Key Concepts in Assessment


**Reliability** – “refers to the extent to which the scores produced by a test or an assessment are consistent, dependable and replicable.” (p.123) “The extent to which an assessment task is deemed reliable is indicated by whether or not it would produce the same or very similar scores for the same students at different times or different places and regardless of who is marking the assessment.” (p.124)

**Validity** – “refers to accuracy in assessment design – the extent to which any assessment measures what it has been designed to measure.” (p.135) “Bias can act as a source of invalidity if, for example, an assessor makes incorrect assumptions about the linguistic ability of candidates . . . [e.g] an assessment aimed at assessing numeracy that contains ‘wordy’ problem-style questions that students could answer if they were expressed using mathematical symbols instead of verbally.” (p.139)

**Norm-Referencing** – “An educational assessment procedure can be identified as norm-referenced when the score that an individual achieves is converted into a statement or grade indicating how the individual compares with others who have undergone the same assessment.” (p.97) “A sub-set of norm referencing is cohort referencing. This is where a normative-type distribution of scores is applied only to the cohort in question and thus the actual pass grade might vary from cohort to cohort. An example would be the former 11+ exam taken by pupils in the UK to sort those going from grammar schools from those going to technical or secondary modern schools.” (p.98)

**Criterion-Referencing** – “The individual’s understanding of specific subject matter is evaluated against pre-specified criteria. The outcome is judged on the student’s own performance and not with reference to performance by others.” (p.40)

**Norm- and Criterion-Referencing** - There are examples where the scoring system is both norm- and criterion-referenced . . . current A Levels, which changed from largely norm-referenced (or cohort-referenced) to criterion-referenced but where the final attribution of grades is still influenced by normative expectations.” (p.41)