The pearls and pitfalls of setting High Quality MCQs
Prof Mergan Naidoo
AL: T&L: SNPH

Outcomes for the workshop

- Provide evidence of using the MCQ assessments as a reliable and valid method
- Demonstrate competency in blueprinting an examination
- Develop skills in writing good quality SBAs/EMQs
- Develop skills in standard setting
- Develop skills in using working with psychometrics

<table>
<thead>
<tr>
<th>Key issues that underpin any test</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summative/formative</td>
<td>Be clear on the purpose of the test.</td>
</tr>
<tr>
<td>Blueprinting</td>
<td>Plan the test against the learning objectives of the course or competencies essential to the speciality.</td>
</tr>
<tr>
<td>Validity</td>
<td>Select appropriate test formats for the competences to be tested. This action invariably results in a composite examination.</td>
</tr>
<tr>
<td>Reliability</td>
<td>Sample adequately. Clinical competencies are inconsistent across different tasks. Test length is crucial if high-stakes decisions are required. Use as many examiners as possible.</td>
</tr>
<tr>
<td>Standard setting</td>
<td>Define endpoint of assessment. Set the appropriate standard—eg, minimum competence—in advance.</td>
</tr>
</tbody>
</table>

5. The Policy

5.1 Assessments in the University are underpinned by the principles set out in the Policy on Teaching and Learning.

5.2 Assessments must draw on criteria that are clearly related to the purpose and outcomes of the curriculum by ensuring coherence between assessment criteria, the purpose, outcomes and assessment methods.

5.3 Monitoring, reviewing and improving assessment in practice must be incorporated into all programme and School review processes.

5.4 Valid and reliable assessments must be included as an integral part of the teaching/learning cycle.

5.5 Assessments must enhance engagement with the learning task and be aligned to the pedagogy and teaching methods.

5.6 The standards of UKZN qualifications must be maintained by ensuring that assessment is appropriate to and fits the specified module outcomes and the exit level outcomes of the programme, including the generic and disciplinary outcomes expected of a tertiary level qualification.

5.7 In adhering to principles of best practice, assessments must be transparent, consistent, practical, fair and feasible.

5.8 Academics are accountable for the quality of the assessments they implement and must be able to explain and justify their assessment judgements to students, examiners or any stakeholders.

---

Summative assessment:

- Is an end point examination
- Can block intended career progression (high stakes)
- Is perceived as threatening

Formative Assessment:

- Breaks learning into manageable modules
- Allow repeated attempts to master the each module
- Is not perceived as threatening (low stakes)

---

Competence

The ability to handle a complex professional task by integrating the relevant cognitive, psychomotor and affective skills
Miller’s model of competence

- Does
- Shows how
- Knows how
- Knows

Professional behaviours

Bloom’s taxonomy of knowledge

- Evaluation
- Synthesis
- Analysis
- Application
- Comprehension
- Knowledge

Bloom’s taxonomy

- Recall questions: * are thought to test examinees’ knowledge of related facts.
- Interpretation questions: * require examinees to review information (often in tabular or graphical form) and reach a conclusion (e.g., a diagnosis).
- Problem-solving questions: * present a situation and require examinees to take a course of action (e.g., the next step in current management).
- Memory questions: * ask for simple mastery of the knowledge base, similar to recall.
- Comprehension questions: * ask for demonstration of understanding of ideas and phenomena.
- Reasoning questions: * test the examinees’ thinking processes when faced with information.

Professional Authenticity

- Define, describe
- Interpret, discuss
- Analyse, order
- Evaluate, discriminate
- Synthesis, integrate, design
- Application, demonstrate
- Professional Authenticity

UKZN INSPIRING GREATNESS
**Reliability**: A measure of the reproducibility or consistency of a test.

“Sample adequately. Clinical competencies are inconsistent across different tasks. Test length is crucial if high-stakes decisions are required. Use as many examiners as possible.”

---

**Reliability: Consistency**

1. Test-retest
2. Internal consistency
3. Parallel forms
4. Split forms
5. Inter-rater reliability

Aim for value ≥ 0.7

---

**Reliability measures**

- Coefficient alpha
- Kruder Richardson 20: KR 20
- Standard error of measurement: SEM is the difference between a person's actual ability and the test score they achieve
Validity

Has the test measured what it set out to measure

“Select appropriate test formats for the competencies to be tested. This action invariably results in a composite examination”

Validity: Truthfulness

Content validity
Construct validity
Criterion–related validity
Valid and Reliable Assesments

- **Reliable**: dependable, repeatable, consistent
- **Valid**: measures appropriate knowledge and skills

Figure 1: Reliable but not Valid
Figure 2: Not Reliable, not Valid
Figure 3: Reliable and Valid


Limpopo students who walked out during exam could face suspension, says VC
Clinical competence is:
- Complex
- Highly integrated
- Content and situation dependent

Assessment requires:
- Quantitative and qualitative information
- Professional judgment
- From different sources

You need:
- Adequate sampling
- Different judges
- A range of contexts
Acknowledgements

- Prof Val Wass and the RCGP