

Chapter 6: 21st Century Evaluation

The concept of validity has assumed a pivotal role across decades of debate on the characteristics of quality in educational and psychological measurement. Despite this, it has proved extremely resistant to definition. In Chapter 6, we respond to the concerns of the more conservatively minded, who object that the concept of validity is becoming so large as to present an obstacle to validation practice. We do so by proposing a new framework for the evaluation of testing policy. In fact, we see this as a revision of the original progressive matrix from Messick (1980), which we have redesigned to dispel some of the confusion engendered by its original presentation. After first defending the new framework we then provide a more detailed analysis of technical and social evaluation, before considering evaluation within each of the cells respectively.

Validity in Educational and Psychological Assessment will be available from March 2014. The authors believe that this book will be of interest to anyone with a professional or academic interest in evaluating the quality of educational or psychological assessments, measurements and diagnoses.

ISBNs: Paperback: 9781446253236 Hardback: 9781446253229

References

- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1999). *Standards for Educational and Psychological Testing*. Washington, D.C.: American Educational Research Association.
- American Psychological Association, American Educational Research Association, and National Council on Measurements Used in Education. (1954). Technical recommendations for psychological tests and diagnostic techniques. *Psychological Bulletin*, **51**, 2, Supplement.
- Anastasi, A. (1950). The concept of validity in the interpretation of test scores. *Educational & Psychological Measurement*, **10**, 1, 67–78.
- Brennan, R. L. (Ed.). (2006). *Educational Measurement* (4th ed.). Westport, CT: Praeger Publishers.
- Cronbach, L.J. (1971). Test validation. In R.L. Thorndike (Ed.). *Educational Measurement* (2nd Edition) (pp.443–507). American Council on Education. Washington: DC.
- Cronbach, L.J. & Meehl, P.E (1955) Construct validity in psychological tests. *Psychological Bulletin*, **52**, 4, 281–302.
- Fast, E.F. & Hebbler, S. with ASR-CAS Joint Study Group on Validity in Accountability Systems. (2004). *A Framework for Examining Validity in State Accountability Systems*. Washington, DC: Council of Chief State School Officers.
- Geisinger, K.F. (1992). The metamorphosis of test validation. *Educational Psychologist*, **27**, 2, 197–222.
- Goldstein, J. & Behuniak, P. (2011). Assumptions in alternate assessment: An argument-based approach to validation. *Assessment for Effective Intervention*, **36**, 3, 179–191.
- Kane, M.T. (1992). An argument-based approach to validity. *Psychological Bulletin*, **112**, 3, 527–535.
- Kane, M.T. (2001). Current concerns in validity theory. *Journal of Educational Measurement*, **38**, 4, 319–342.
- Lindquist, E.F. (Ed.) (1951). *Educational Measurement*. Washington, DC: American Council on Education.
- Messick, S. (1980). Test validity and the ethics of assessment. *American Psychologist*, **35**, 11, 1012–1027.
- Moss, P. A., Girard, B. J. & Haniford, L. C. (2006). Validity in Educational Assessment. *Review of Research in Education*, **30**, 1, 109–162.
- Newton, P.E. & Shaw, S.D. (2013). Standards for talking and thinking about validity. *Psychological Methods*, **18**, 3, 301–319.
- Shepard, L.A. (1993). Evaluating test validity. *Review of Research in Education*, **19**, 405–450.
- Shepard, L.A. (1997). The centrality of test use and consequences for test validity. *Educational Measurement: Issues and Practice*, **16**, 2, 5–8.

Research News

Karen Barden Research Division

Conferences and seminars

The Future of Education International Conference

In June, Sanjana Mehta attended The Future of Education Conference in Florence, Italy. The conference aims to promote transnational cooperation and share good practice in the field of innovation for education. Sanjana presented a paper on *Thrown in at the deep end? Exploring students', lecturers' and teachers' views on additional support lessons at university*.

The Assessment in Higher Education Conference

Held in Birmingham in June, this fourth biennial conference provided an opportunity to debate key issues and developments in current assessment, policy and practice. Simon Child presented a paper entitled *"I've never done one of these before". A comparison of the assessment 'diet' at A level and the first year of university*.

British Education Studies Association (BESA)

The ninth BESA Annual Conference took place at Swansea Metropolitan University in June. The key theme of the conference was Education: Past, Present and Future. Jackie Greatorex presented on *Using scales of cognitive demand in a validation study of Cambridge International A and AS level Economics*.

Journal of Vocational Education and Training (JVET)

The JVET tenth international conference was held in July at Worcester College, Oxford. Colleagues from the Research Division presented the following papers:

Jackie Greatorex: *How can major research findings about returns to qualifications illuminate the comparability of qualifications?*

Martin Johnson: *Insights into contextualised learning: how does feedback on performance contribute to professional examiners' shared understanding?*

Carmen Vidal Rodeiro: *Do Cambridge Nationals support progression to further studies at school or college, to higher education courses and to work-based learning?*

International Meeting of the Psychometric Society (IMPS)

Tom Benton presented a paper at the 78th Annual Meeting of the Psychometric Society in Arnhem, The Netherlands, in July on *An empirical assessment of the Guttman's Lambda 4 reliability coefficient*.

European Association for Research on Learning and Instruction (EARLI)

Tom Bramley presented a paper at the 15th Biennial EARLI Conference held in Munich, Germany on *Maintaining standards in public examinations: why it is impossible to please everyone*.

British Educational Research Association (BERA)

The BERA Annual Conference was held from 3–5 September at the University of Sussex. Colleagues from the Research Division and OCR presented the following papers:

Tom Benton: *Examining the impact of tiered examinations on the aspirations of young people*.

Jackie Greatorex: *How can major research findings about returns to qualifications illuminate the comparability of qualifications?*

Simon Child, Sanjana Mehta, Frances Wilson, Irenka Suto and Sally Brown: *Beyond freshers' week: An exploration of universities' assessment practices and additional support lessons*.

Sanjana Mehta, Martin Johnson, Nicky Ruston and Simon Child: *Controlled Assessment and Modern Foreign Language (MFL) speaking and listening: using a mixed methods approach to evaluate the effects of assessment arrangements*.

Magda Werno: *Addressing the needs of non-native speakers entering English education at the secondary stage*.

Tom Benton of the Research Division won the BERA poster competition with his entry entitled *Calculating the reliability of complex qualifications*. Tom was helped with the poster design by Research Division colleague Jo Ireland.

The following posters were also presented:

Nicky Rushton: *Changing times, Changing qualifications*.

Jackie Greatorex, Stuart Shaw, Phineas Hodson and Jo Ireland: *Do the A and AS level Economics examination papers elicit responses that reflect the intended construct?*

European Conference on Educational Research (ECER)

In September, Frances Wilson and Nicky Rushton attended the ECER conference in Istanbul, Turkey. The main theme of the conference was *Creativity and Innovation in Educational Research*. Frances presented a paper entitled *A comparison of assessment at school and university: More than just increasing demand*. Nicky gave a poster presentation on *The Register of Change*.

European Conference on Computer Supported Cooperative Work (ECSCW)

Martin Johnson presented a paper at the 13th European Conference on Computer Supported Cooperative Work in Paphos, Cyprus in September on *'Seeing what they say': mapping the characteristics of effective remote feedback*.

European Association of Test Publishers (E-ATP)

The E-ATP Annual Conference took place in September in St Julian's, Malta. One of the key themes was *Innovations in Assessment* and Irenka Suto presented a paper entitled *The Cambridge Approach to 21st Century skills: advances in teaching, assessment and support for learners*.

Association for Educational Assessment – Europe (AEA-Europe)

The AEA-Europe Annual Conference took place in Paris in November with the theme of *International surveys, policy borrowing and national assessment*. Several colleagues from Cambridge Assessment attended the conference and the following papers were presented:

Tom Benton and Tim Gill: *Investigating the relationship between aspects of countries' assessment systems and achievement in PISA*.

Victoria Crisp: *The judgement processes involved in the moderation of teacher-assessed projects in a national assessment*.

Gill Elliott: *Method in our madness? The advantages and limitations of mapping other jurisdictions' educational policy and practice*.

Tim Oates: *Why Grand Theory and detailed narrative are equally essential in drawing from transnational comparisons*.

Stuart Shaw, Martin Johnson and Paul Warwick: *Assessment for learning in international contexts: approaches and challenges in researching teacher values and practices (Phase 2)*.

Further information on all conference papers can be found on the Cambridge Assessment website: <http://www.cambridgeassessment.org.uk/our-research/all-published-resources/conference-papers/>

Publications

The following articles have been published since Issue 16 of *Research Matters*:

Crisp, V. & Grayson, R. (2011) Modelling question difficulty in an A level Physics examination. *Research Papers in Education: Policy and Practice*, **28**, 3, 346–372.

Greatorex, J. and Shaw, S. D. (2013). The validity of tutor assessed Independent Research Reports contributing to a pre-university qualification *Educational futures*, BESA, **6**, 1, 58–78. Available online at: http://educationstudies.org.uk/?post_type=journal&p=1113

Johnson, M (2013). 'Seeing what they say': mapping the characteristics of effective remote feedback, In: Korn, M., Colombino, T., and Lewkowicz, M. (Eds) *The 13th European Conference on Computer Supported Cooperative Work Adjunct Proceedings*. Aarhus: Aarhus University, pp. 141–147. Available online at: <http://ojs.statsbiblioteket.dk/index.php/daimipb/article/view/13587/11586>

Keating, A. & Benton, T. (2013). Creating cohesive citizens in England? Exploring the role of diversity, deprivation and democratic climate at school. *Education, Citizenship and Social Justice*, **8**, 2, 165–184. Available online at: <http://esj.sagepub.com/content/8/2/165.abstract>

Mehta, S., Child, S., Brown, S. & Suto, I. (2013). Thrown in at the deep end? Exploring students', lecturers' and teachers' views on additional support lessons at university. *Future of Education Conference Proceedings 2013*. Available online at: http://conference.pixel-online.net/foe2013/common/download/Paper_pdf/132-SOE38-FP-Mehta-FOE2013.pdf

Shaw, S. D. and Imam, H. (2013) Assessment of International Students Through the Medium of English: Ensuring Validity and Fairness in Content-Based Examinations. *Language Assessment Quarterly*, **10**, 4, 452–475.