



Valid Assessment

Purpose:

To explore the concept of valid assessment.

What you need:

- people grouped in planning or teaching teams of 3-6 people
- 1 copy of pages 1 & 2 (back-to-back) per pair of teachers.

Step 1:

- DISCUSS examples of tasks that were probably not valid because:
 - some intervening skills/knowledge got in the way of students being able to show their learning on what was being assessed (such as reading skills, graphing skills)
 - students were not given a fair chance because the task was poorly designed or did not assess all important aspects that were required - but only one segment
 - claims are made about students progress on the Standards but the task did not actually monitor the Standards well at all.
- DISCUSS how inappropriate, unfair and inaccurate assessment can be invalid assessment - and how assessment can be made more valid.

Step 2:

- READ the research on page 2.

Step 3:

- DISCUSS what the key facets of validity mentioned mean:
 - What is 'consequential validity'? How would you recognise it?
 - What is 'content validity'? How would you recognise it?
 - What is 'construct validity'? How would you recognise it?
- AGREE on where each perspective on validity matters:
 - e.g. for which assessment purpose: OF learning? FOR learning? AS learning?

Step 4:

- In small teams AGREE on what 'criteria' you would apply to check if an assessment process is valid for a particular purpose?

Step 5:

- DISCUSS: Are there procedures in place here, now, to ensure the validity of our assessment practices? i.e. To avoid us making inaccurate, unfair and inappropriate judgements on students' learning progression?
 - IF YES: What are they? How do they ensure that individual assessment practices are accurate? fair? and also appropriate?
 - IF NO: What procedures might we use to ensure that individual assessment practices are: accurate? fair? and also appropriate?

For Validity, Fairness, Accuracy and Appropriateness Matter

*Quality and trustworthiness in assessment requires
both valid and consistent judgements on students' learning.*

Assessment has two “main purposes - to help learning and to summarize what has been learned” (Wynne Harlen 2006, p. 103) and assessment needs to be valid for whichever purposes it is used.

“Validity refers to the appropriateness, truthfulness and accuracy of your assessment” and “validity is specific to some use or interpretation (e.g. ... placement, diagnosis) - no result is valid for all purposes” (James Athanasou and Iasonas Lamprianou 2002, p. 167-8).

“All assessment involves judgement and will therefore be subject to some error and bias. While this aspect has been given attention in the context of teachers' assessments for summative uses, it no doubt exists in teachers' assessments for formative purposes.” (Harlen 2006, p. 117).

Validity and assessment for/as/of learning

If “the purposes of the assessment is a key element in validity - ‘it does what it claims to do’ - then the validity argument differs for formative and summative assessment ... In formative assessment it is about consequences - has further learning taken place as a result of the assessment? In summative assessment it is the trustworthiness of the inferences that are drawn from the results - does our interpretation of students' results do justice to their understanding?” (Gordon Stobart 2006, p.134)

“For formative assessment to be valid it must lead to further learning. The validity argument is therefore about the consequences of assessment” (Stobart 2006, p. 133).

This perspective on validity is called ‘consequential validity’.

Valid Portfolio assessment

Ask yourself:

“Does your portfolio measure the behaviours you want it to?” [This is about ‘construct validity’.]

For example, “You may want your math[s] portfolio to assess growth in problem solving ability; yet your evaluation criteria may place too heavy an emphasis on the final solution to the exclusion of the process used to get there ... [such] pitfalls ... can undermine the validity of the portfolio.

In general there are three challenges to validity that you need to address:

representativeness, rubrics and relevance.

Representativeness [‘content validity’ means a good coverage/sampling of the important aspects]

The best way to ensure representativeness is to be clear at the onset about the cognitive learning skills and dispositions that you want to assess and to require a variety of products that reflect these ...”

Rubrics [‘content validity’ and ‘equity validity’]

You need to go about “designing clear criteria” and “You will want criteria for assessing individual entries and the portfolio as a whole”

“Relevance

Assembling the portfolio should not demand abilities of the learner extraneous to the ones you want to assess. A second-grade geography portfolio whose purpose is to reflect skill in map reading and map making should not demand fine motor skills beyond what you would expect a seven-year-old to possess.” (Gary Borich and Martin Tombari 1999, pp. 234-5) [Relevance is ‘construct validity’.]

In order to be fair we need to take some steps to ensure the validity (or fairness, accuracy and appropriateness) of our assessments. Some of these steps can be taken as we design the assessment task and rubric. Other aspects we can only check later, once the students have completed their work and we have assessed it.