Assessing Clinical Competence with a Computerised Vignette Matching Tasks: A comparison of Trainees in the UK & Australia

Dr Gavin Richardson
Prof Hamish J McLeod

Challenges in Assessing Trainee Competence

- "Accreditation standards tend to increase in number over time…however even when criteria are specified, expected competence *levels* often remain elusive to both supervisors and supervisees"

  Bernard, J. M., & Goodyear, R. K. (2009). Fundamentals of clinical supervision. Upper Saddle River, NJ: Pearson Education, Inc.

- Trainees learn to do more "stuff" across training but they also need to progress and do "stuff" that they previous learned but to a higher standard
- This introduces complexity to competence ratings using simple Likert scales: supervisors often adjust ratings to take account of stage of training
- Using Likert scales for competency ratings may measurement & error variance due to variations in anchoring, implicit assumptions, biases, and adjustments about developmental stage

Insights from Supervisor Training

- Feedback from NES supervisor training points to a need for early career supervisors to explicitly learn the performance benchmark that needs to be acquired and displayed by trainees
- The “calibration set” that early career supervisors directly acquire is commonly gained in one placement setting with one level of trainee.
- Hence, there may be considerable variance in the anchors and mental models of skill/competence that supervisors develop
- One solution is to make more explicit the information used in expertise based judgements that have been internalized by experienced supervisors

Sources of Competency Assessment Bias

- Leniency – a tendency for supervisors to under-use average and below average grades
- Halo – high correlations between competence domains that are rater specific (i.e. do not replicate with new raters in subsequent placements)

Using Vignettes of Clinical Competencies to Reduce Leniency and Halo Biases

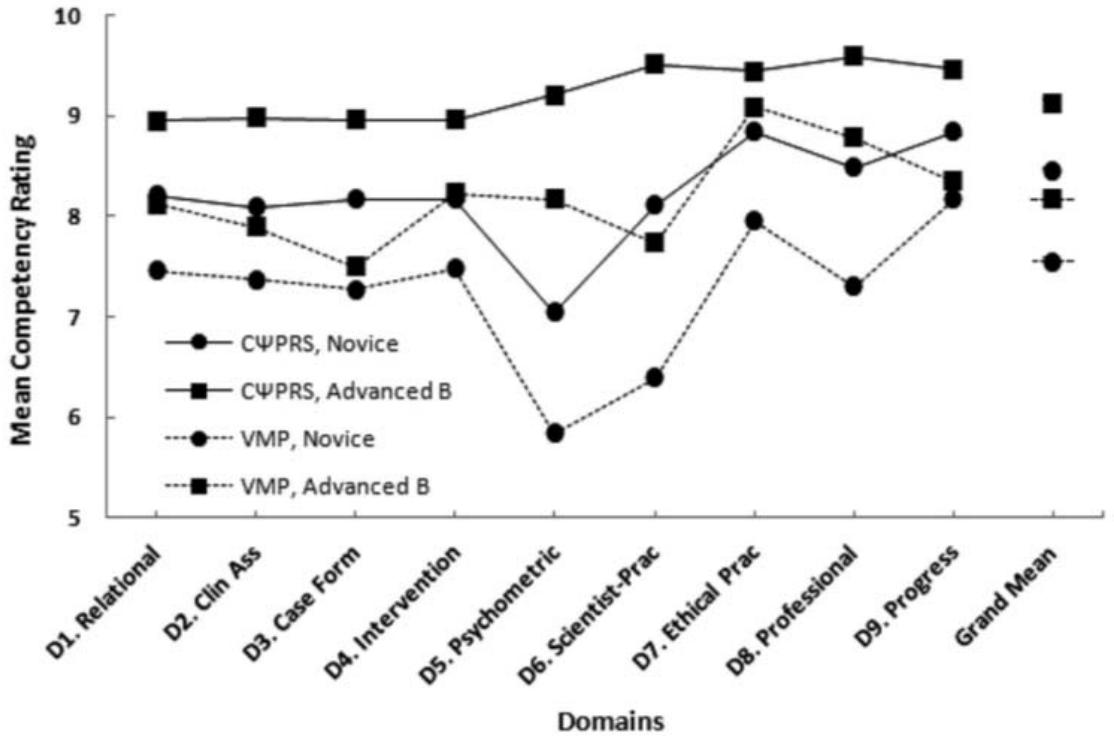


Figure 2. Mean domain scores obtained by novices and advanced beginners on the CΨPRS and the VMP. Clin Asst = Clinical Assessment; Case Form = Case formulation; Scientist-Prac = Scientist-Practitioner; Ethical Prac = Ethical Practice; Progress = Progress during placement; Advanced B = Advanced Beginner.

Gonsalvez et al., (2013). Assessment of psychology competencies in field placements: Standardized vignettes reduce rater bias. *Training and Education in Professional Psychology*, *42*(2), 99–111.

Vignette Matching as a Solution

- Clinical skills vignettes provide an alternative to Likert scales as a way of grading Trainee competence
- This approach reduces both leniency and halo biases
- But first, there is a need to obtain suitably calibrated vignettes that reflect the levels of competence needed to evaluate training progress and acquire rights to practice in clinical psychology

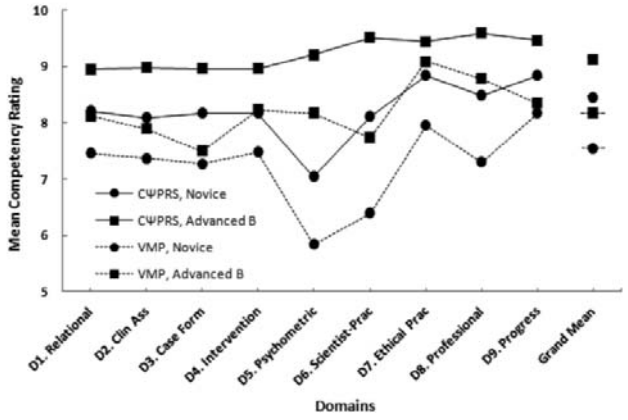


Figure 2. Mean domain scores obtained by novices and advanced beginners on the CΨPRS and the VMP. Clin Asst = Clinical Assessment; Case Form = Case formulation; Scientist-Prac = Scientist-Practitioner; Ethical Prac = Ethical Practice; Progress = Progress during placement; Advanced B = Advanced Beginner.

Scientist Practitioner Competencies

Please indicate whether at end-placement your trainee attained a developmental stage that is higher, equal to, or lower than that depicted in each vignette.

Trainee FL consistently demonstrates a commitment to bringing the scientific method to her/his clinical work. She/he uses a systematic hypothesis generation and testing approach in her/his work with clients, appropriately seeking information through interview, observation or psychometric testing to test her/his clinical formulations. She/he seeks evidence of reliability and validity in making decisions about which assessment methods (e.g., tests) to use. She/he routinely accesses scholarly scientific resources (e.g., journals) to guide decisions about the most effective treatments to use. When research is lacking or unclear regarding the best treatment approach, she/he shows the ability to tailor a treatment program for the client based on an analysis of the available evidence or scientific principles. She/he values and usually implements systematic assessment of treatment progress in clients.

My trainee has reached a developmental stage that is:

Higher than depicted in this vignette

Equal to that depicted in this vignette

Lower than depicted in this vignette

The Vignette Matching Assessment Task
www.westernsydney.edu.au/vmp/about#vmat

>>

Aims of the Vignette Matching Assessment Task Project

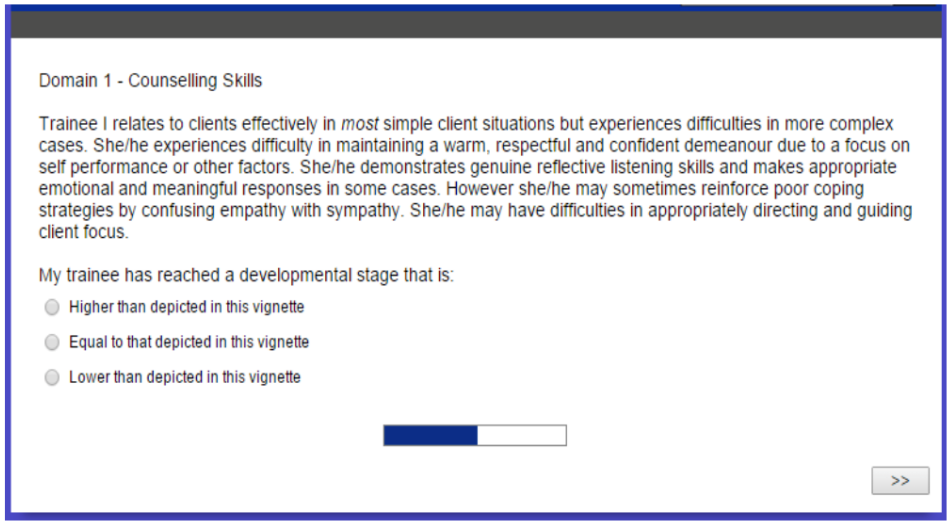
- Develop and test a comprehensive set of clinical skills and knowledge vignettes
- Compare expected competence levels for trainees across training jurisdictions (U, GB, N)
- Compare vignette matching to standard clinical skills ratings for both measurement performance and user-acceptability

Vignette Matching Assessment Task

Supervised field placements are essential in the training and credentialing of psychologists and other clinical practitioners. However, recent research indicates that supervisor assessments are influenced by leniency and halo rating biases, and multiple attempts to remedy this through better definition of the anchors used in the Likert scaling and by additional training to supervisors, have proven largely ineffective.

A Novel Approach to Trainee Assessment

The V-MAT takes an alternative approach to assessing trainee competence. Rather than using a Likert-scale, it presents detailed vignettes of student competencies that are expected at various training stages. A vignettes is a capsule summary of about 100 words (see below for an example) that profiles essential features of competencies a trainee possesses and does not possess in reference to a specific domain (e.g. counselling skills). Supervisors read the vignettes and match their trainees' performance with the vignettes in each of the domains in order to determine if they have acquired the level of competence expected at their stage of training.



https://www.westernsydney.edu.au/vmp/about

Participants

- 59 senior clinicians and/or trainers (37 females, 22 males)
- 45 = U; 14 = GB
- All registered with the Psychology Board of Australia or the UK's Health and Care Professions Council (HCPC)
- Eligibility for inclusion:
  - (1) be course directors or deputy course directors, practicum coordinators, training clinic directors or deputy directors, ***or***
  - (2) have 10 or more years' experience as clinical supervisors and five or more years' experience in evaluating the practicum competencies of students.
- Sample characteristics :
  - Years of practice: *X* = 20.29 years (*SD* = 7.50, Range = 5 - 35+ years)
  - Years of supervision experience: *X* = 15.10 years (SD = 7.78, Range = 2 – 33 years)
  - Years of clinical trainee supervision experience: *X* = 12.49 years (SD = 7.94, Range = 1 – 33 years)
  - UK and Australian rater samples were equivalent across all characteristics

Methods: Operationalisation of Developmental Stages

Stage 1: Beginner	Knowledge and skills are at an early stage or yet to be developed. Inadequate knowledge and/or difficulty applying knowledge to practice. Several problems or inadequacies occur during sessions. There may be an absence of key features, inability to prioritise issues or to make appropriate judgements. Little awareness of process issues. On par with trainees commencing training without any practicum experience. Regular and intensive supervision required.
Stage 2	Some basic competencies in assessment and intervention, manages narrow range of clients with low levels of severity, using structured therapeutic activities. Performance is variable; major problems may occur occasionally; regular supervision required.
Stage 3	Moderate repertoire of basic competencies in both assessment and intervention leading to management of a wider range of clients. Demonstrates understanding of underlying principles and a moderate ability to generalise these to new cases/situations. Performance can be improved in minor ways; less frequent supervision required
Stage 4: Competent	Large repertoire of basic to advanced competencies in both assessment and intervention, applied across range of clients and severity levels. Performance has reached competency levels on a par with a clinical psychology working in their first job upon qualification.

Methods: Assessed Competence Domains

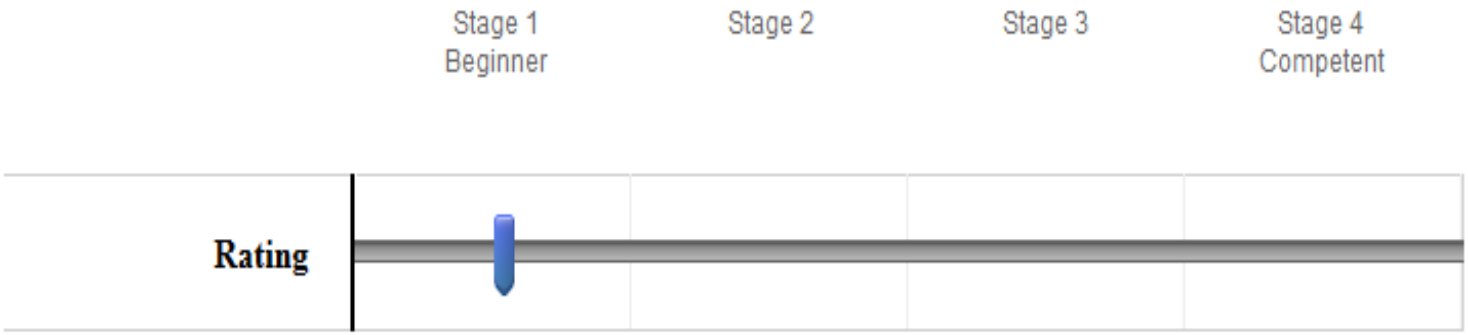
1. *Relational Skills*
2. *Clinical Assessment*
3. *Case Formulation*
4. *Generic Intervention*
5. *Cognitive-Behavioural Intervention*
6. *Psychometrics*
7. *Scientist-Practitioner Approach*
8. *Ethical Behaviour*
9. *Professional Skills*

Vignette Development

- Phase 1: Vignette templates for 4 developmental stages were generated from a standardised competency assessment scale (the CΨPRS).
- Phase 2: The initial set of 82 initial vignettes was reviewed by 4 experts and condensed to a set of one vignette per level for each domain
- Phase 3: Anonymous peer commentary was obtained for all vignettes
- Phase 4: Core project team incorporated feedback
- Phase 5: The vignettes underwent preliminary calibration and unreliable items were discarded
- Phase 6: This study

Rater Calibration Judgements

Trainee TJ demonstrates a good knowledge of rationales for and good skills to conduct a fairly large range of CBT techniques. She/he efficiently identifies unhelpful cognitions and beliefs, and poses relevant socratic questions after appropriate preparation. A collaborative style ensures that modest gains within and across sessions are typically achieved. When this does not occur it is because of client resistance or because an implicit or more subtle belief was not targeted for change. Despite demonstrating good CBT skills, aspects that could improve include fluency, timing, and improved consolidation of high impact moments during sessions.



Results

We examined:

- Differentiation of levels of competence depicted in the vignettes
- Cross-national differences in what constitutes each level of competency
- Patterns of competency progression across the 9 assessed domains

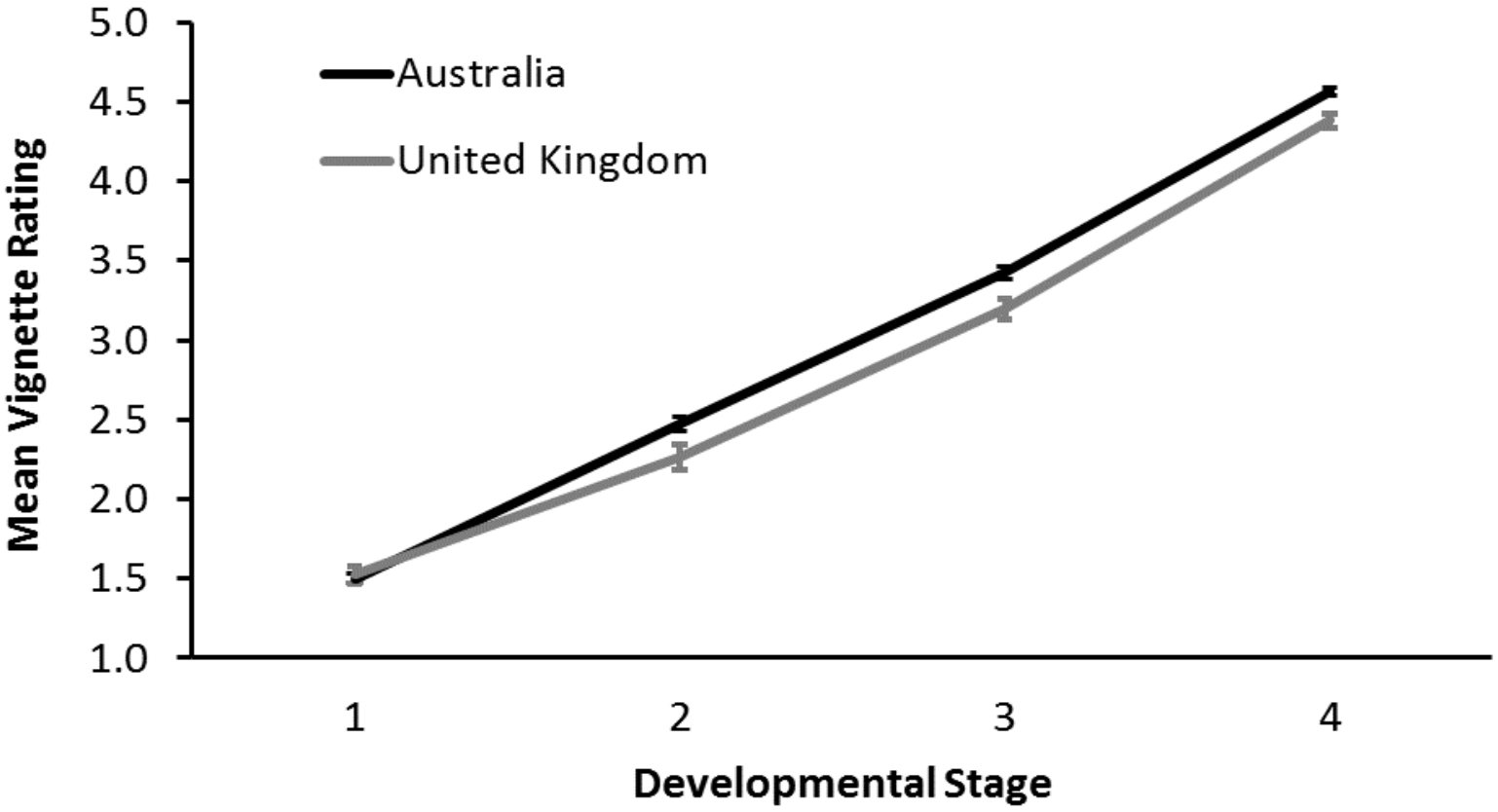
Summary of Main Effects Analysis

- There was a significant main effect of Developmental Stage [$F(2.36, 134.569) = 2108.120$, $p = .00$, $\eta^2 = .97$].
  - Pairwise comparisons with a Bonferroni adjustment for multiple comparisons ($p < .008$) revealed that all stages were rated differently to one another ($ps < .008$).
  - These differences reflected a linear increase from Stage 1: Beginner to Stage 4: Competent [$F(1, 57) = 4337.68$, $p = .00$, $\eta^2 = .99$].
- There was a significant main effect of Country [$F(1, 57) = 7.30$, $p = .01$, $\eta^2 = .11$]
  - Ratings provided by Australian calibrators ($M = 2.99$, $SD = 1.21$) were significantly higher than those provided by UK calibrators ($M = 2.84$, $SD = 1.16$)
- There was a main effect of Domain [$F(6.28, 358.15) = 13.52$, $p = .00$, $\eta^2 = .19$].
  - Bonferroni adjustment for multiple pairwise comparisons ($p < .001$) showed that the mean rating for D7: *Scientist Practitioner Approach* was significantly higher than all other domains ($ps < .001$). In addition, the mean rating for D5: *Cognitive-Behavioural Intervention* was significantly higher than for D2: *Clinical Assessment* ($p < .001$).

Interactions

- Country x Developmental Stage [$F(2.36, 134.57) = 4.43$, $p = .01$, $\eta^2 = .07$]
- Domain x Developmental Stage [$F(14.90, 849.47) = 19.16$, $p = .00$, $\eta^2 = .25$]
- Country x Domain x Developmental Stage [$F(14.90, 849.47) = 1.74$, $p = .04$, $\eta^2 = .03$]

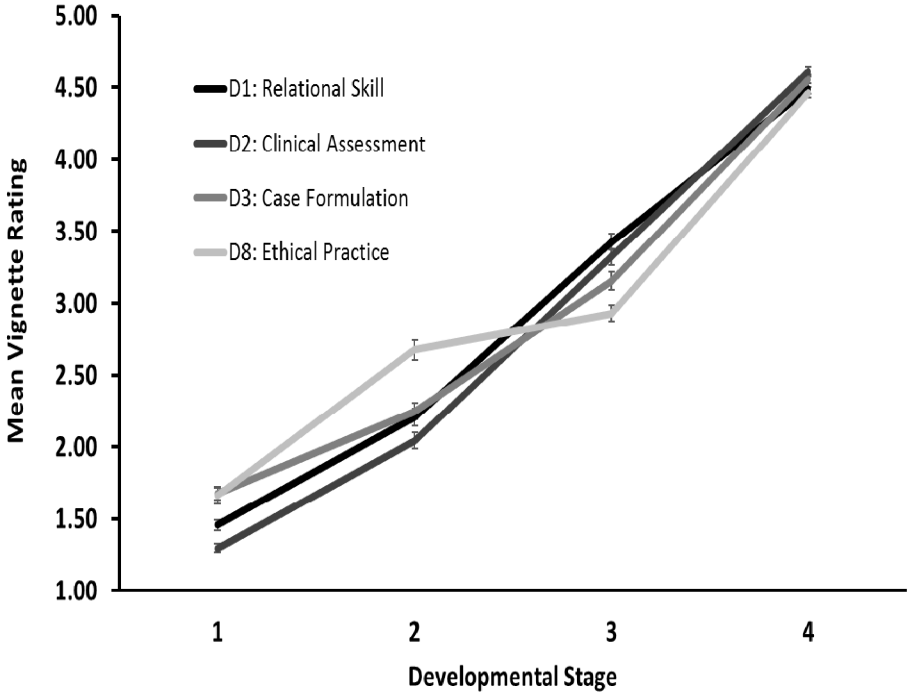
Calibration Ratings: U vs GB



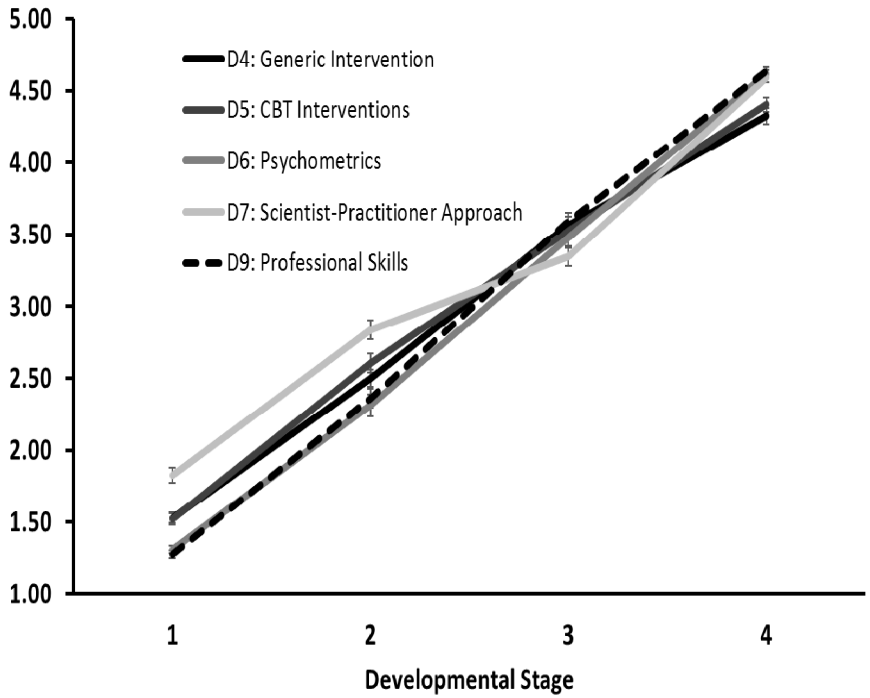
Development Trajectories Across Competence Domains

- The linear and quadratic trends evident in the Country x Developmental Stage interaction were present in each of the domains. But, the results revealed no differences in either linear or quadratic trends as a function of Country, but there were significant trends evident across Developmental Stages (i.e. main effect).
- With Bonferroni adjustment for multiple comparisons ($p < .0055$) there were significant linear increases across developmental stages for all domains and four domains showed quadratic trends:
  - *D1: Relational Skills*
  - *D2: Clinical Assessment*
  - *D3: Case Formulation*
  - *D8: Ethical Behaviour*.

Differences in Expected Rate of Competence Development



Quadratic Trends



Linear/Non-Quadratic Trends

Interpretation and Conclusions

- Expert raters show high degrees of agreement for behavioural descriptions of different levels of trainee competence
- Expected competence development curves vary as a function of the domain being assessed
- The perceived threshold for qualification level competency varies slightly across training jurisdictions

Thank you

gavinrichardson@nhs.net
hamish.mcleod@glasgow.ac.uk