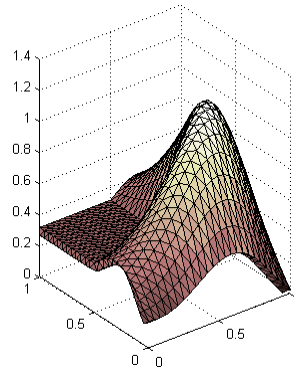


Worldwide Access



Future Work

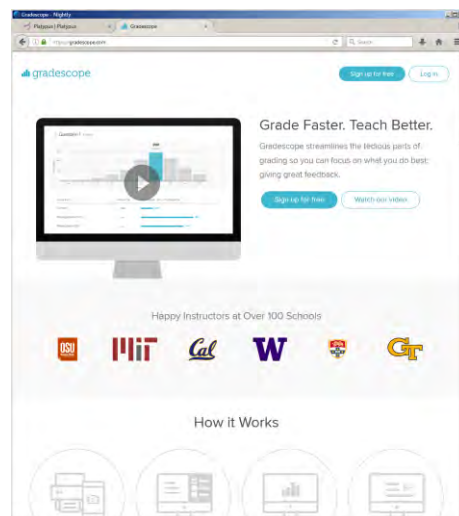
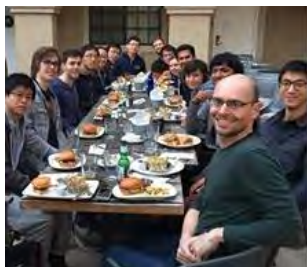
- Richer Inputs
 - [Matlab] Figures
 - Simple Diagrams/Sketches
 - ➔ Rich media in general
- Alternate Question Types
 - Lab Reports, etc.
 - ➔ User-Defined?
- Team Submissions/Review
- Visualization Tools



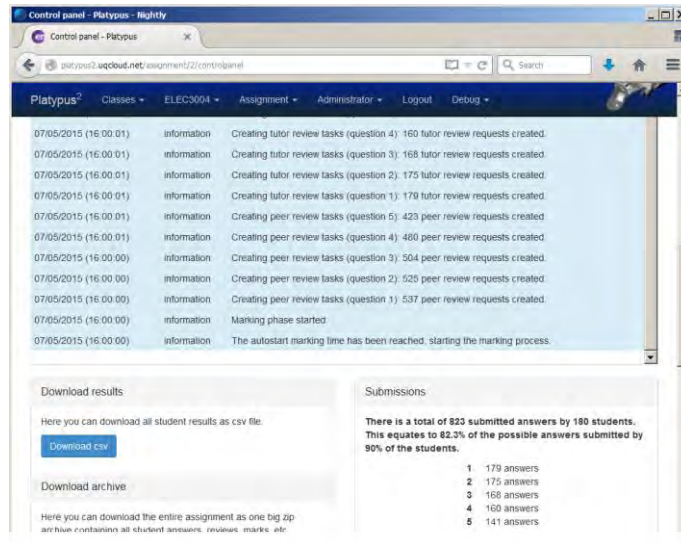
Will it Replace the Teacher?

Today we piloted [Gradescope](#)'s AI-assisted grading.

So here we are, CS 188 all done with **650 18-page** exams after just **three** hours of grading! AI graded AI 😊



Platypus: Available For Your Class Too



<http://openplatypus.org/>

Students will be students

THE SINGH-O-SCOPE




**INTRODUCING
SURYA'S NON-LINEAR PAYMENT SYSTEM**
(PATENT PENDING)

Just 50 easy inputs of \$14.99
OR
A simple single input of \$59.99

 (Additional charges apply for QUT students)

**CALL 1800-66-SINGH NOW
AND RECEIVE A FREE
ELMO T-SHIRT**



**AS SEEN ON
TV**

SURYA SINGH



**GRAND PRIX
FESTIVAL D'AVORIAZ
1985**

GENERATOR

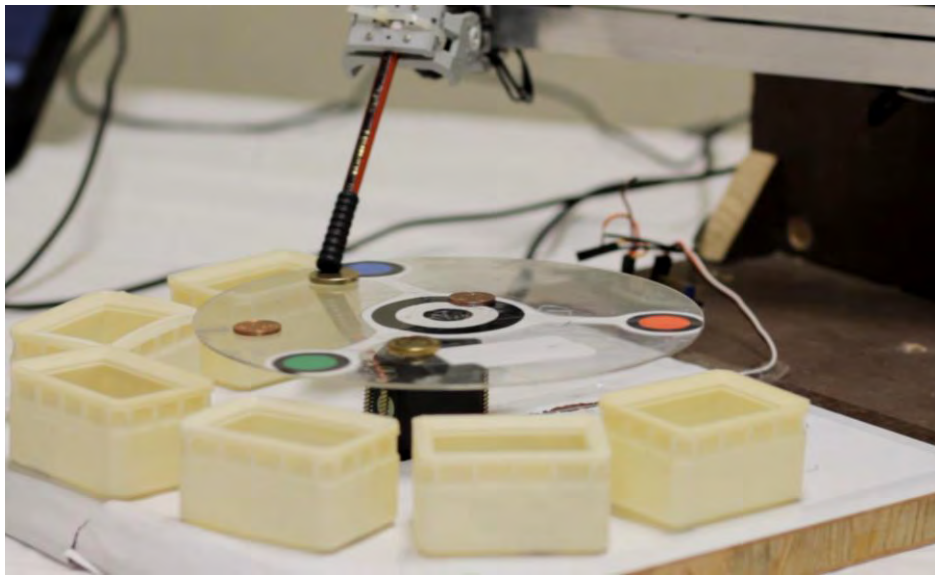
PREPARE TO BE GAUSSIAN ELIMINATED

MICHAEL BIEPIN • LINDA HAMILTON • PAUL WINFIELD

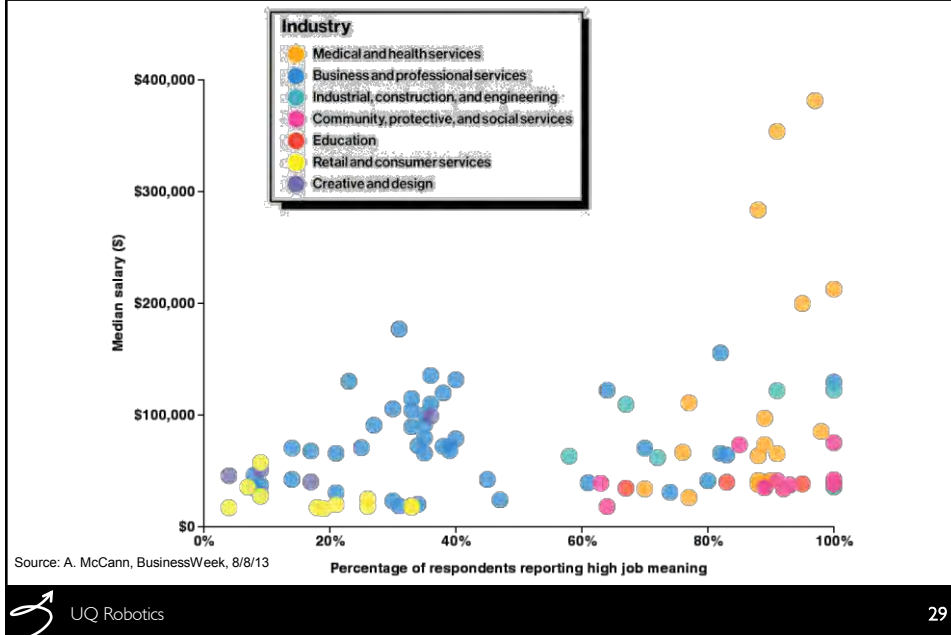
Collaboration: More than Just for Robots



The Goal: Create Change



Does Your Job Make the World a Better Place?



For More Information:

Get Started | Platypus - Nightly

Get Started | Platypus

platypus.upcloud.net/get-started

PLATYPUS Create Assignment Write Questions Assign Tutors Begin Marking Release Results

Get Started

Getting Started managing an assignment with Platypus only requires the following 5 steps

Create Assignment

New assignment for DEMO1001 Demo Data

<http://openplatypus.org/>

UQ Robotics

30

UQ Robotics: Dynamic Systems in Motion

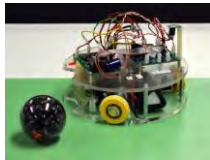
Diverse international
research group

Hanna Kurniwati (NUS/MIT)

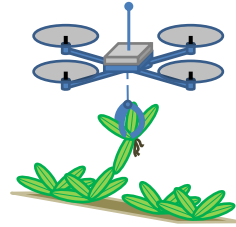
Paul Pounds (ANU/Yale)

Surya Singh (Stanford/Syd)

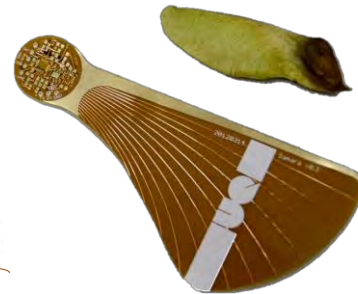
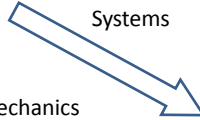
Planning Under Uncertainty
& Decision-Making



Aerial Robotics



Systems



Mechanics
of Motion



UQ Robotics

31

UQ Robotics

32