Slide 1

UNIVERSITY OF
KWAZULU-NATAL ™
INYUVESI
YAKWAZULU-NATALI

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

EDGEWOOD CAMPUS · HOWARD COLLEGE CAMPUS · NELSON R MANDELA SCHOOL OF MEDICINE · PIETERMARITZBURG CAMPUS · WESTVILLE CAMPUS

UKZN INSPIRING GREATNESS

Slide 2
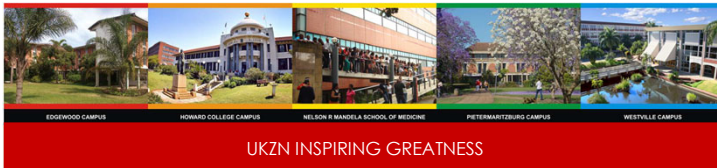
Outcomes for the workshop

- 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

UKZN INSPIRING GREATNESS

Slide 3

Key issues that underpin any test	
Key issues	**Description**
Summative/formative	Be clear on the purpose of the test.
Blueprinting	Plan the test against the learning objectives of the course or competencies essential to the speciality.
Validity	Select appropriate test formats for the competencies to be tested. This action invariably results in a composite examination.
Reliability	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.
Standard setting	Define endpoint of assessment. Set the appropriate standard—eg, minimum competence—in advance.

Wass V, Van der Vleuten C, Shatzer J, Jones R. Assessment of clinical competence. The Lancet. 2001 Mar 24;357(9260):945-9.

UKZN INSPIRING GREATNESS

Slide 4

UNIVERSITY OF
KWAZULU-NATAL
INYUVESI
YAKWAZULU-NATALI

POLICY ON ASSESSMENT
CO/01/0312/2012

Name of document:	Policy on Assessment
Reference number: *(supplied by Office of the Registrar)*	CO/01/0312/2012
Originator/Author: *(name and position)*	Prof Renuka Vithal (DVC: Teaching & Learning) Teaching & Learning Strategy Group
Custodian: *(position/office)*	University Teaching & Learning Office
Policy approved by:	**Structure:** T&L Strategy Group / CAABs / Senate / Council — **Date:** 23/10/2012 / 1-5/10/2012 / 07/11/2012 / 03/12/2012
Policy effective date:	Revised – 1 January 2013

Slide 1

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 purposes 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 flexible.

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.

Slide 2

5.10 Students are responsible for reading, understanding and complying with the rules and regulations related to assessments in the modules and the programme for which they are registered; for using assessments to engage in critical self-evaluation of progress towards learning outcomes; and for behaving ethically and responsibly in the conduct of assessment tasks as stipulated in module outlines, College Handbooks and University Academic Rules.

5.11 In module planning a range of assessment options should be considered such as: peer and self-assessment; criterion- and norm-referenced assessment; formative and summative assessment; and continuous assessment, as appropriate to the outcomes of the particular module. Assessment planning must be guided by the notional study hours for the module to avoid over- or under-assessment.

5.12 Integrated assessments are important in relation to the question of whether the exit level outcomes of the programme or qualification have been achieved.

5.13 Continuous assessments, when approved for assessing an entire module must be:
  5.13.1 systematic – involving assessment tasks that are carefully planned, timed, recorded and communicated to students;
  5.13.2 comprehensive – using a range of assessment methods for formative and summative purposes, with different weightings; levels of complexity and explicitly assessing all outcomes;
  5.13.3 cumulative – so that each assessment builds on the previous and provides scaffolding for the next, giving students multiple opportunities to be assessed on selected tasks; and
  5.13.4 formative – in that each assessment informs students learning, further teaching and assessment.

While no formal summative examinations are written, assessment tasks under test or exam-like conditions may be included. There are no supplementary examinations or special examinations when a whole module is approved for assessment by continuous assessment.
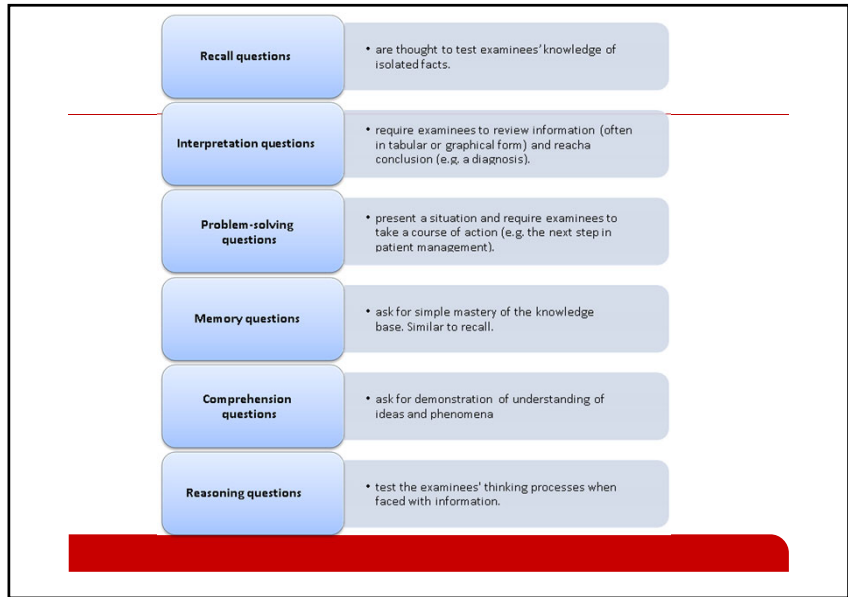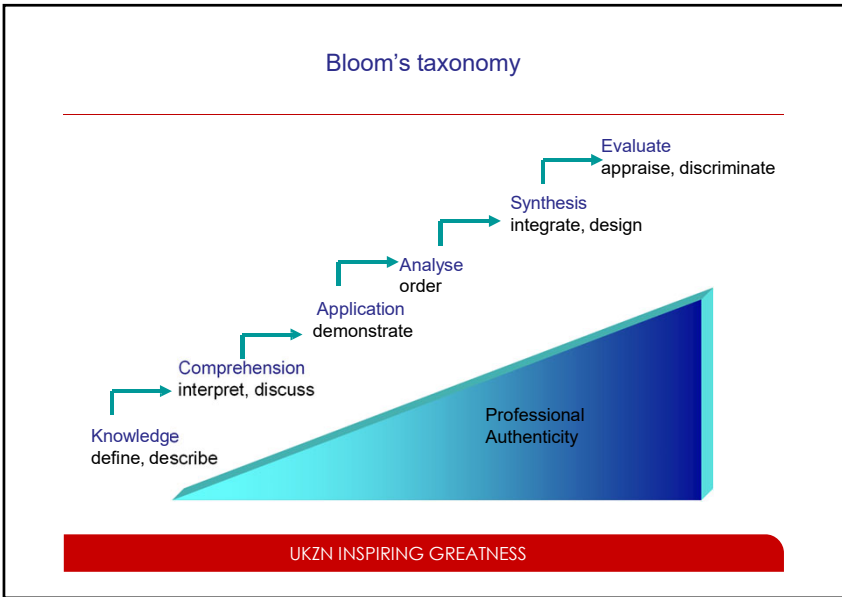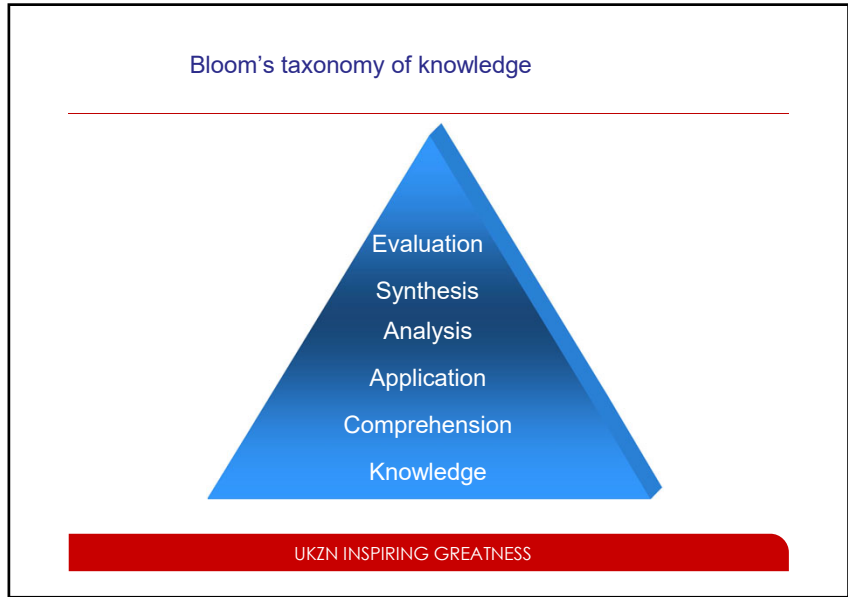
Slide 3

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)

UKZN INSPIRING GREATNESS

Slide 4

Competence

The ability to handle a complex professional task by integrating the relevant cognitive, psychomotor and affective skills

UKZN INSPIRING GREATNESS

Miller's model of competence
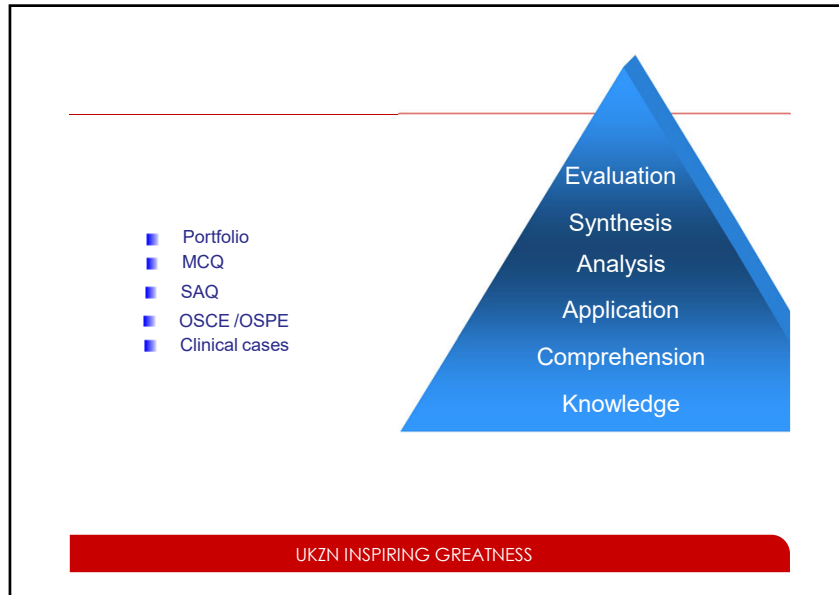
Professional expertise

Does
Shows how
Knows how
Knows

Professional behaviours

UKZN INSPIRING GREATNESS

Bloom's taxonomy of knowledge

Evaluation
Synthesis
Analysis
Application
Comprehension
Knowledge

UKZN INSPIRING GREATNESS

Bloom's taxonomy

Evaluate
appraise, discriminate

Synthesis
integrate, design

Analyse
order

Application
demonstrate

Comprehension
interpret, discuss

Knowledge
define, describe

Professional Authenticity

UKZN INSPIRING GREATNESS

Recall questions	• are thought to test examinees' knowledge of isolated 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 patient 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.

3

Evaluation

Synthesis

Analysis

Application

Comprehension

Knowledge

- Portfolio
- MCQ
- SAQ
- OSCE /OSPE
- Clinical cases

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."
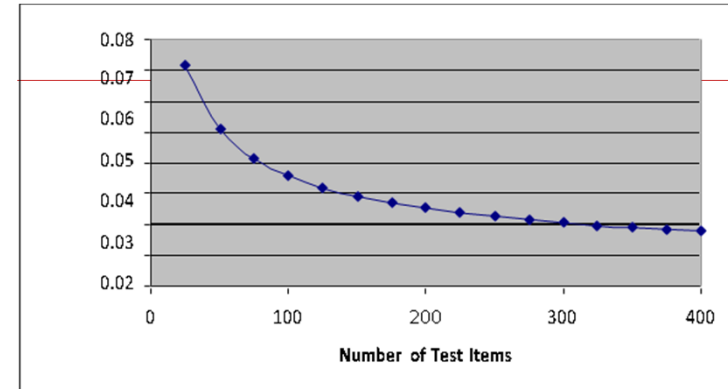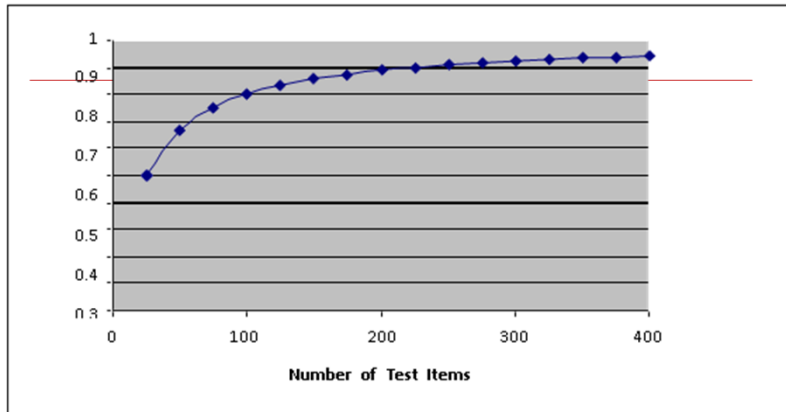
Reliability: Consistency

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

Slide 1

Valid and Reliable Assessments

‣ **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

UKZN INSPIRING GREATNESS

Slide 2

Table 1 Reliability estimates of different assessment instruments as a function of testing time

Instrument	Description	1 hour	2 hours	4 hours	8 hours
Multiple choice[12]	Short stem and short menu of options	0.62	0.76	0.93	0.93
Patient management problem[12]	Simulation of patient, full scenarios	0.36	0.53	0.69	0.82
Key feature case (write-in)[13]	Short patient case vignette followed by write-in answer	0.32	0.49	0.66	0.79
Oral examination[14]	Oral examination based on patient cases	0.50	0.69	0.82	0.90
Long case examination[15]	Oral examination based on previously unobserved real patient	0.60	0.75	0.86	0.90
OSCE[16]	Simulated realistic encounters in round robin format	0.54	0.69	0.82	0.90
Mini-clinical exercise (mini-CEX)[17]	Short follow-up oral examination based on previously observed real patient	0.73	0.84	0.92	0.96
Practice video assessment[16]	Selected patient–doctor encounters from video recordings in actual practice	0.62	0.76	0.93	0.93
Incognito standardised patients[18]	Real consultations scored by undetected simulated patients 0.86	0.61	0.76	0.82	0.86

* One-facet all random design with items crossed with persons (pxi).
† Two-facet all random design with judges (examiners) nested within items within persons (j:i:p).
‡ One-facet all random design with items nested within persons (i:p).

Van Der Vleuten CP, Schuwirth LW. Assessing professional competence: from methods to programmes. Medical education. 2005 Mar 1;39(3):309-17.
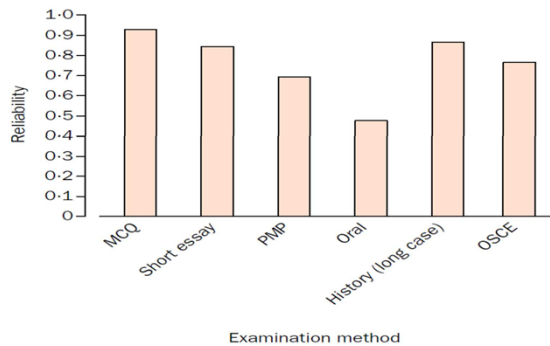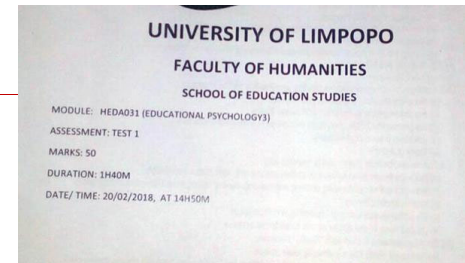
UKZN INSPIRING GREATNESS

Slide 3

Figure 1: Reported reliability when 4 h testing times are used for different test formats
MCQ=multiple-choice examination; PMP=patient management problem; OSCE=objective structured clinical examination.

Wass V, Van der Vleuten C, Shatzer J, Jones R. Assessment of clinical competence. The Lancet. 2001 Mar 24;357(9260):945-9.

UKZN INSPIRING GREATNESS

Slide 4

UNIVERSITY OF LIMPOPO

FACULTY OF HUMANITIES

SCHOOL OF EDUCATION STUDIES

MODULE: HEDA031 (EDUCATIONAL PSYCHOLOGY3)

ASSESSMENT: TEST 1

MARKS: 50

DURATION: 1H40M

DATE/ TIME: 20/02/2018, AT 14H50M

Limpopo students who walked out during exam could face suspension, says VC

Slide 1

Table 1. Illustrations of potential assessment strategies related to qualitative research methodologies for making robust assessment decisions.

Strategies to establish trustworthiness	Criteria	Potential assessment strategy
Credibility	Prolonged engagement	Train assessors. People who know that the learner best (coach, peers) provides information for assessment. Incorporate intermittent feedback cycles in the procedure
	Triangulation	Involve many assessors and different credible groups. Use multiple sources of assessment within or across methods. Organise a sequential judgement procedure where conflicting information necessitates the gathering of more information
	Peer examination (sometimes called peer debriefing)	Assessors talk about benchmarking, the assessment process and results before and halfway an activity. Separate assessors' multiple roles by removing summative assessment decisions from the coaching role
	Member checking	Incorporate the learner's point of view in the assessment procedure. Incorporate intermittent feedback cycles
	Structural coherence	Assessment committee discusses inconsistencies in the assessment data
Transferability	Time sampling	Sample broadly over different contexts and patients
	Thick description (or Dense description)	Assessment instruments facilitate inclusion of qualitative, narrative information. Give narrative information a lot of weight in the assessment procedure
Dependability	Stepwise replication	Sample broadly over different assessors
Dependability/ confirmability	Audit	Document the different steps in the assessment process (a formal assessment plan approved by an examination board, overviews of the results per phase). Quality assessment procedures with external auditor. Learners can appeal the assessment decision

C. P. M. van der Vleuten, L. W. T. Schuwirth, E. W. Driessen, J. Dijkstra, D. Tigelaar, L. K. J. Baartman & J. van Tartwijk (2012) A model for programmatic assessment fit for purpose, Medical Teacher, 34:3, 205-214, DOI: 10.3109/0142159X.2012.652239

UKZN INSPIRING GREATNESS

Slide 2

Table 1
Tips to align questions with course goals and objectives (content validity)

- Base questions on important content and skills/learning outcomes.[5,8,25]
- Questions should be written at an appropriate difficulty level (appropriate does not necessarily mean difficult). Avoid irrelevant difficulty.[8,26]
- Include questions that test multiple levels of Bloom's Taxonomy (remembering, understanding, applying, analyzing, and evaluating).[25]
- Develop an examination blueprint to ensure all concepts that are considered essential and important are given fair and balanced representation on the examination. This will also help overcome writer's block.[9,27]
- Involve curriculum planners, content experts, course coordinators, and medical educators in developing the examination blueprint and bank.[27]
- Avoid testing trivial and non-essential information (factoids).[26,28,25]
- Questions should test the application of clinical knowledge rather than the recall of information, making the question clinically relevant and valid to the candidates.[26-29]
- Define and track question attributes in the bank of questions.[27,28]
- Perform a post-hoc statistical analysis across more than one assessment. Use the results to modify/refine/correct/remove questions. Each question should be critically reviewed before reuse.[9,27]

Dubins DN, Poon GM, Raman-Wilms L. When passing fails: Designing multiple choice assessments to control for false positives. Currents in Pharmacy Teaching and Learning 2016 Oct 31; 8(5):598-608.

UKZN INSPIRING GREATNESS

Slide 3

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

UKZN INSPIRING GREATNESS

Slide 4

- What is currently good? What is less good?
- What would you like to change ?
- What are the opportunities and threats to doing so?
- Summary on flip back: appoint some one to feedback

UKZN INSPIRING GREATNESS

Acknowledgements

- Prof Val Wass and the RCGP