Are clicker response timings associated with grades?

Ms Suzan Orwell[1],
Miss Patience Mhindu[2] (Mathematics student),
Dr Peter Soan[2], Dr James Denholm-Price[2]

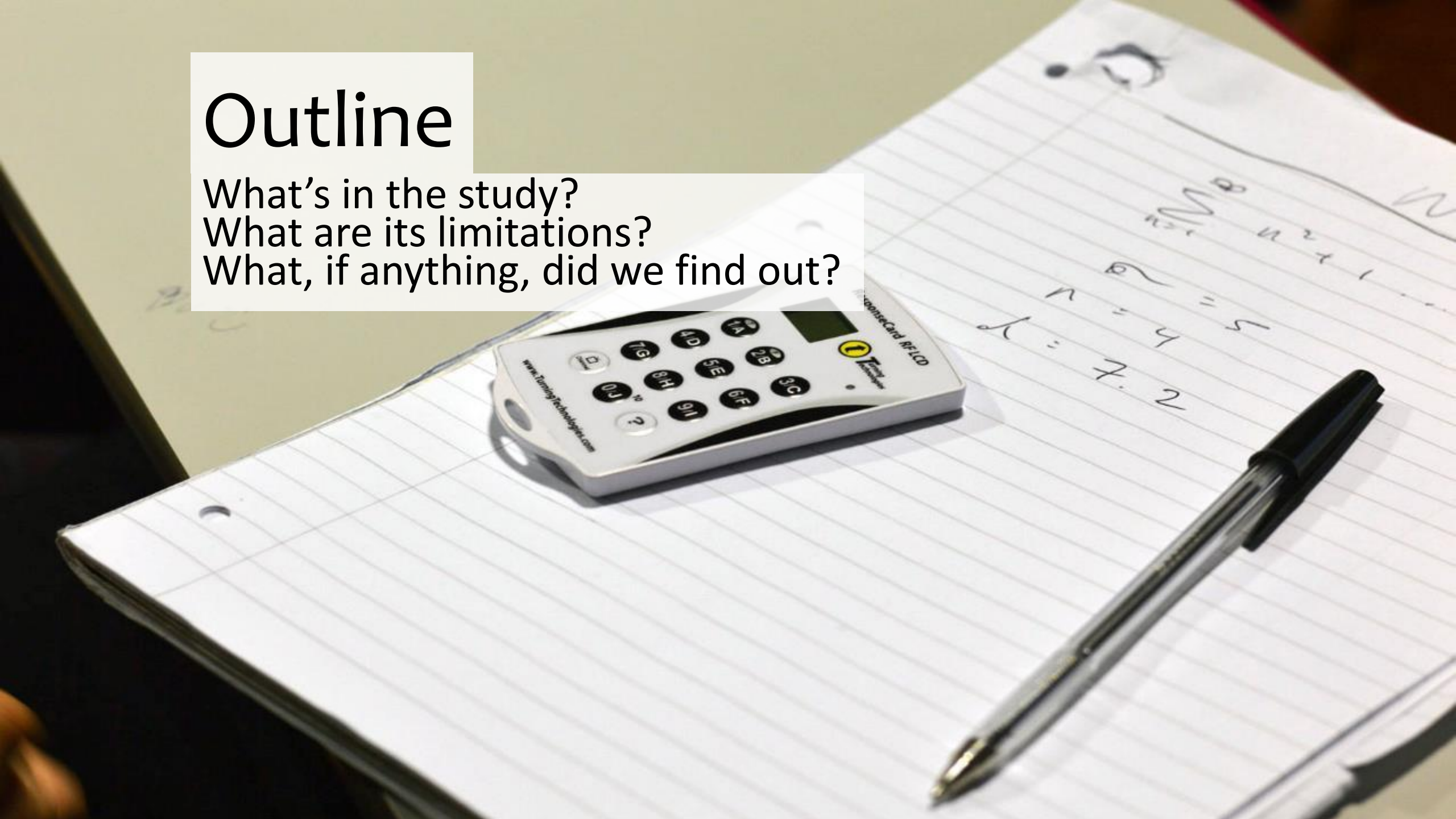
[1] Centre for Higher Education Research and Practice (CHERP)

[2] School of Computer Science and Mathematics,
Faculty of Science, Engineering and Computing

Kingston University London

Outline

What's in the study?
What are its limitations?
What, if anything, did we find out?



Kingston University London

Project Overview: Clickers

- Early decision to go with *hardware* clickers based on
  - research suggesting students' own devices are a distraction and less effective than a dedicated clicker
  - and anecdotal suggestion that it lowers the barrier to entry for busy/reluctant staff (this is a *staff engagement* project as well as being student-focused)
- This decision was eventually backed-up by our own evaluation:
  - Student survey: 50:50 divided between "happy to use own device" and "would prefer a clicker"
  - Staff focus group: "phones would be a distraction"
  - Student focus group – divided: some wanted a hybrid hardware and software option, some were (like staff) in favour of hardware only for similar reasons

Our staff views on clickers

- Staff would prefer for the students to use clickers and not their mobile phones.
  - mobile phones can distract the students
  - students may not have a mobile phone or may not be able to install the relevant app on their phone and students may not have reliable access to Wi-Fi.

"The University should provide Clickers or whatever technology they choose, the student shouldn't worry about it."

Kingston University London

Student and Staff views

- 2015: in-class survey of 216 first year students
  - Over 96% of responding students found the clicker easy to use.
  - 44% of the Mathematics students chose “Yes, to make learning more active.”
  - When asked if they felt that the quizzes have been beneficial to the learning, only 15 (7%) students somewhat/disagreed

“I like Clickers because they give immediate feedback”

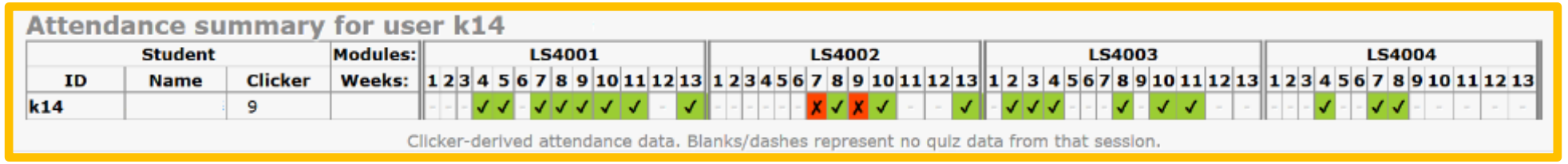
- 2015: staff focus group
  - the majority of staff said that using clickers in their teaching increased participation, reflection, feedback and peer discussion

‘Comparing to when we don’t use clickers, students will not necessarily put their hand up in case they get something wrong.’

‘The Clickers are a platform for quiet students.’

Web site: Data views for students

- Might showing students a record of their “engagement” influence their behaviour?
- Simple “widget” embedded into the VLE



- Student focus group answers: Predominantly *No!*

“Attendance monitoring will make no difference, motivation will make a difference … I know what I missed and I don’t need to look at my attendance”

(Majority representative view.)

“Having a lot of red crosses made me attend a particular module; it has motivated me to attend”

(Just one student!)

If not attendance, then what?

Clickers Project: Latest quiz scores for ANON STUDENT (k0000000) from MA5200

You attempted the most recent quiz on 19/10/15 where you scored 5/5:

Question: If a sequence is strictly monotonic increasing
Your answer: each term is smaller than the next ✓
You scored: 1/1

Question: Which two of the following together mean a sequence of real numbers converges?
Your answers:
The sequence is monotonic increasing ✓
The sequence is bounded above ✓
You scored: 4/4

Details for MA5200>
Your Clickers Modules
Your Engagement Profile

› So the students can see their responses, can they give us useful information about their in-class engagement?

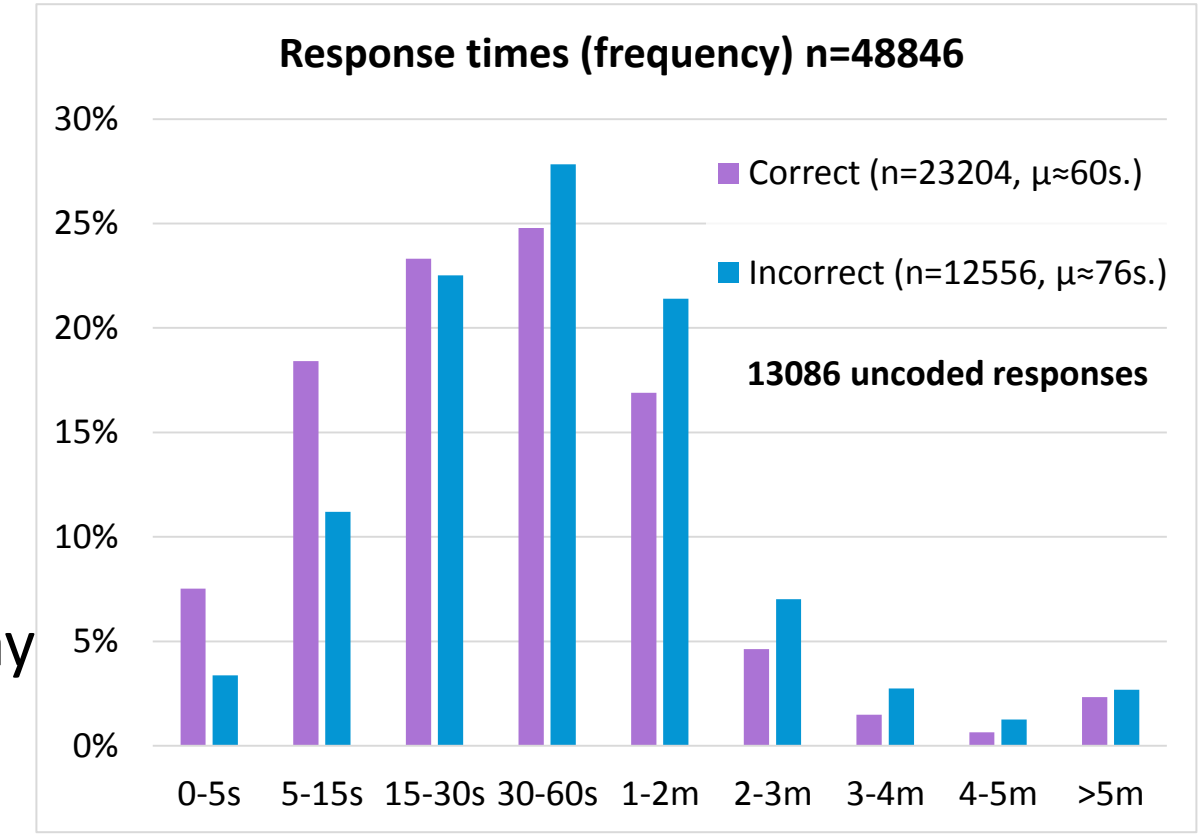
Hypothesis

Monitoring students' weekly online and CRS data can help to identify those who are disengaged and at risk of failing.

Kingston University London

What are response times typically?

- This is *everything* from our "clickers" database in 2015/6, all the messy data gathered from
  - 7 subject areas
  - 28 distinct members of staff
  - 30 different modules
  - 268 teaching sessions
- Overall, correct responses are faster ($p<0.01$) which agrees with one recent study but it may be case-specific…
- We'll look in detail at just a few modules :-)



Kingston University London

Retrospective analyses…

- Can we identify “at risk” students from their clickers usage pattern?
  - They’re “at risk” in this context if they fail the summative assessment.
  - Are clickers or quiz results correlated with this?
- Other research:
  - Prof Eric Mazur (Harvard) in Miller *et al.* (2014) with “Peer Instruction”
    - response times for correct answers are significantly faster than for incorrect answers,
    - in contrast to Heckler *et al*. (2010) with the converse result
  - Response times and student “certainty”
    - Probability of “high certainty” *decreases* with an *increase* in response time (Gvozdenko, 2010)

Retrospective analysis
Part 1:

5 weeks of “flipped” mathematics (linear algebra)

Online e-assessments (“Numbas” quizzes) before each class

Peer Instruction using clickers questions in class

Summative assessment a few weeks after the classes end

Kingston University London

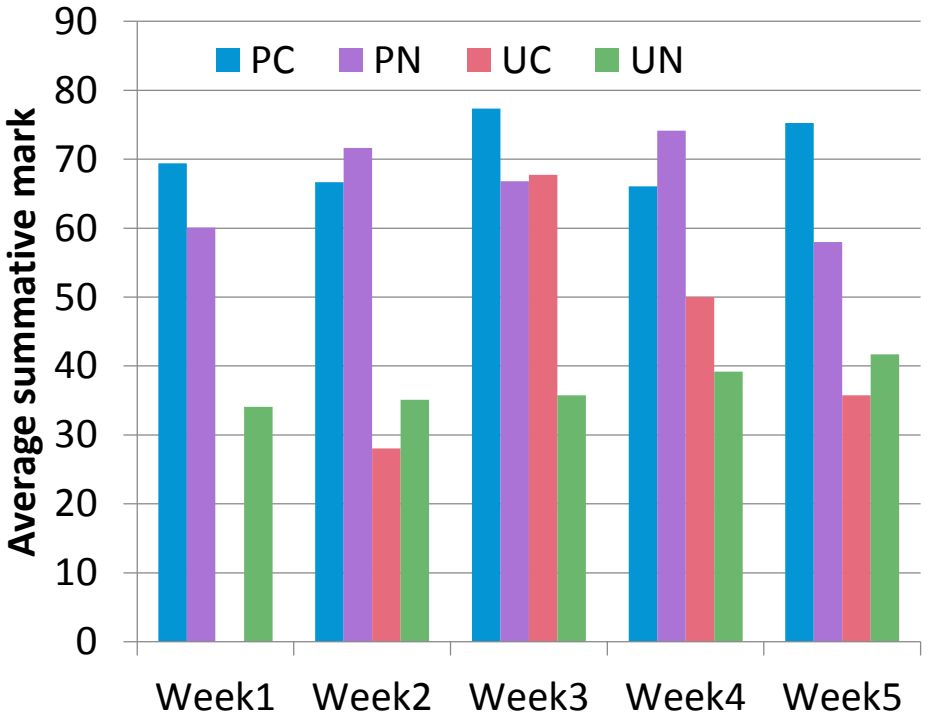
Data cleaning & preparation

- TEL researchers working with “live” modules and “real” (*aka* undisciplined?) academics :-# means datasets seldom “clean”
  - Out of roughly 50 “clicker” questions only 20 were usable
- Students classified as
  - P = “Prepared” if they did the *Numbas* quiz before class (otherwise “U”)
  - C = “Correct” if “clicker” responses were largely correct (otherwise “N”)
  - Students’ responses categorised 4 ways
    - PC = Prepared and generally Correct
    - PN = Prepared but Not generally Correct
    - UC = Unprepared but generally Correct
    - UN = Unprepared and Not generally correct

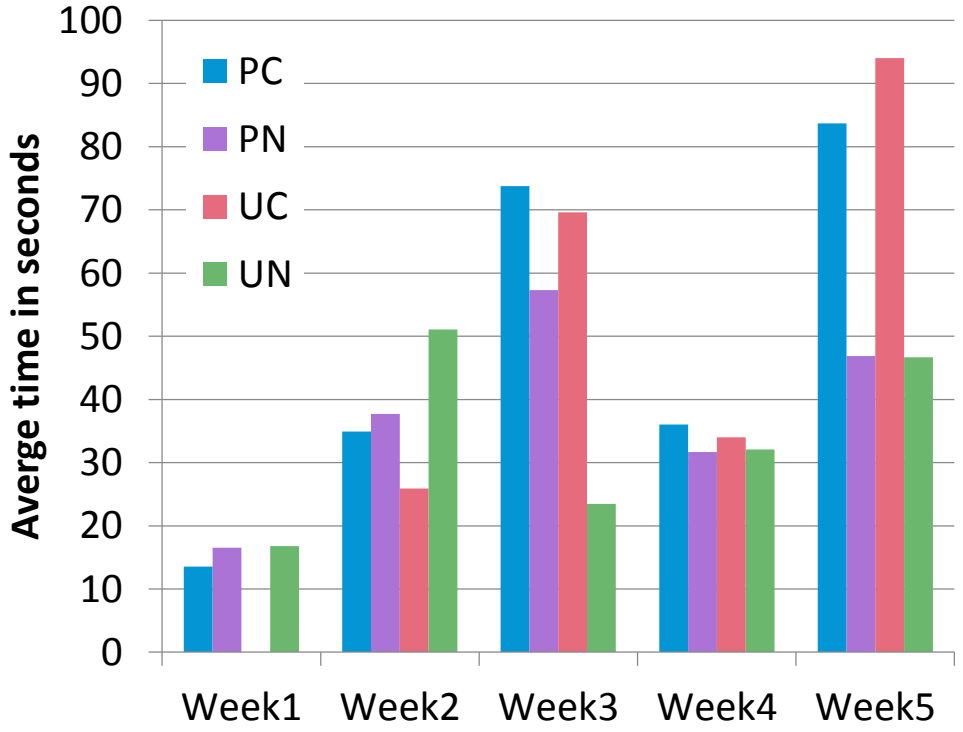
Kingston University London

Weekly data conclusions

PREPAREDNESS TRUMPS CRS



CORRECT CRS RESPONSES ARE SLOWER



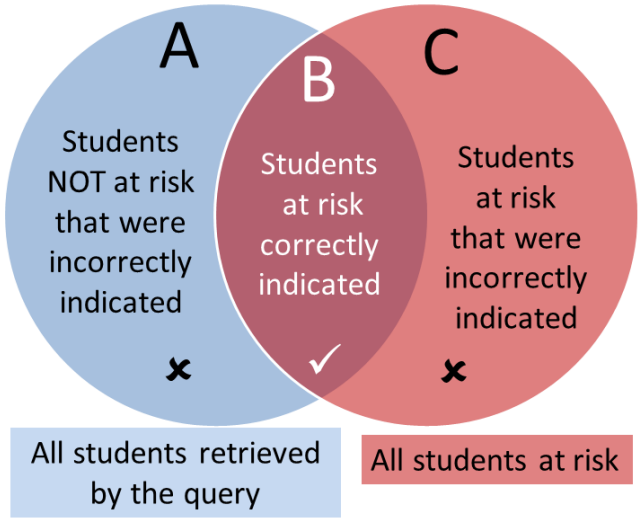
Kingston University London

Can monitoring students' weekly online and CRS data in a Flipped Learning environment help to identify those who are at risk of failing?

- Absence of preparation could be a factor in a risk indicator.
  - Varies week-by-week (question difficulty and/or direct link to summative assessment) but *e.g.* the mean mark for students who were prepared in week 4 is significantly higher than for unprepared students ($p<0.01$; $n=69$).
- Responding incorrectly is not a risk indicator.
  - No statistically significant difference was found in the mean summative marks between students who responded correctly in-class *versus* those who didn't.
- Responding correctly takes longer.
  - The difference in mean response time for correct responses was 73.7s as opposed to incorrect responses with a mean time of 69.5s, which was not quite statistically significant ($p=0.074$; **n=60**) but interesting nonetheless.

Is prediction possible from a really simple indicator?

- Unprepared students may be at risk of eventual failure – is such a naïve measure useful?
  - This is a relatively small study (n=80) but of the 6 “at risk” students (defined as taking the summative assessments but scoring less than 40% overall), 4 were unprepared and 2 were prepared.
  - To take this further more data are needed where students’ “preparedness” can be measured and compared to their final results.
    - (Flipped Learning is a useful testbed here.)



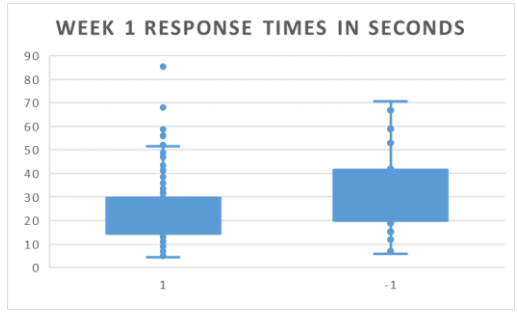
Retrospective analysis Part 2:

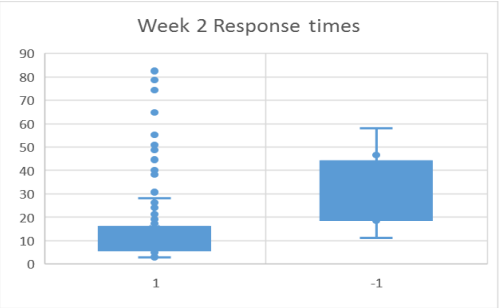
Looking for patterns in response times from one module each in Actuarial, Mathematics, Statistics, Life Sciences and Civil Engineering.

(Only the MA is “flipped”, the rest use clickers for varying degrees of active learning; the AM, MA & ST cohorts overlap.)

Zooming-out: Different subjects & cohorts

- The big picture was Mazur-like (correct responses were quicker) whilst the flipped classroom responses were the opposite so expand the picture we're now looking at module level.
- One 1st year module each from
  - Actuarial, Mathematics and Statistics – with overlapping cohorts of students
  - Life Sciences and Civil Engineering – with distinct students
    - between 18 and 113 individual questions
    - over 12 to 19 weeks of clicker usage
    - from 73 to 365 participants (Actuarial to Life Science)

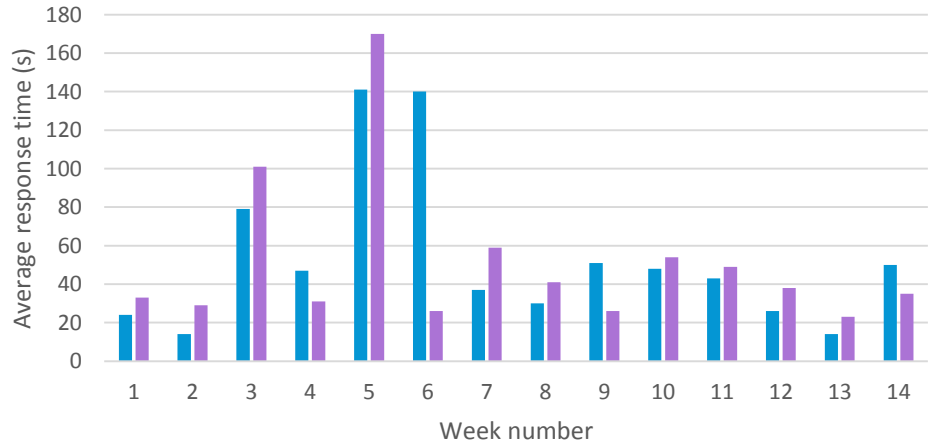




Response times for correct and incorrect answers are mostly similar (*e.g.* ST, AM, MA, LS) but not CE

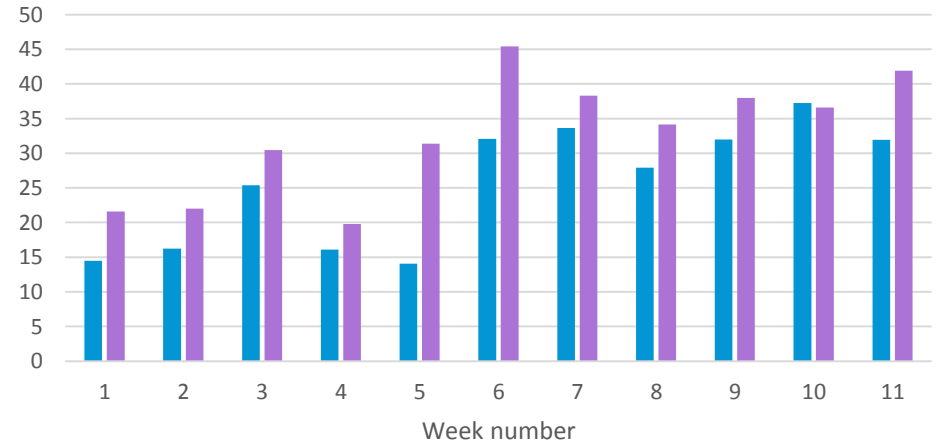
STATS MODULE RESPONSE TIMES

	Correct	Incorrect
Questions	*n*=14	
Mean	53s	51s
S.D .	41s	40s
T-test- two tail	*p*=0.89	



CIVIL ENGINEERING RESPONSE TIMES

	Correct	Incorrect
Questions	*n*=12	
Mean	26s	33s
S.D .	76s	73s
T-test- two tail	*p*=0.067	



Time taken to answer correctly

Time taken to answer incorrectly

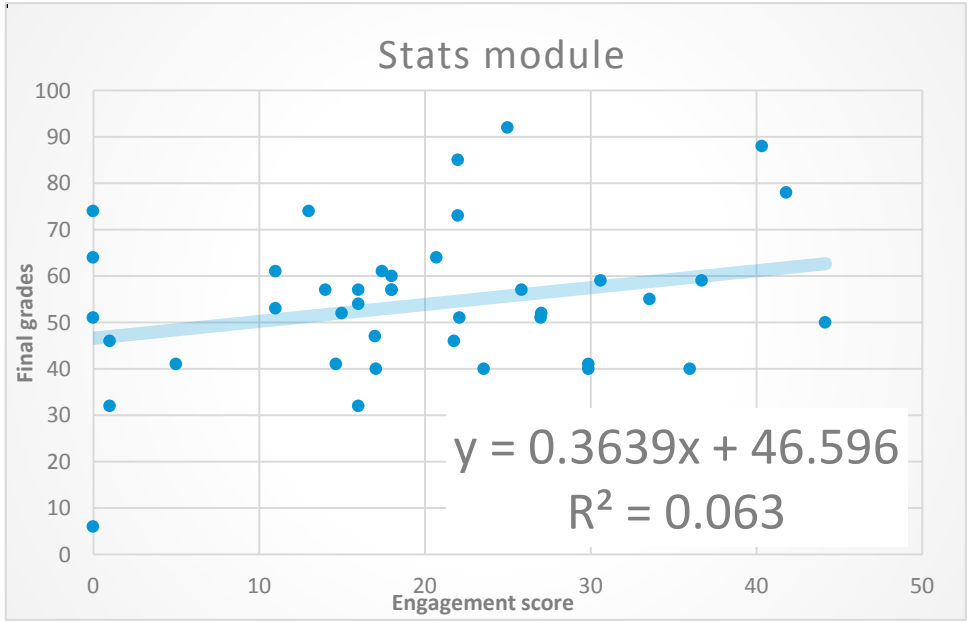
Module scores correlate (somewhat) with response patterns

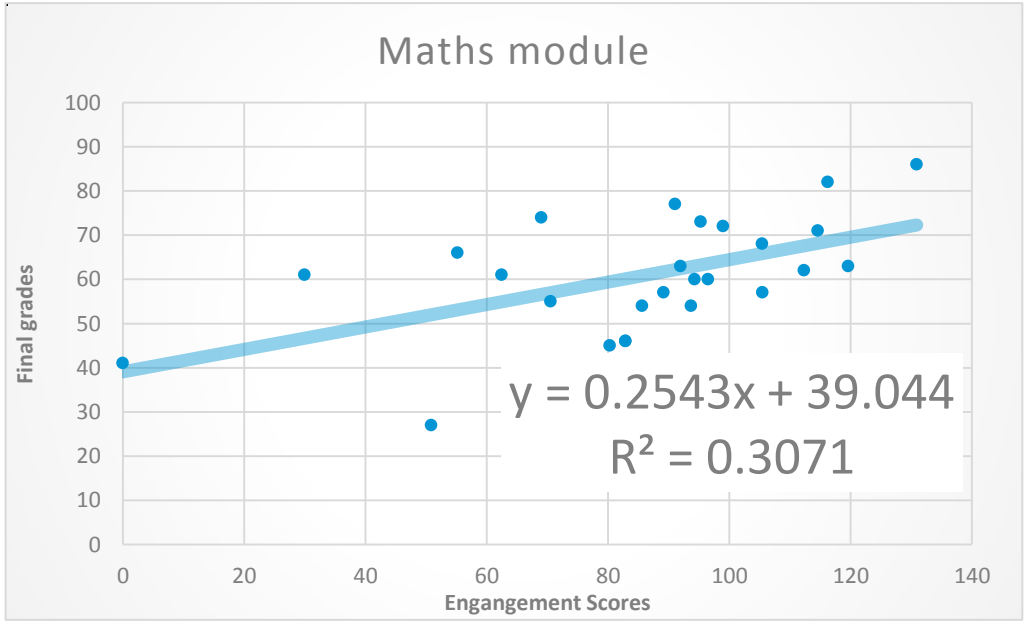
- FREQUENCY
  - There is a positive correlation between students who use the clickers often (whether correct or incorrect) and their grades.
- CORRECT RESPONSES
  - There is a stronger positive correlation between students who use clickers often whilst also getting correct responses with their grades.
- RESPONSE TIME
  - There is a positive correlation between final grades and shorter responses times. This is consistent with the analysis that students who tend to answer “correctly” take a shorter time to respond to questions compared to students that answer “incorrectly”.

Kingston University London

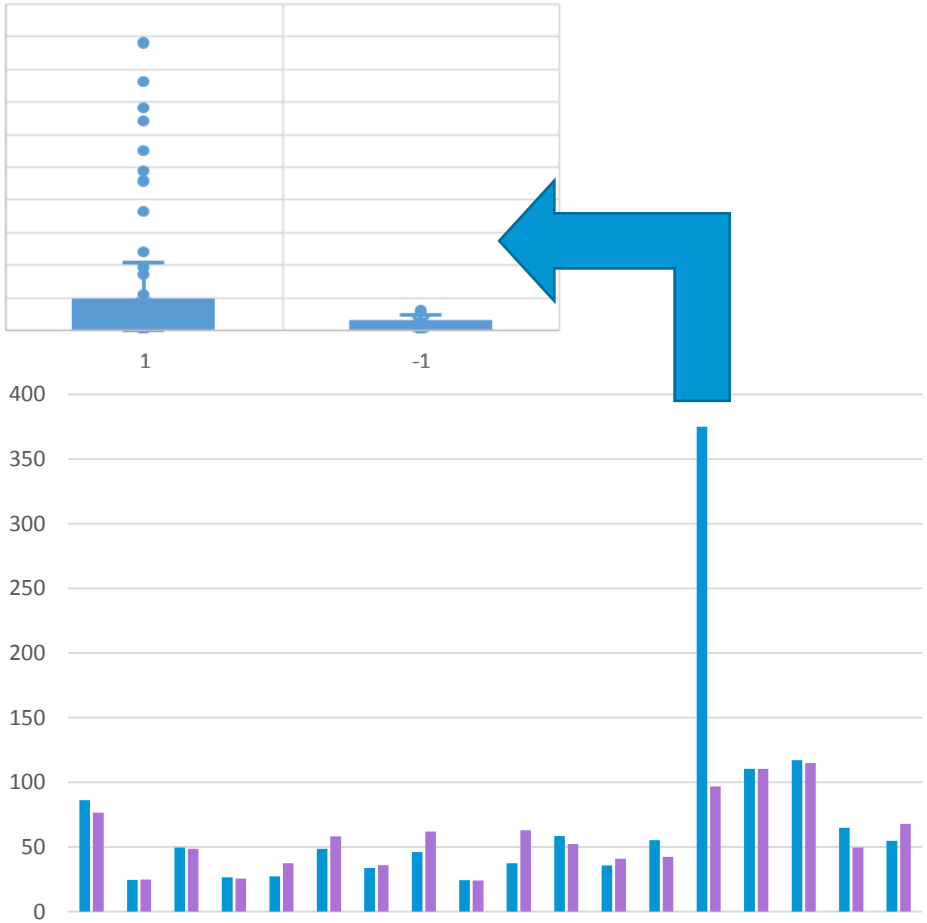
Measuring clicker "engagement" *vs.* grades

- Responses are optional – students have strongly varying response rates and may "click once" if they think that may simply "register their attendance"
- Define an engagement score: average correct response frequency × total responses
- Compare with final module results (which don't include any "marks" for clicker questions).
- Correlation varies – perhaps a measure of clicker question "quality" *i.e*. discrimination…





However … an "analytics" system needs to be aware that sometimes things "look wrong"



- *E.g.* it turns out that this "question" logs progress on tasks during a 2h. assessed lab session …
  - "correct" responses indicate progress throughout the lab
  - "incorrect" responses indicate students logging the first task only
  - these "outliers" were not included in this analysis but an automated system might not know better…

Kingston University London

Conclusions

- Clickers work ☺ Students *and* staff like them and find them easy enough to use.
- Flipped Learning and Peer Instruction environments with clickers can provide interesting/useful data
  - *E.g.* for "learning analytics" approaches
- Whether or not response times are correlated with
  - in-class correctness (maybe)
  - eventual summative progression (possibly)

it's clear that even "medium-sized" cohorts (n=80) are too small at this stage to draw statistically-significant conclusions … more data/more participating institutions are needed

Kingston University London

Thanks for your interest ☺

- Any questions?
- Anyone interested in clickers and working together?
  - j.denholm-price@kingston.ac.uk
  - http://sec.kingston.ac.uk