

Workshop 1: An introduction to multilevel models in education

Harvey Goldstein (Bristol University, UK)

The workshop will introduce the idea of multilevel modelling through simple illustrations and worked examples using educational data sets. A minimal level of statistical background equivalent to a working knowledge of multiple regression will be assumed.

Topics to be covered will include:

- Simple variance component models
- The ‘value added’ model for estimating school effects
- Analysis of data containing ‘contextual’ effects on achievement
- Modelling binary data
- More complex classifications (crossing and multiple membership)
- Modelling segregation and diversity
- A detailed example analysing the effects of school mobility on achievement using a very large English longitudinal data set (PLASC).

Demonstrations will be given of running data analyses.

Workshop 2: Applications of Item Response Theory (with special attention to applications in examinations)

Frans Kleintjes and Anton Béguin (CITO, The Netherlands)

Item Response Theory (IRT) is used to analyse test data on item level. IRT provides a statistical process for estimating characteristics of items and examinees and defining how these interact in describing item and test performance. When used properly IRT can increase the efficiency of the testing process, enhance the information provided and make detailed predictions about unobserved testing situations.

IRT is used for many measurement applications including item banking, test construction, adaptive test administration, scaling, linking and equating, standard setting, test scoring and score reporting. In the workshop, the emphasis will be on using IRT in examinations.

The workshop starts with some theory on IRT, including a comparison with classical test theory. Major properties of IRT will be highlighted using illustrative examples. IRT output will be explained, discussed and interpreted based on materials that will be provided. Several concepts used in IRT will be explained using examples from the test construction experience of the presenters and if available from participants.

The application of IRT in the following will be covered:

- Information function and its use in test construction
- Content related interpretation of the ability scale.
- Standard setting and Maintaining standards in Examinations.

Workshop 3: Validation **(with an emphasis on using examination or test results)** *Christina Wikstrom and Simon Wolming (Umeå University, Sweden)*

Validity theory is one of the fundamental cornerstones of educational measurement. The purpose of this workshop is to make the participants familiar with the concept of validity, what it involves and how validity evidence typically is gathered for different types of instruments and assessment outcomes.

The workshop is open for AEA-Europe members and AEA-Europe conference participants who want to learn about the concept of validity and who are, or will be, conducting different types of validity studies in their professions as test developers or test users.

The first part of the workshop will mainly be theoretical, and presentations will be made of the development of the concept of validity, from traditional approaches to the modern view on validity, but also how the concept has been challenged and debated,

foremost in recent years. Various sources such as articles and books will be used as references during the presentation, but emphasis will be on how validity is described in the chapters on validity in *Educational Measurement and The Standards for Educational and Psychological Testing*.

The second part of the workshop will be more practical, and focus especially on describing and discussing different types of analyses that could be suitable for validating the outcomes of different types of assessment instruments and systems. Participants will be provided with typical examples of validation problems but also encouraged to contribute with own examples.

Workshop 4: A Beginner's Guide to Educational Assessment

Chris Whetton, Marian Sainsbury (NFER) and Steven Bakker (Dutchtest)

The workshop is aimed at participants who would like to have an understanding of the scale and scope of educational assessment. It will not go into any topic in detail, but will give a broad introduction across a wide range of topics.

Tests and examinations are an important tool in education. They can provide stakeholders such as students, parents, teachers, members of boards, Ministries of Education and society generally with all kind of information regarding the development of learners, the performances of schools, the quality of education on a national level.

The workshop will cover the purposes of assessment, what measurements need in order to be reliable and valid (and what this means), how assessments are developed and improved, the role of teachers in assessment, how tests and examinations differ, what results mean and how they can be used and finally how assessment systems can be set up.

Practical exercises and issues for discussion will be presented throughout the workshop to ensure the active involvement of participants.