(COMFORT) FOOD FOR THOUGHT ON ASSESSMENT

General Principles

“Components of competence show great variability across tasks, they can not be differentiated empirically, and growth in competence is more capricious than expected.”


**************************************************************************************

“Strength [in assessment] will come from a carefully balanced combination of methods.”


Reliability

“The predominant condition affecting reliability of assessment is domain- or content –specificity, because competence is highly dependent on context and content. This means that we will only be able to achieve reliable scores if we use a large sample across the content of the subject to be tested. If the assessment involves other conditions with a potential effect on reliability- such as examiners and patients- careful sampling across those conditions is equally essential…So far, this is nothing new. What is new, however, is the recent insight that reliability is not conditional on objectivity and standardization … In recent years many studies have demonstrated that reliability can also be achieved with less standardized assessment situations and more subjective evaluations, providing the sampling is appropriate.”


**************************************************************************************

“There is a widespread misconception about the relationship between subjectivity and reliability. It is often believed that subjective measures are inevitable unreliable, and that objective measures are reliable by definition. This is not true.”

“Any reliability analysis is a means and not an end….The point is that lower reliability will necessitate more prudence in drawing conclusions from a test than higher reliability. We should be interested not in the reliability of a test but in the reproducibility of the decisions made on the basis of the test.”

“Ideally, one should calculate reliability, then go back to the original data and use the reliability to decide which decisions can be made and which decisions cannot be made with sufficient certainty.”

‘Depending on the sampling strategy applied, like checklists, objective measures may produce unreliable test scores and subjective measures such as more holistic and global professional judgments may yield more reliable test information. “

Van der Vleuten, CPM. The assessment of professional competence: Developments, research and practical implications. Advances in Health Sciences Education 1996; 1; 41-67.

“The consistency of candidate performance across cases (intercase reliability) is perhaps the most important issue in clinical competence testing. Doctors do not perform consistently from task to task. Broad sampling across cases is needed to assess clinical competence reliably.”


Validity

“Several authors have cautioned for sacrificing validity as a compromise to objectivity, particularly in complex professions such as the health sciences. Assessment techniques that avoid professional judgment in the name of objectivity may lead to an atomization of complex skills thereby trivializing the content of the assessment. For example, to break down communication skills into its smallest possible behavioral components in order to be able to check them be on a performance list may enhance objectivity but will not reflect the intended complexity of the skill.’

Van der Vleuten, CPM. The assessment of professional competence: Developments, research and practical implications. Advances in Health Sciences Education 1996; 1; 41-67.

Utility

This is a conceptual model that proposes:
Utility = reliability X validity X acceptability (practicality) X cost X educational impact

The relationship among variables is shown as multiplicative intending that if one variable is equal to zero there is no utility to the method. The other important concept is that these variables are weighted depending on the purpose of the assessment. For example, more cost may be tolerable for high stakes assessment than formative assessment. Assessment that is formative however, should be weighted to have higher educational impact.

Van der Vleuten, CPM. The assessment of professional competence: Developments, research and practical implications. Advances in Health Sciences Education 1996; 1; 41-67.