



# Comparative Judgement supporting learning & assessment in higher education

Matt Wingfield, Managing Director, Edtech Horizons  
and Chief Executive, The eAssessment Association

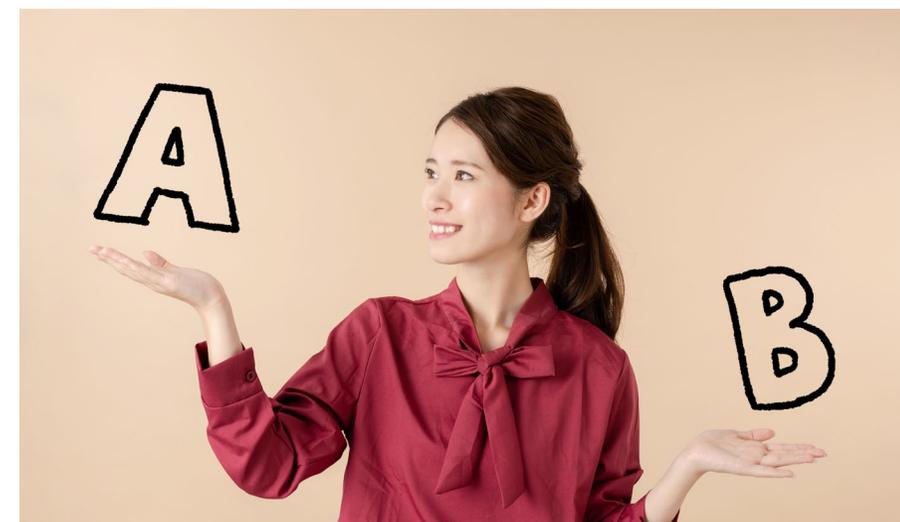
Dr. Jill Barber, Reader in pharmacy at  
University of Manchester and National Teaching Fellow

# Agenda

- Introduction to Comparative Judgement (CJ)
- The rationale behind adding ‘adaptivity’ – the ‘A’ in ACJ
- How the approach is being used to support assessment and feedback:
  - Summative and formative assessment of essays within the School of Pharmacy at the **University of Manchester, UK**
  - Early-stage learning interventions at **Purdue University, USA**
  - The student perspective – a short interview with a student at the **University of Edinburgh, UK**
  - Using comparative judgement as a formative peer assessment tool in schools - research by the **KTH Royal Institute of Technology, Sweden**
- Sourcing comparative judgement software
- Questions

# What is Comparative Judgement

- First conceived by Louis Thurstone, a psychometrician, in the 1920's
- Articulated through Thurstone's 'Law of Comparative Judgement'  
[https://en.wikipedia.org/wiki/Law\\_of\\_comparative\\_judgment](https://en.wikipedia.org/wiki/Law_of_comparative_judgment)
- Humans find it much easier to evaluate things in pairs, rather than in isolation
  - Which is heavier...?
  - Which is clearer...?
  - Which is better...?



# What is Comparative Judgement

The optician's eye-test is a great example of a modern application of Thurstone's law:

- Minimises uncertainty
  - Not how sharp or fuzzy is this image, but...
  - Which lens gives the clearest vision?
- Whilst still ending-up with a very accurate prescription





# Adaptive comparative judgement – a chance to change the paradigm?

Jill Barber, Reader in Pharmacy and National Teaching Fellow  
University of Manchester



# Headline use – the removal of bias



Student / Mark	Leon from Sydney	Jill from Croydon	Elena from Siberia	Doug from Kentucky	Jo from Bradford
Abdol	5	5	3	6	7
Betty	9	8	6	10	10
Clive	7	6	4	8	9
Doris	5	5	3	6	7
Ethan	3	1	0	4	4
Frederika	7	8	6	9	9
George	10	10	10	10	10

# It's good to remove bias



Student / Mark	Leon from Sydney	Jill from Croydon	Elena from Siberia	Doug from Kentucky	Jo from Bradford
Abdol	5				7
Betty				10	10
Clive	7	6			
Doris		5	3		
Ethan			0	4	
Frederika	7		6		
George		10			10

ACJ can remove bias from staff assessments but is quite time consuming: "An hour spent in assessment is an hour less of life" Sargeant

# Pilot question

- Imagine that you have been awarded a grant of £1 million to reduce premature deaths anywhere in the world. Choose a specific disease, group of diseases or other major cause of death and describe how you would spend the money to reduce the death rate.
- Marked by RM Compare (ACJ software); the most important criterion is a description of the problem. Criterion 2 is a description of the action. Criterion 3 is the costing.

# Results – staff and students marking fourth year student work

Accuracy 0.95, compares with 0.65 by conventional marking.

Name	Number of Judgements Made	WMS-SD
PD	3	-0.8
SE (PSS)	16	-1.0
FH (student)	37	-0.9
PD1	9	0.1
JH	26	-0.9
JB	37	1.3
DB	32	-0.6
SC	37	1.5
CD	37	2.0
PD2	12	-0.1
PD3	13	-0.8
EB	12	0.2

# Peer Assessment 1

- Imagine that you have been awarded a grant of £1 million to reduce premature deaths anywhere in the world. Choose a specific disease, group of diseases or other major cause of death and describe how you would spend the money to reduce the death rate. **Marked by RM**  
**Compare; the most important criterion is a description of the problem. Criterion 2 is a description of the action. Criterion 3 is the costing.**
- Students were required to carry out 9 comparisons.
- They were required to leave feedback.
- They were persuaded to fill in a short questionnaire about the process.



# Peer assessment 1

Question	Summary of answers (n=50)
Ease of use (compared with Turnitin Grademark)	ACJ easier (28), more difficult (2), similar (20)
Useful (compared with conventional revision)	Useful (37) quite useful (8) less useful than revision (5)
Fairness (were you convinced the marking was fair?)	Yes (27), probably but a bit uneasy (19) no (4)
Nine judgements per student were required to get a result. Was this number OK?	Yes, still learning from later judgements (25), prefer fewer (19) last few are a waste of time (6)
How many exercises would be appropriate per semester?	Two or three during Reading Week (27) was the most popular response.

Accuracy (reproducibility) was the same for students as staff (0.94).

# Peer Assessment 2 – Case Study

## Case Presentation

Bahar, a 27-year-old woman presents with numbness, fatigue, tiredness and cognitive impairment. Bahar is a heavy smoker, does not drink alcohol but lives on coffee. After MRI scans she is diagnosed with relapsing-remitting MS. Her doctor initiates interferon beta. When this fails, she is prescribed fingolimod.

## Answer the following questions (each answer not more than 100 words):

Describe Multiple Sclerosis epidemiology and risk factors

How many types of multiple sclerosis are there and how do they differ to each other?

What counselling should the pharmacist provide to Bahar as a patient new to interferon-beta treatment?

Which are the most important counselling points the pharmacist should provide to Bahar for fingolimod therapy?

How pharmacists can contribute in improving MS outcomes?

# Feedback to one student

Detailed description of risk factors and types of MS. Clear counselling points written as though talking directly to patient, with recommendations of how to manage side effects and monitoring required. Good integration of pharmacy practice. Only primary research sources used.

Good answers here and literature referenced for each point made. Lots of points covered for all questions and relation to pharmacy practice from sources is clearly demonstrated. One or two points on certain questions missed. Though these answers were very good, what I was given to compare it to was just better although this paper was still very good and above average.

You have provided good information on MS epidemiology, and suggested factors that contribute to incidence and prevalence. You are very thorough at counselling patients, and have provided the necessary side effects that patients should be educated on. Perhaps discussing the mechanism of action would be appropriate, so the patient knows fingolimod therapy works differently than previous therapy.

Accurate information with good layout, extensive referencing.

All four types of MS are mentioned, a good description of how they differ

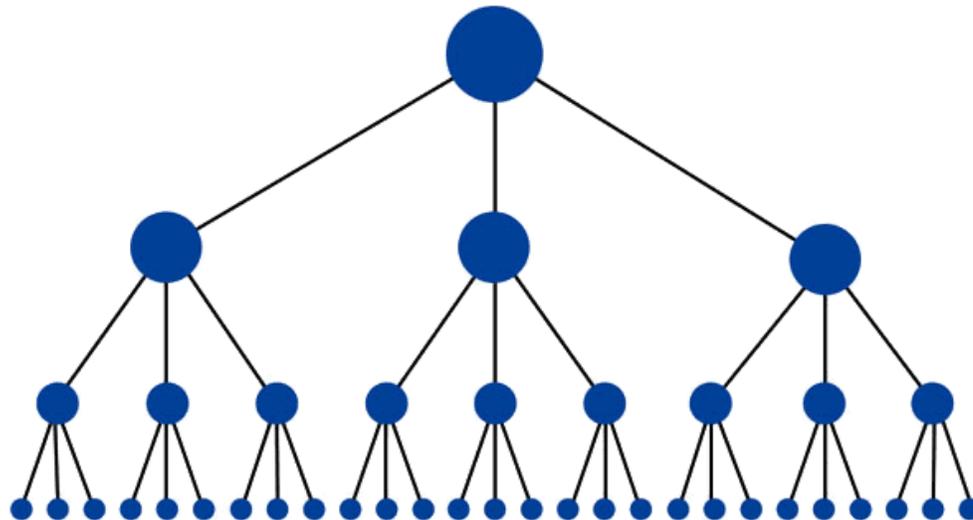
Counselling points are good, you may want to make them more patient-friendly and reassure the patient that these side effects can be managed or provide a safety net.

Good source of references, statements made are backed up to a good standard.

Good referencing, well written answers. Patient has been prioritised in the answers.

# Peer Assessment 2

- High quality feedback
- Accuracy only 0.55!
- Assessment not sufficiently simple or hierarchical.



# Changing the learning paradigm

- ACJ forces students to cooperate by critiquing one another's work.
- It gives insights into what good looks like.
- It enforces hierarchical learning. Eg
  - DNA is the genetic material
  - DNA is a double stranded helix
  - The DNA molecule has a diameter of 20 Å

But it's a new method and needs some work. So far, we have achieved excellent reproducibility and excellent feedback, but not in the same assessment!

# Acknowledgements

Dr Costas Demonacos

Steve Ellis (e-learning)

Matt Wingfield (RM consultant)

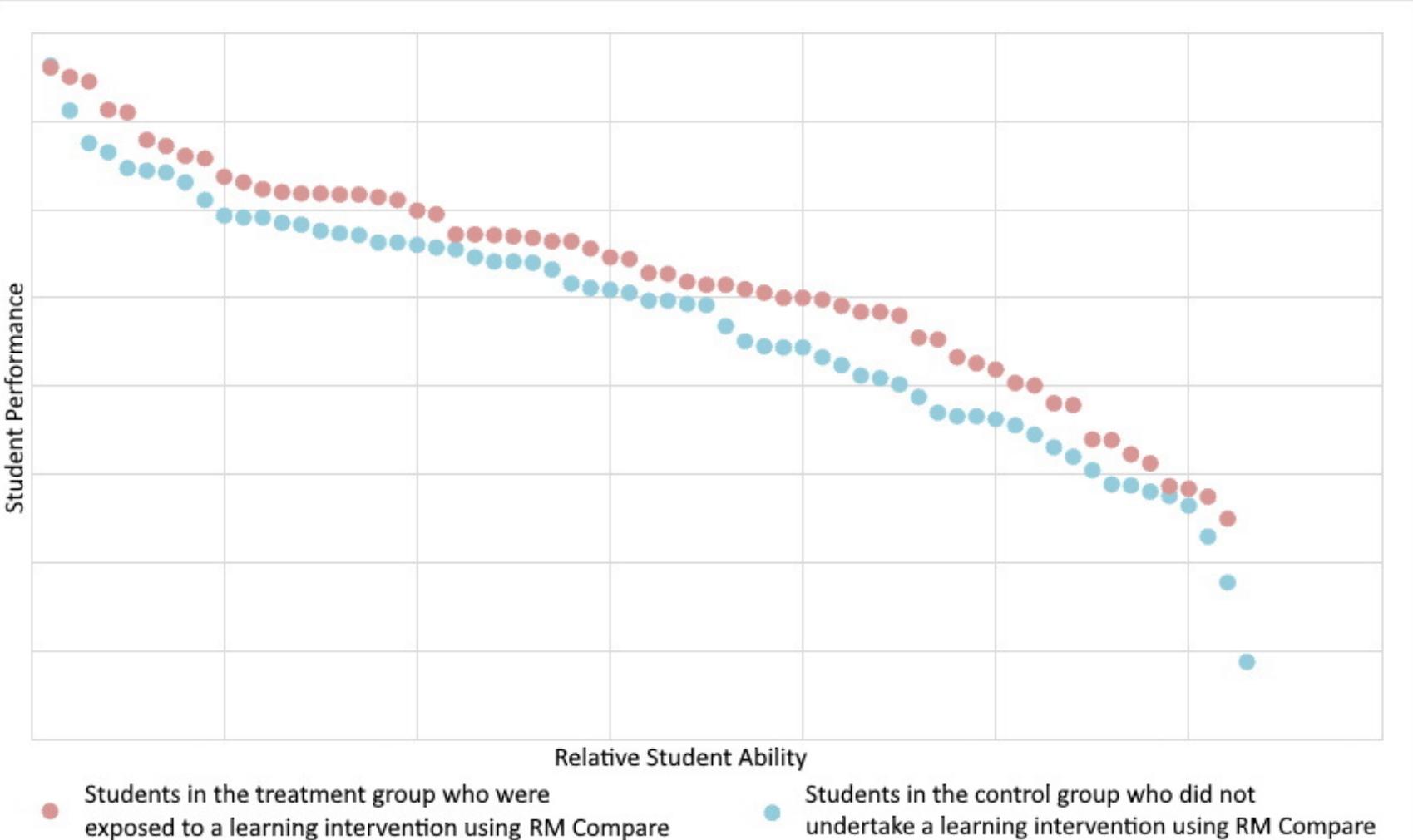
Many staff and students

# Learning by Evaluating

Purdue University led innovative use of ACJ to support learning through formative, peer-to-peer learning - academic paper [here](#)

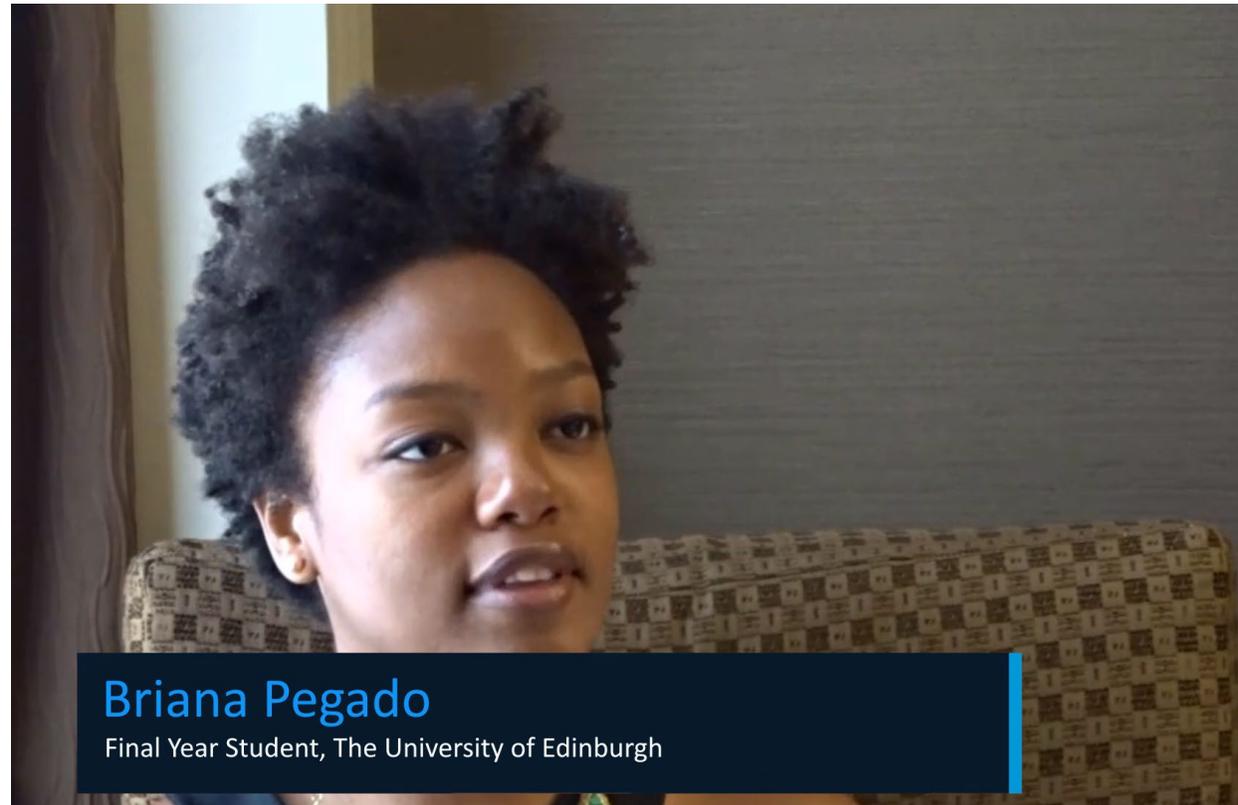
- 500 first year undergraduate students
- 50% of the group given a 10-20 minute learning intervention using ACJ at the beginning of the course using prior cohort example work
- Focus on STEM – design thinking
- All students undertook a peer assessment of their final work using ACJ at the end of the course
- Researchers:
  - Dr Nathan Mentzer, Associate Professor, Purdue University – [profile](#)
  - Dr Scott Bartholomew, Assistant Professor, Brigham Young University - [profile](#)

# Learning by Evaluating



# The Student Perspective

Using ACJ to peer assess essays at the University of Edinburgh, UK



**Briana Pegado**

Final Year Student, The University of Edinburgh



# ACJ in Sweden

- KTH Royal Institute of Technology, Stockholm:
  - Summative Assessments - National Tests in Design Technology & Science (Primary)
  - Formative Assessment - Design Technology (Middle School)
  - Research led by Dr Eva Hartell – [profile](#)



# In Summary

- There are several comparative judgement tools on the market



- RM Compare is the most mature Adaptive Comparative Judgement solution - <https://www.rm.com/products/rm-compare>
- RM Compare free trial here: <https://compare.rm.com/freetrial>
- ACJ can make a tangible impact to:
  - the reliability of assessing open-ended student work
  - maintaining assessment standards
  - support intuitive, student-led peer learning
  - boosting student attainment – across all ability levels
- researchED Haninge, Sweden 4<sup>th</sup> December 2021:
  - Register here: <https://researched.org.uk/event/researched-haninge/>





# Join the eAssessment Association - it's free:

<https://www.e-assessment.com/>

## Thank You Any Questions?

### Contacts:

Matt Wingfield:

[matt.wingfield@edtechhorizons.com](mailto:matt.wingfield@edtechhorizons.com)

Dr. Jill Barber:

[jill.barber@manchester.ac.uk](mailto:jill.barber@manchester.ac.uk)