

THE CEFR IN EUROPE AND BEYOND

Tim McNamara
The University of Melbourne
tfmcna@unimelb.edu.au

What is assessment?



Assessment is a process of observing students' behaviour in order to reach conclusions about what they know or can do

Compare similar activities:

- Police use evidence to reach a conclusion about an event - leads to a person being charged
- Doctors use evidence to reach a conclusion about a person's state of health
- Scientists use evidence to reach a conclusion about some physical, chemical or biological process

Assessment provides evidence only, and relies on interpretation



- Assessment only gives us evidence, not certainty
- We *interpret* the evidence - we draw *inferences* - we *make claims* about students based on these inferences
- e.g. ‘this student can speak and make herself understood on simple topics’ or ‘this student can follow a simple conversation’
- Our claims remain claims

Mislevy: claims and evidence

An assessment is a machine for reasoning	ASSESSMENT ARGUMENT
about what students know, can do or have accomplished	CLAIMS
based on a handful of things they say, do, or make in particular settings	OBSERVATIONS/ EVIDENCE

Evidence Centred Design Stage 1: Defining the construct - domain modelling

Claims	Evidence	Task
What you want to be able to claim that students know or can do	What a student would need to be able to say or do as a basis for the claim	Situations that might make it possible to obtain evidence needed

What is the CEFR?

- It represents a construct definition; it is an exercise in domain modelling
- It provides a set of claims
- It provides a general characterization of evidence and tasks
- It is not a test - it allows different kinds of tests to be realizations of this construct

What happens when a particular set of claims is uniquely privileged?

- The privileging of a construct represents an exercise of power
- It gives power to those who mandate that outcomes be reported in terms of this set of claims - governments, ministries, authorities, funding agencies, testing agencies
- It gives power to those who adhere to the construct over those who do not

If the CEFR is the only construct in town...

- All claims in all tests are interpretable only in terms of this prescribed set of claims
- Any claims that are not interpretable in terms of this set of claims are inadmissible

All tests need to show that they are making the same claim

- TOEFL iBT could not be marketed for use in Europe unless its claims were interpretable in terms of the claims of the CEFR
- Taiwan's Government realized the power that could accrue to its assessment if its claims were interpretable in terms of the CEFR
- TestDaF had no choice, as a European test

What's wrong with this?

- The adequacy of these claims cannot be disputed
- The limits to thought about the goals of language education and the possibilities of other claims is unhealthy
- The values underlying this set of claims cannot be disputed, unless through complex policy processes

The CEFR: A universal language?

- Arguments in favour of ease of interpretability remind one of arguments in favour of spread of English
- Impacts of spread of CEFR need to be studied, just as in the case of the spread of English
- Different kinds of studies of the US, Taiwanese and German cases would result
- Policy that encourages linking studies not matched by policies encouraging alternative studies

Messick's validity matrix

	WHAT TEST SCORES ARE ASSUMED TO MEAN	WHEN TESTS ARE ACTUALLY USED
USING EVIDENCE IN SUPPORT OF CLAIMS	What reasoning and empirical evidence support the claims we wish to make about candidates based on their test performance?	Are these interpretations meaningful and useful in particular contexts?
THE SOCIAL CONTEXT OF ASSESSMENT	What social and cultural values and assumptions underlie test constructs and hence the sense we make of scores?	What happens in our education systems and the larger social context when we use assessments?