OFQUAL Argument Based Framework – Key Questions

1) Alignment between assessment and the curriculum/syllabus
Test specification and assessment methods are appropriate

- How does the assessment specification reflect the learning outcomes?
- How do the assessment criteria reflect the performance standards?
- Are the methods of assessment appropriate for the construct to be assessed?

Test content, mark scheme, and component weightings are appropriate

- How are the assessment and the results perceived to be accurate and appropriate by the main stakeholders?
- How appropriate is the mark scheme in terms of the weightings assigned to individual tasks in a component and to individual components in relation to their relative importance within the assessed domain?
- How representative is the assessed domain of content of the target domain of content defined by the entire curriculum/syllabus?

2) Accuracy and reliability of scores
Assessment development and administration procedures are appropriate

- What are the procedures used for creating items and constructing tests?
- Are standardised procedures followed when administering the assessment, and how are they monitored?

Construct-related evidence is adequate

- Internal structure of the assessment: How did the items/tasks perform within individual components and how are components related? If statistical models were used, was their use appropriate (how well do the test data meet the model assumptions and fit the model)?
- Analyses of construct irrelevance variance: How the assessment tasks and individual components performed for different subgroups of candidates, taking into account their differences in ability?

Score reliability / generalizability is adequate

Reliability relates to the consistency of results from replications of an assessment procedure. The following aspects of reliability may be investigated:

- Mark scheme: Is the mark scheme appropriate for consistent interpretation to maximise consistency in marking between markers?
- Marker training, marking standardisation, and marking quality monitoring: What procedures are followed to ensure marker reliability?
- Component reliability: Estimates of marker-related, test-related and overall component level reliability.

3) Alignment of boundary scores with performance standards
The standard setting procedures are appropriate

- What procedures are followed by the standard setting process?
- How are statistical information and professional judgement used to ensure that the standards are set appropriately? The standard setting process should pay attention to the defined performance standards.

The outcomes are accurate

- What is the likely error in boundary scores resulting from potential inconsistency in standard setting?
- What is the level of classification accuracy at the overall qualification level?

4) Indicator of future performance
The performance on the assessment is a good predictor for future performance

- How well can the results from the assessment predict outcomes from assessments that measure similar constructs in the future (if data is available)?