Comparing standards of examination papers when there are no archived scripts

Ian Jones Loughborough University
Colin Foster Loughborough University
Jodie Hunter Massey University, New Zealand

13th Annual UK Rasch User Group Meeting
Cambridge 2019

BERJ British Educational Research Journal

BERA BRITISH EDUCATIONAL RESEARCH ASSOCIATION

British Educational Research Journal
Vol. 42, No. 4, August 2016, pp. 543–560

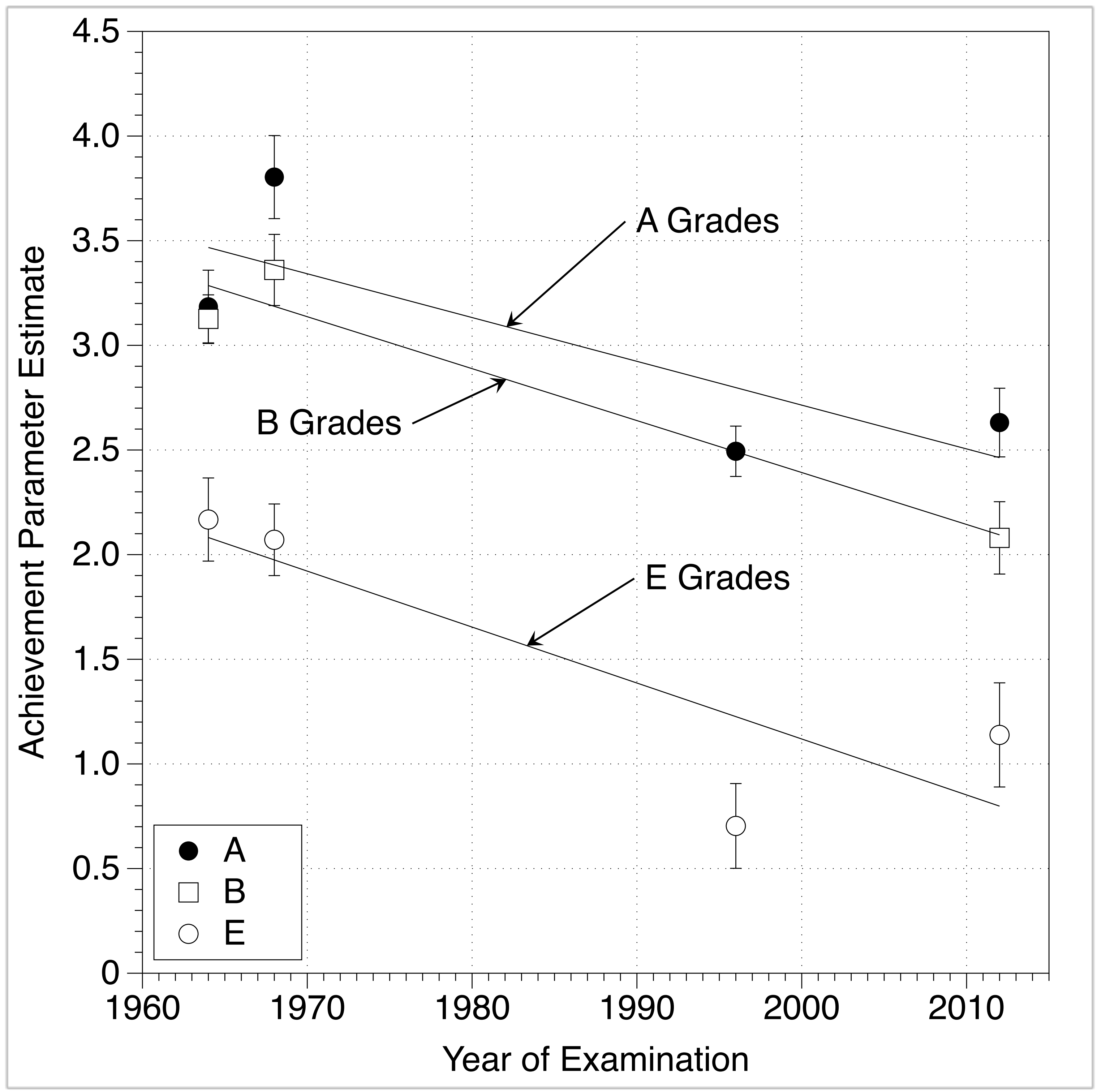
DOI: 10.1002/berj.3224

Fifty years of A-level mathematics: have standards changed?

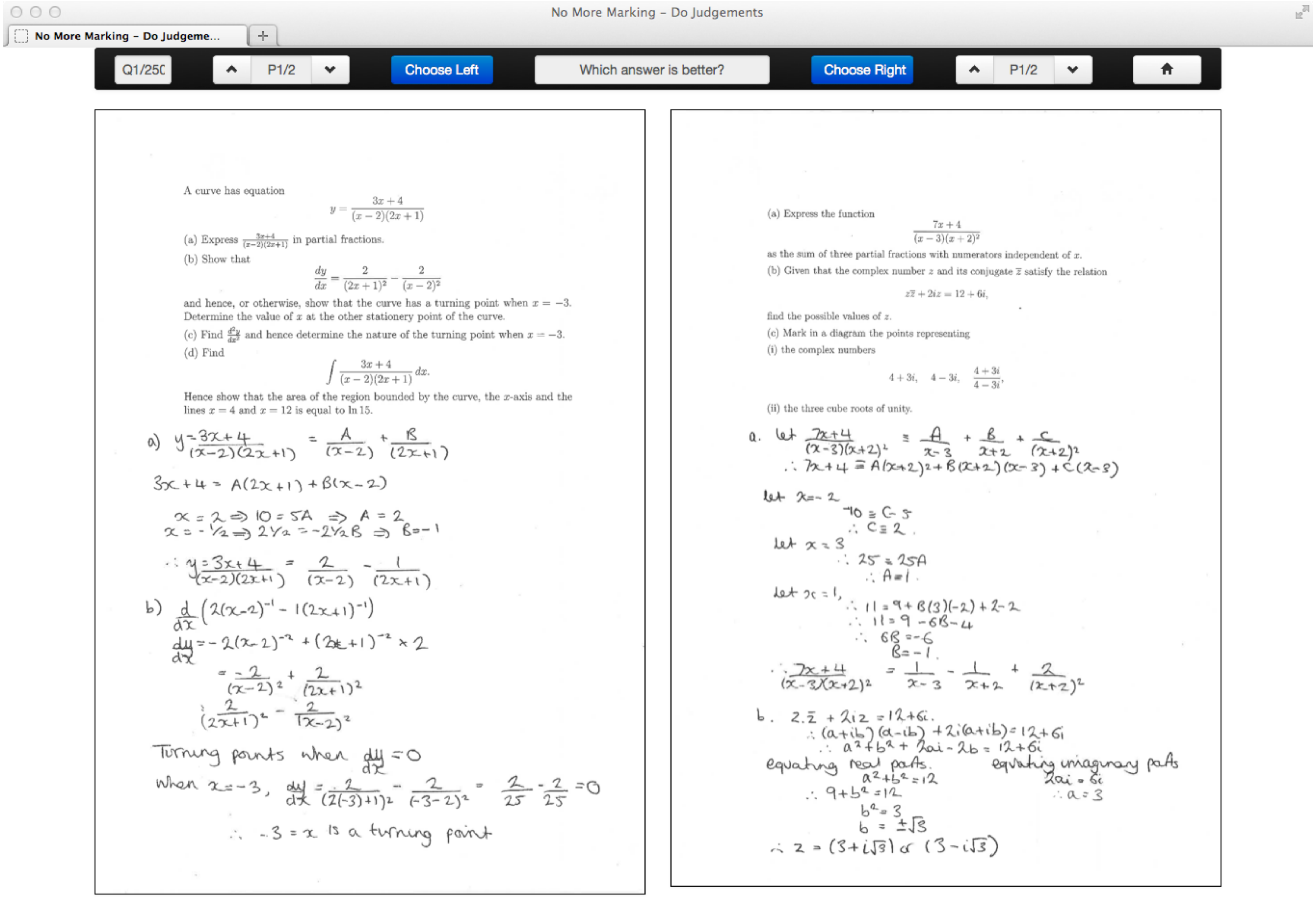
Ian Jones[a,*], Chris Wheadon[b], Sara Humphries[c] and Matthew Inglis[a]

[a]*Mathematics Education Centre, Loughborough University, UK;* [b]*No More Marking Ltd.;* [c]*Ofqual*

BERJ (2016) Results



BERJ (2016) Method



A curve has equation

$$y = \frac{3x+4}{(x-2)(2x+1)}$$

(a) Express $\frac{3x+4}{(x-2)(2x+1)}$ in partial fractions.

(b) Show that

$$\frac{dy}{dx} = \frac{2}{(2x+1)^2} - \frac{2}{(x-2)^2}$$

and hence, or otherwise, show that the curve has a turning point when $x = -3$. Determine the value of x at the other stationary point of the curve.

(c) Find $\frac{d^2y}{dx^2}$ and hence determine the nature of the turning point when $x = -3$.

(d) Find

$$\int \frac{3x+4}{(x-2)(2x+1)}\,dx.$$

Hence show that the area of the region bounded by the curve, the x-axis and the lines $x = 4$ and $x = 12$ is equal to $\ln 15$.

a) $y = \frac{3x+4}{(x-2)(2x+1)} = \frac{A}{(x-2)} + \frac{B}{(2x+1)}$

$3x+4 = A(2x+1) + B(x-2)$

$x = 2 \Rightarrow 10 = 5A \Rightarrow A = 2$

$x = -1/2 \Rightarrow 2 1/2 = -2 1/2 B \Rightarrow B = -1$

$\therefore y = \frac{3x+4}{(x-2)(2x+1)} = \frac{2}{(x-2)} - \frac{1}{(2x+1)}$

b) $\frac{d}{dx}\left(2(x-2)^{-1} - 1(2x+1)^{-1}\right)$

$\frac{dy}{dx} = -2(x-2)^{-2} + (2x+1)^{-2} \times 2$

$= \frac{-2}{(x-2)^2} + \frac{2}{(2x+1)^2}$

$\frac{2}{(2x+1)^2} - \frac{2}{(x-2)^2}$

Turning points when $\frac{dy}{dx} = 0$

When $x = -3$, $\frac{dy}{dx} = \frac{2}{(2(-3)+1)^2} - \frac{2}{(-3-2)^2} = \frac{2}{25} - \frac{2}{25} = 0$

$\therefore -3 = x$ is a turning point

(a) Express the function

$$\frac{7x+4}{(x-3)(x+2)^2}$$

as the sum of three partial fractions with numerators independent of x.

(b) Given that the complex number z and its conjugate $\bar{z}$ satisfy the relation

$$z\bar{z} + 2iz = 12 + 6i,$$

find the possible values of z.

(c) Mark in a diagram the points representing

(i) the complex numbers

$$4+3i, \quad 4-3i, \quad \frac{4+3i}{4-3i},$$

(ii) the three cube roots of unity.

a. let $\frac{7x+4}{(x-3)(x+2)^2} \equiv \frac{A}{x-3} + \frac{B}{x+2} + \frac{C}{(x+2)^2}$

$\therefore 7x+4 \equiv A(x+2)^2 + B(x+2)(x-3) + C(x-3)$

let $x = -2$

$-10 \equiv C \cdot 5$

$\therefore C \equiv 2$

let $x = 3$

$\therefore 25 \equiv 25A$

$\therefore A \equiv 1$

let $x = 1$,

$\therefore 11 = 9 + B(3)(-2) + 2 - 2$

$\therefore 11 = 9 - 6B - 4$

$\therefore 6B = -6$

$B = -1$

$\therefore \frac{7x+4}{(x-3)(x+2)^2} = \frac{1}{x-3} - \frac{1}{x+2} + \frac{2}{(x+2)^2}$

b. $z.\bar{z} + 2iz = 12 + 6i$.

$\therefore (a+ib)(a-ib) + 2i(a+ib) = 12 + 6i$

$\therefore a^2 + b^2 + 2ai - 2b = 12 + 6i$

equating real parts: $a^2 + b^2 = 12$; $\therefore 9 + b^2 = 12$; $b^2 = 3$; $b = \pm\sqrt{3}$

equating imaginary parts: $2ai = 6i$; $\therefore a = 3$

$\therefore z = (3 + i\sqrt{3})$ or $(3 - i\sqrt{3})$

BERJ (2016) Preparation

- Question papers typeset for consistency.
- Candidate responses transcribed for consistency.
- 66 scripts divided into 546 questions and uploaded to website for judging procedure.

BERJ (2016) Preparation

- [illegible]pers typeset for consistency.
- Cand[illegible]es transcribed for consistenc[illegible]
- 66 scripts divide[illegible]stions and uploaded to website [illegible]cedure.

TIME-CONSUMING AND EXPENSIVE

BERJ (2016) Preparation

- … pers typeset for … cy.
- Cand…
  consistenc…
- 66 scri… stions and
  … osite … ocedure.

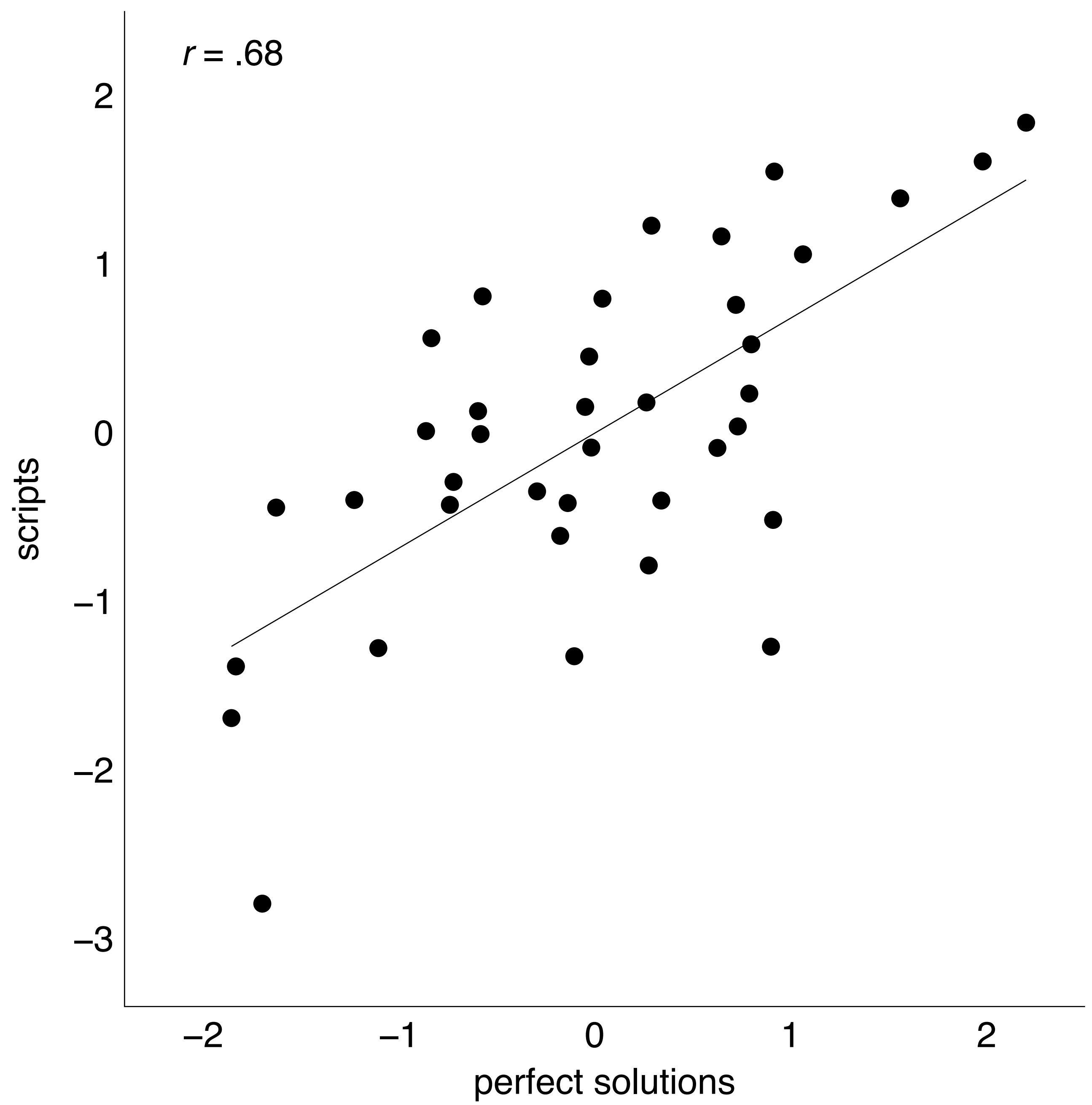
TIME-CONSUM… EXPENSIVE

REQUIRES GRADED ARCHIVE

Can we apply CJ to standards comparison without graded scripts?

Hope 1

Model solutions vs. graded scripts



Hope 2

RESEARCH IN MATHEMATICS EDUCATION, 2017
VOL. 19, NO. 2, 112–129
https://doi.org/10.1080/14794802.2017.1334576

Routledge
Taylor & Francis Group

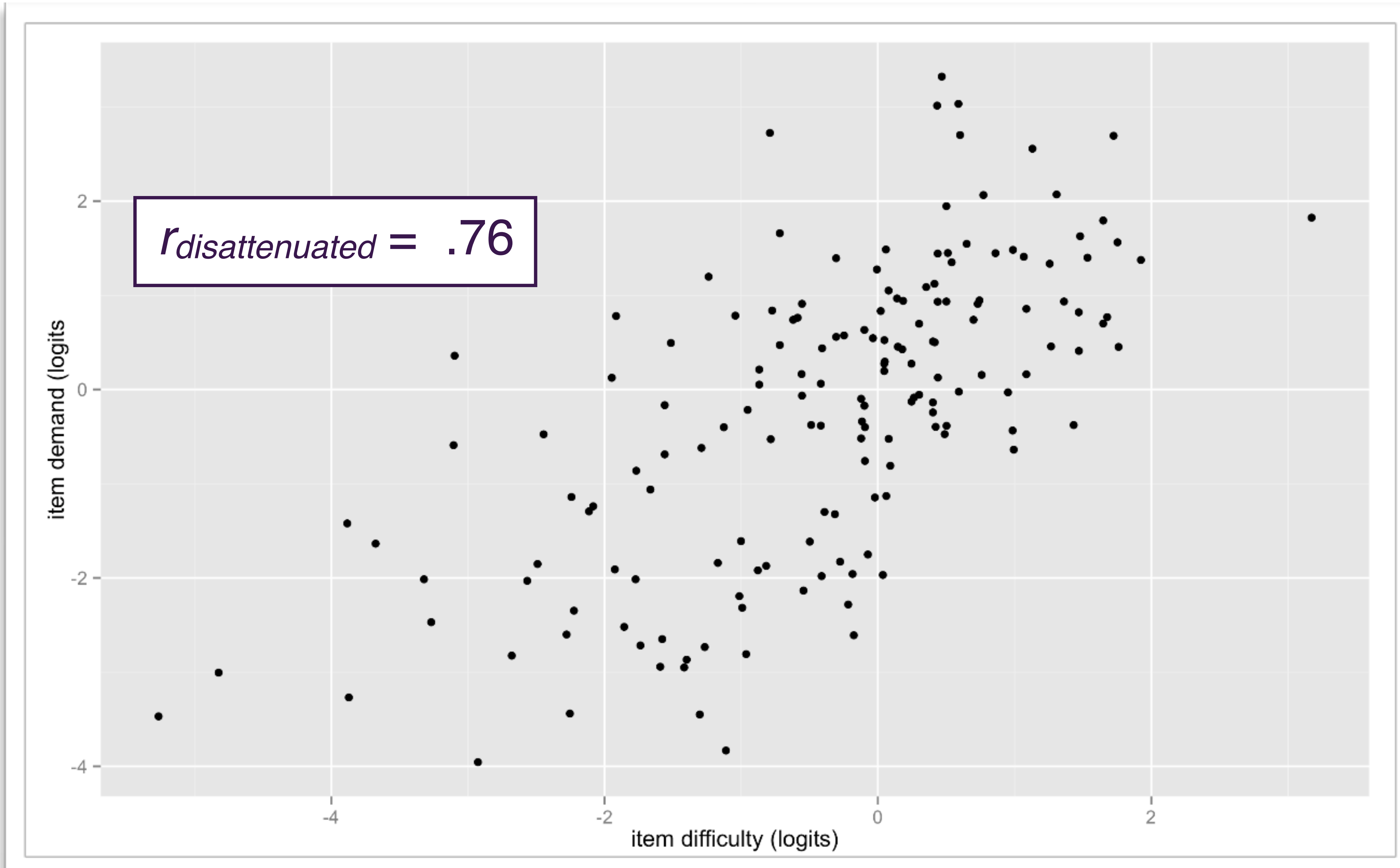
Check for updates

An investigation of construct relevant and irrelevant features of mathematics problem-solving questions using comparative judgement and Kelly's Repertory Grid

Stephen D. Holmes, Qingping He and Michelle Meadows

Office of Qualifications and Examinations Regulation, Coventry, UK

Expected vs Actual Difficulty



From page 64 of Ofqual (2015) *A Comparison of Expected Difficulty, Actual Difficulty and Assessment of Problem Solving across GCSE Maths Sample Assessment Materials*. Report Ofqual/15/5679.

Can we apply CJ to standards comparison without graded scripts?

Study 1

Judging non-typeset items only.

Study 1: comparative judgement

- Exam papers from 1964, 1968, 1996, 2012 (as per BERJ, 2016).
- Split into 42 question items.
- Judged by 8 maths PhD students, total 670 pairwise judgements.
- Internal consistency, $SSR = .91$.
- Inter-rater reliability (split-halves, 100 iterations), $r_{median} = .79$.

Study 1: analysis

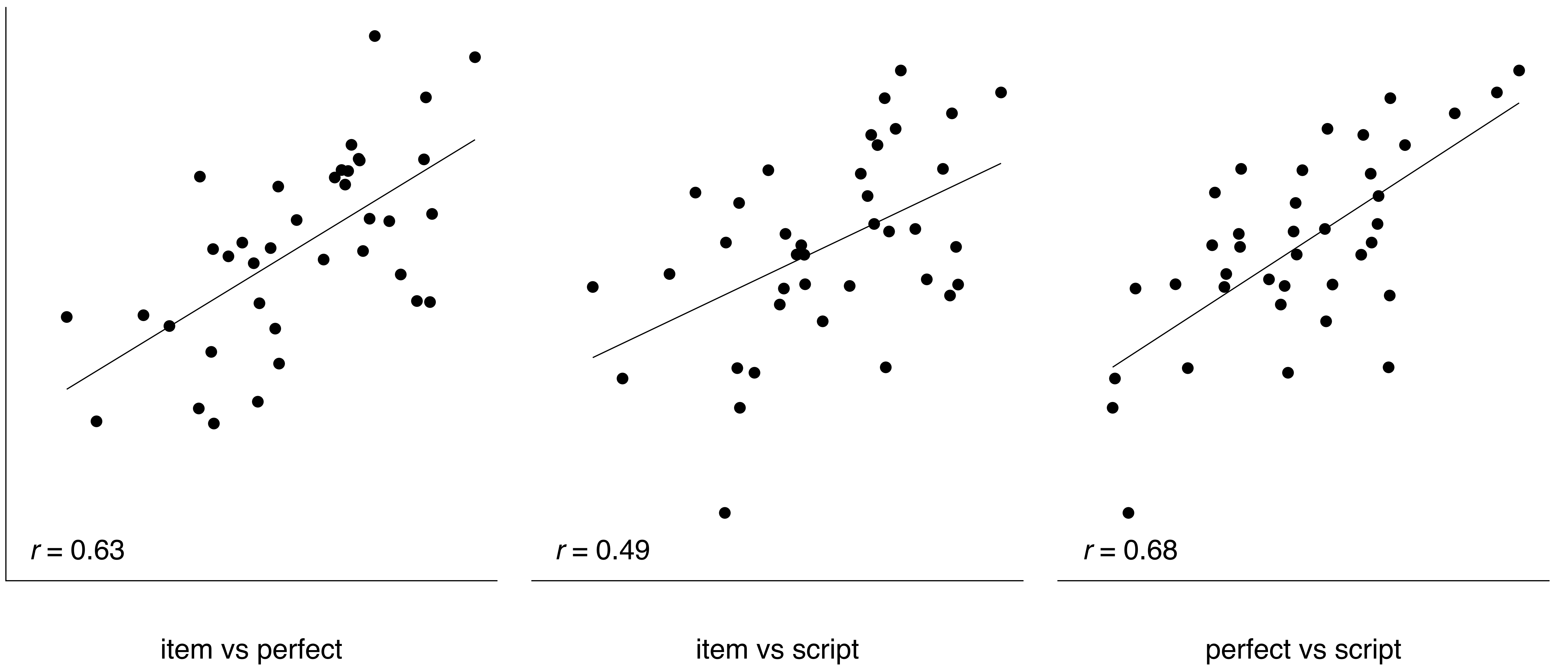
We compared item scores with

(i) the scores of the perfect candidates from the BERJ paper (“perfect scores”), and

(ii) the scores of the real scripts from the BERJ paper (“script scores”).

(Scores were available for 38 of the 42 questions judged for Study 1.)

Study 1: correlations



Study 1: variance explained

- Year as a predictor of item score (BERJ, 2016).
- **Present study**
  $F(1,36) = 35.83$, $p < .001$, $R^2 = .500$, year as predictor: $b = -0.06$.
- **Perfect scores** (BERJ, 2016)
  $F(1,36) = 13.94$, $p < .001$, $R^2 = .279$, year as predictor: $b = -0.03$.
- **Script scores** (BERJ, 2016)
  $F(1,36) = 13.62$, $p < .001$, $R^2 = .274$, year as predictor: $b = -0.03$.

Study 1: variance explained

- Year as a predictor of item score (BERJ, 2016).
- **Present study**
  $F(1,36) = 35.83$, $p < .001$, $\boldsymbol{R^2 = .500}$, year as predictor: $b = -0.06$.
- **Perfect scores** (BERJ, 2016)
  $F(1,36) = 13.94$, $p < .001$, $\boldsymbol{R^2 = .279}$, year as predictor: $b = -0.03$.
- **Script scores** (BERJ, 2016)
  $F(1,36) = 13.62$, $p < .001$, $\boldsymbol{R^2 = .274}$, year as predictor: $b = -0.03$.

Can we apply CJ to standards comparison without graded scripts?

Study 2

Judging (i) typeset papers only, and
(ii) typeset papers with perfect solutions.

Study 2: Exam papers

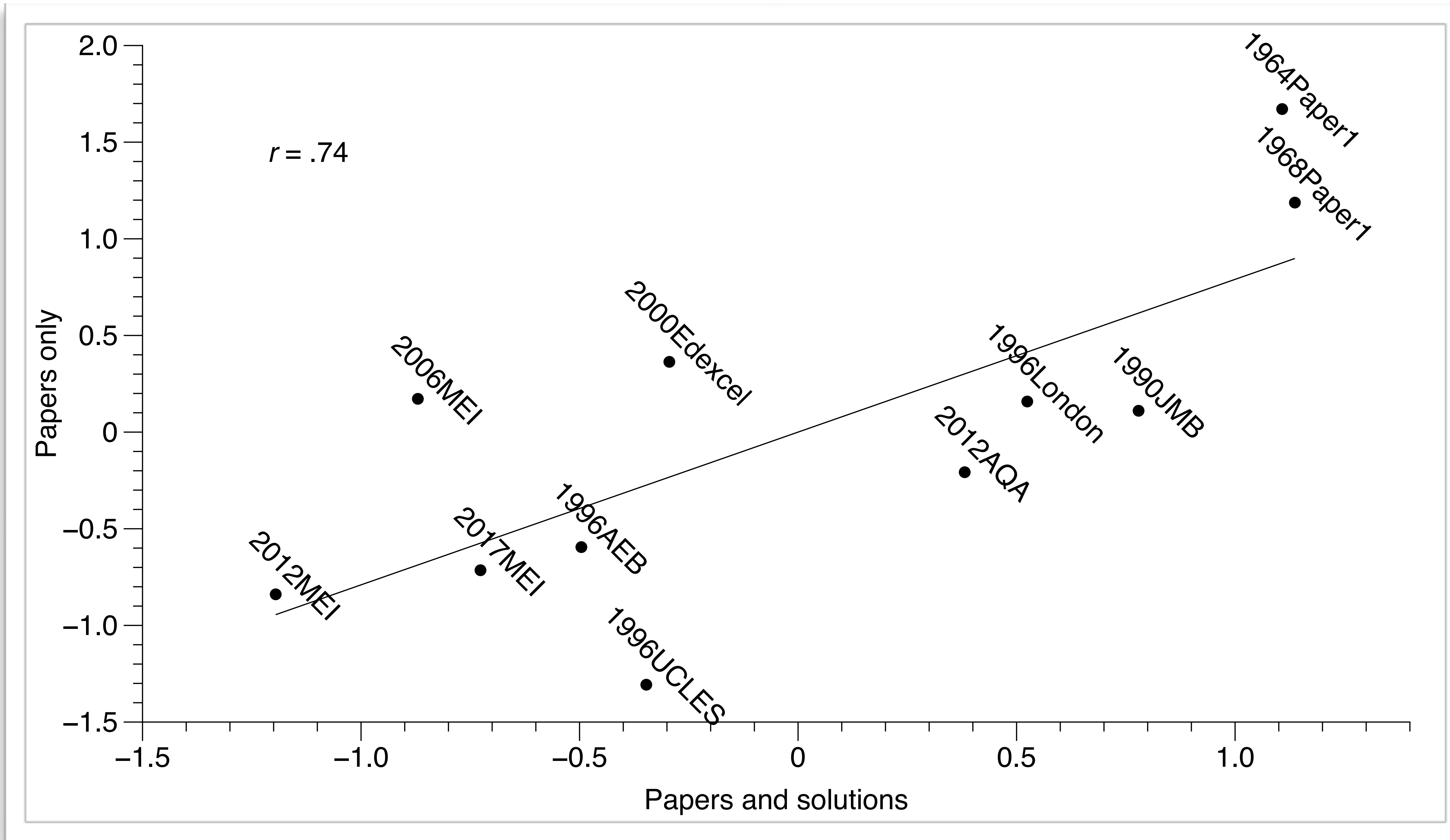
Year	Boards
1964	JMB*
1968	JMB*
1990	JMB
1996	AEB*, London, UCLES
2000	Edexcel
2006	MEI
2012	AQA*, MEI
2017	MEI

* included in BERJ (2016).

Study 2: comparative judgement

- (i) *Papers Only.*
    - Judged by 5 maths PhD students, total 250 judgements, $SSR = .84$.
- (ii) *Papers and Solutions.*
    - Judged by 5 different maths PhD students, total 330 judgements, $SSR = .87$.

Study 2: correlation



Study 2: analysis

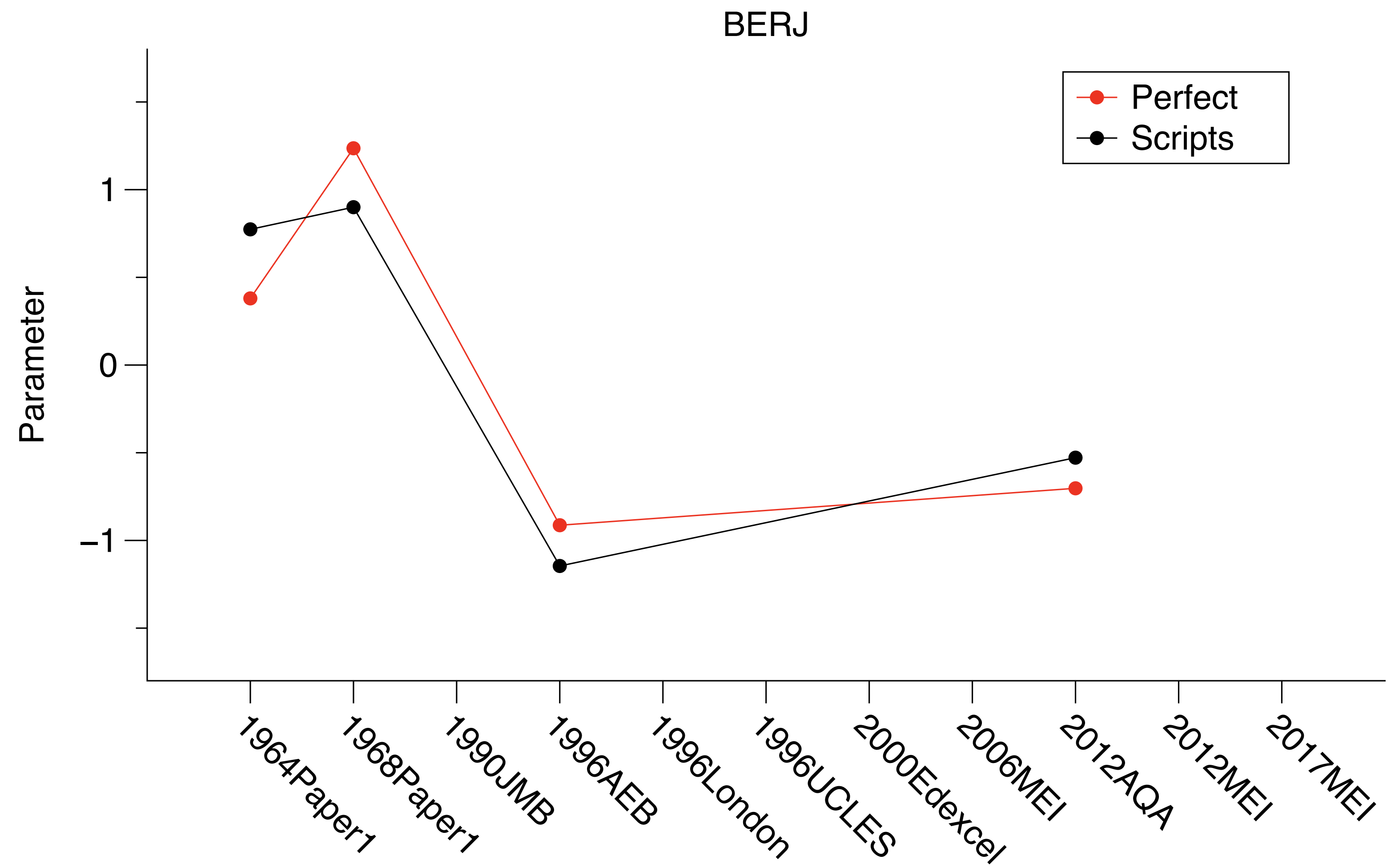
We compared exam paper scores with

(i) the scores of the perfect candidates from the BERJ paper (“perfect scores”), and

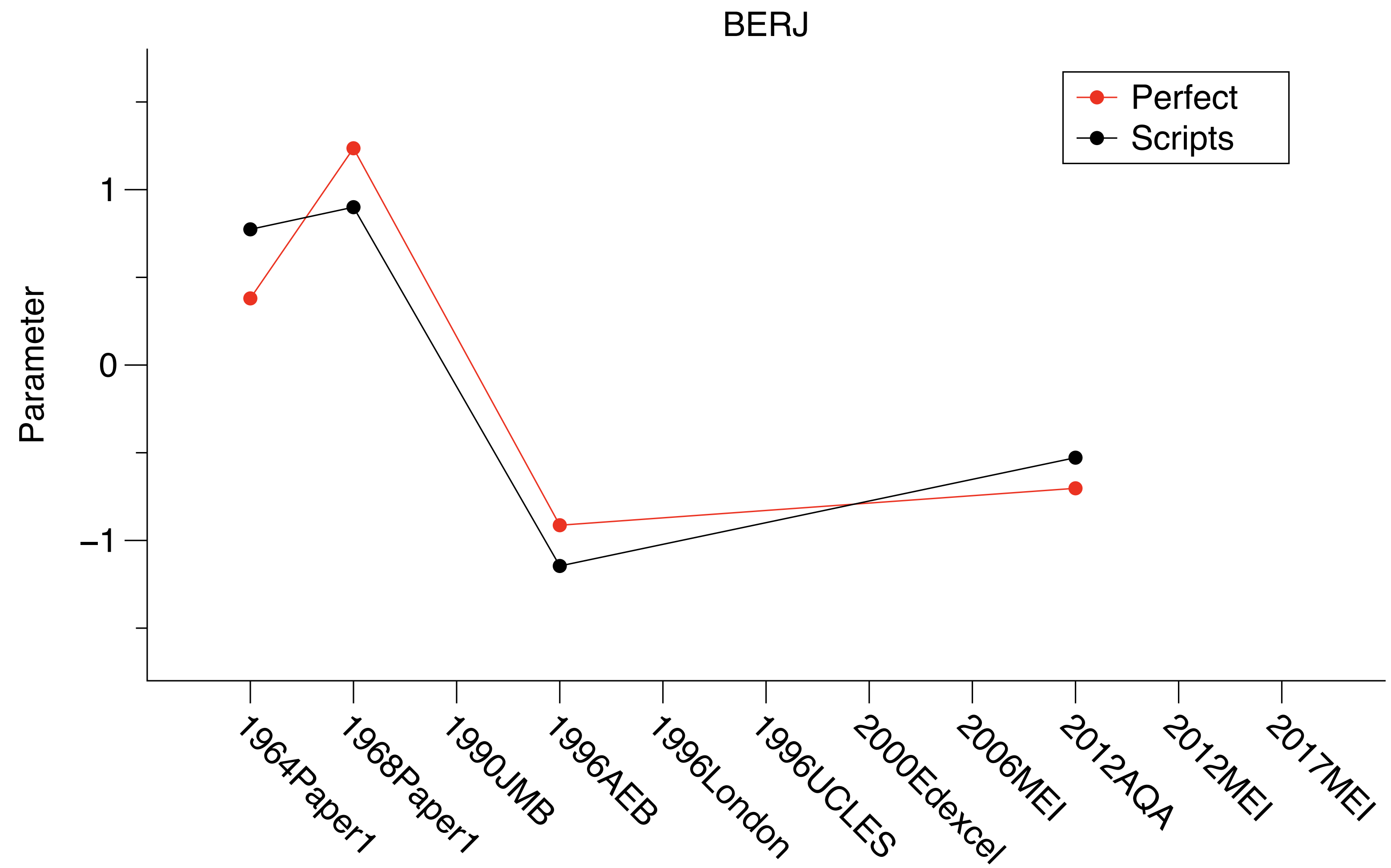
(ii) the scores of the real scripts from the BERJ paper (“script scores”).

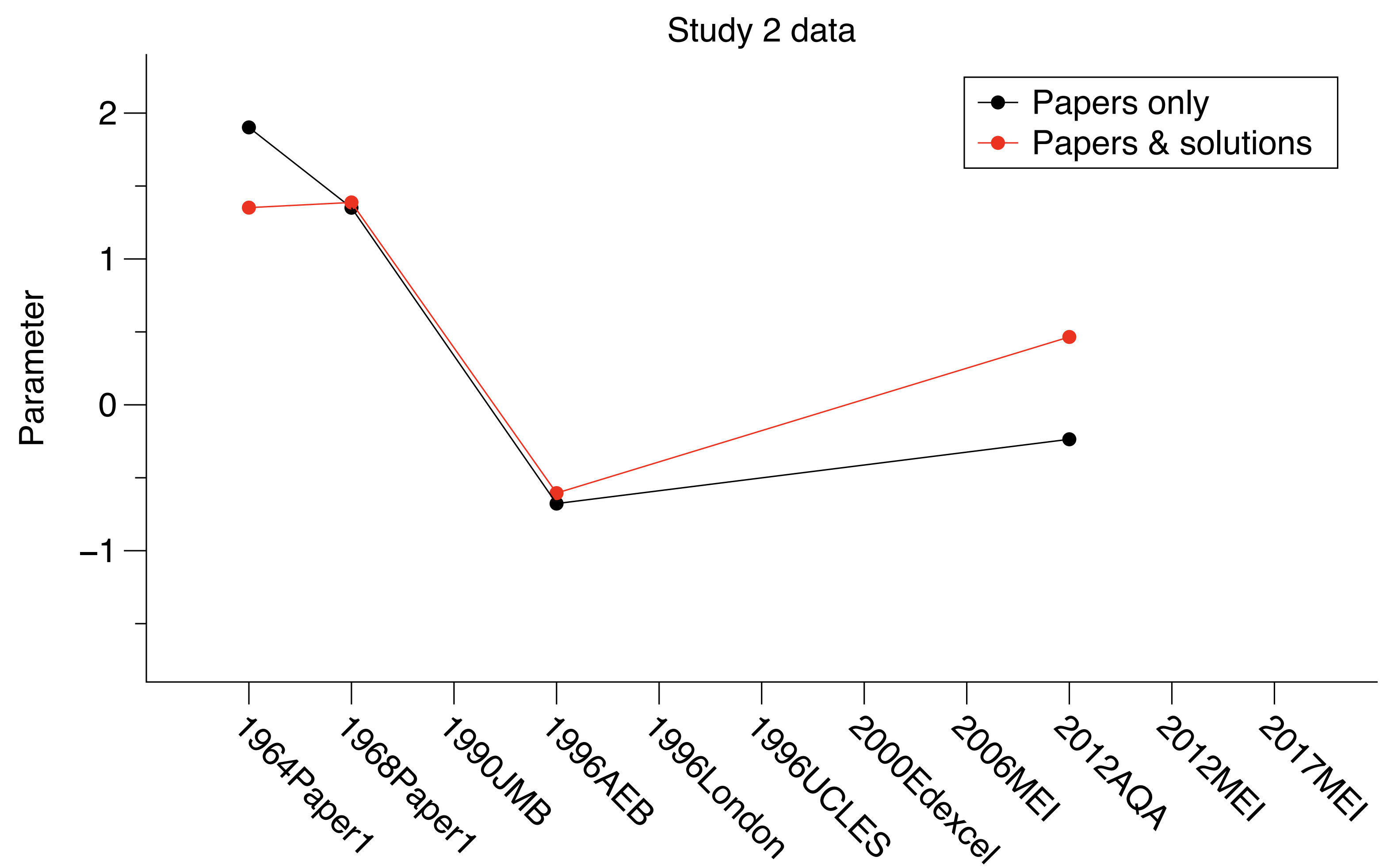
Unlike for Study 1 we did this graphically.

Study 2: graphical analysis

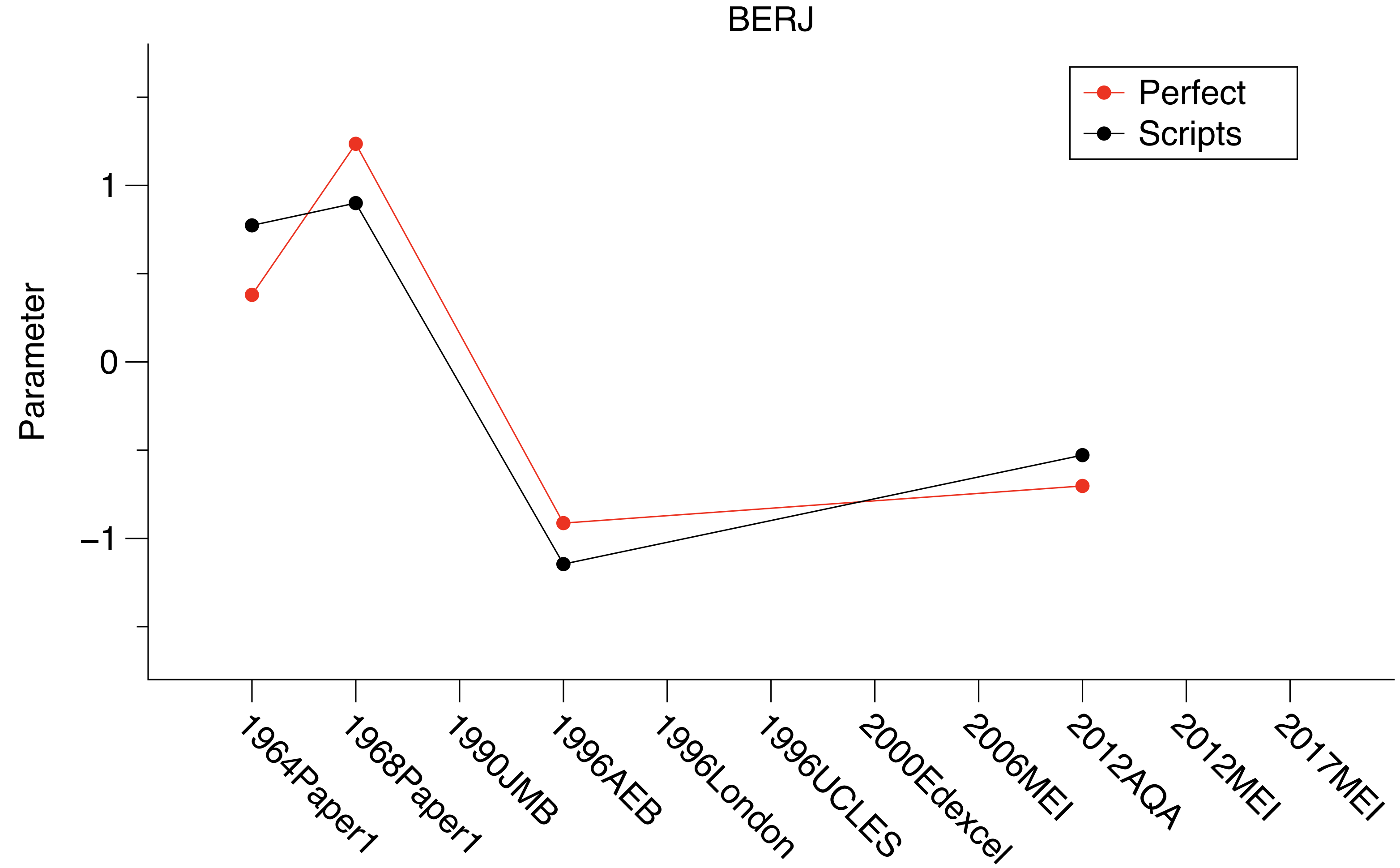


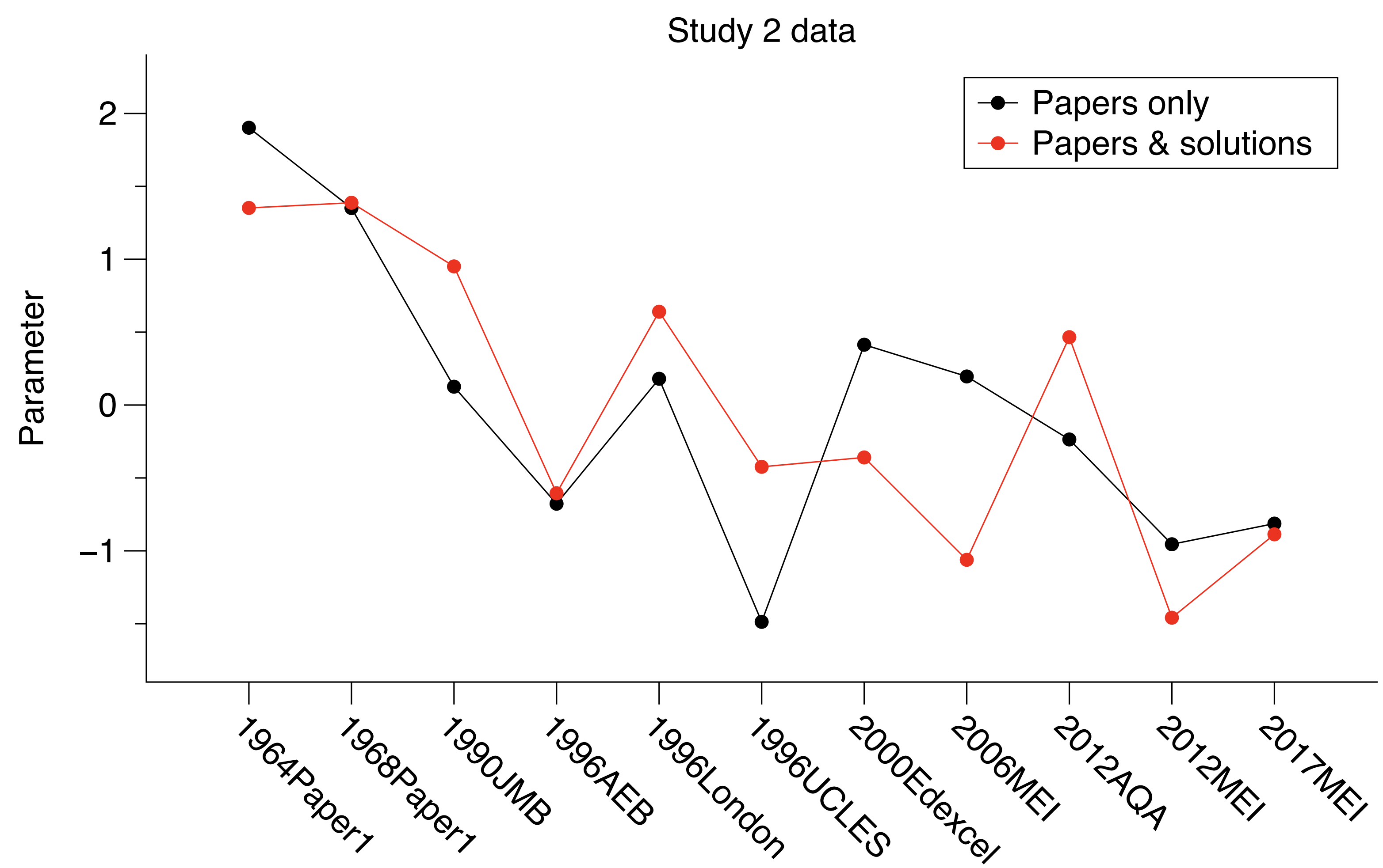
Study 2: graphical analysis





Study 2: graphical analysis





Limitations

- Standards-based assessment research is nonsense (Goldstein, 1979; Newton, 1997).
- Study 2 had only four data points. No estimate available that results are due to chance.
- Papers vary in length from 8 to 40 pages. CJ score vs length: $\rho = -.47$, $p = .15$.
- Cannot say “a candidate who achieved a grade B in 1996 or 2012 appears to have ... performed approximately at the level of a candidate who achieved a grade E in 1964”

Thank you

Ian Jones Loughborough University
i.jones@lboro.ac.uk

Colin Foster Loughborough University
c.foster@lboro.ac.uk

Jodie Hunter Massey University, New Zealand