

Measuring writing quality: The role of the construct in writing assessment project design

Liz Hamp-Lyons
CRELLA

University of Hong Kong
University of Nottingham

What is writing?

- is it a product?
- ‘writing’ is not hard—is it?
- it’s not hard to teach--is it?
- It’s not hard to assess—IS IT?
- How CAN we measure ‘writing quality’?

What is writing 'quality'?

- Measuring writing quality is difficult because we remain unsure exactly what writing *is*.
- In other words, we have a construct issue.
- Research and development work on the assessment of writing has, for most of the time it has existed, concentrated on reliability issues.
- Yet progress towards greater reliability has been stalled for about 30 years.

Unpacking assumptions

- product – process
- expression – substance
- accuracy – fluency
- unigeneric – genre diversity

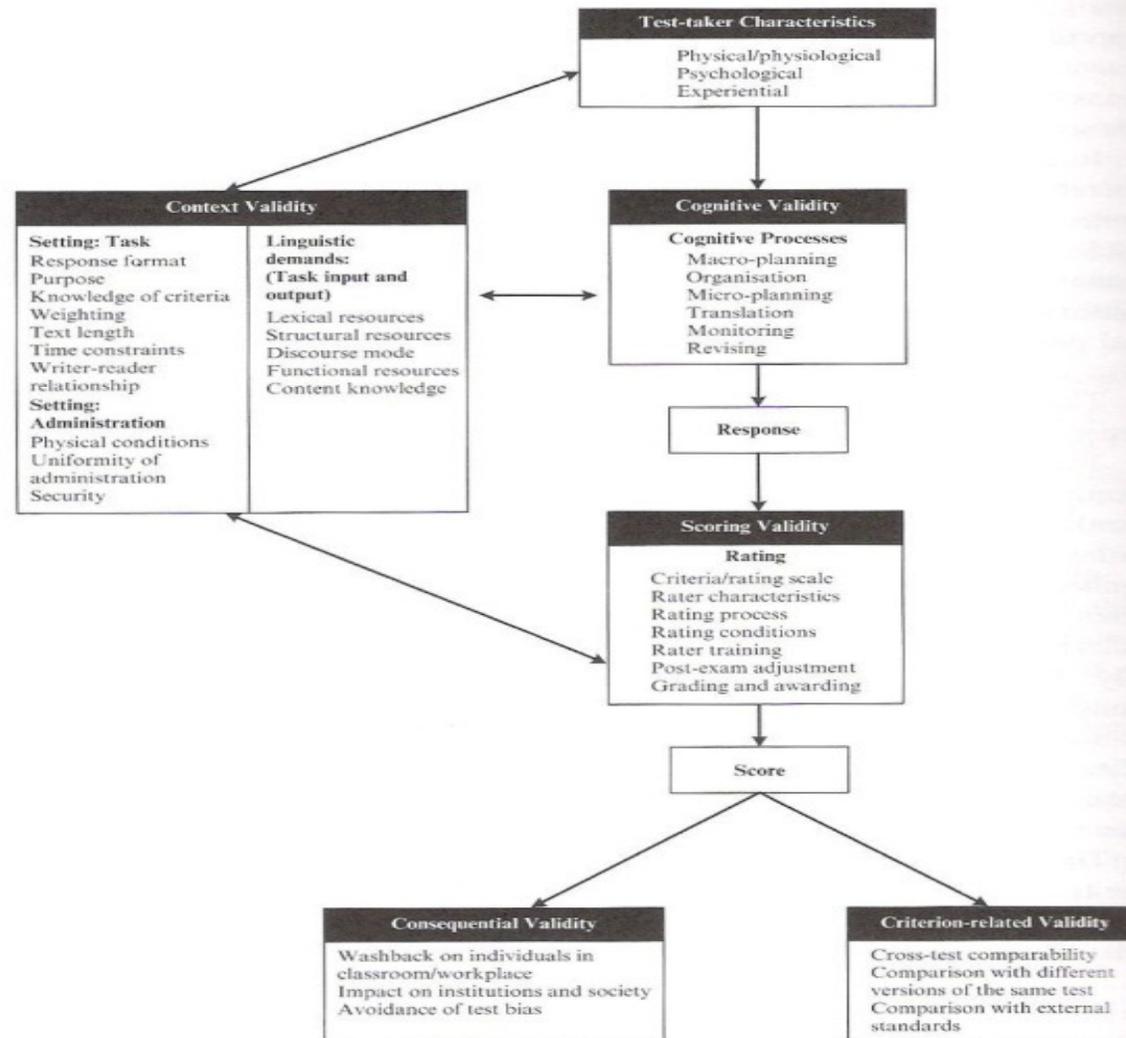
It depends...

- ... depends **on what?**
- Weir's 2005 framework, updated to apply specifically to writing by Shaw and Weir 2007, suggests some of the elements of dependency that argue against making assumptions about what writing quality "is":

Shaw & Weir (2007)

1 Introduction

Figure 1.1 A framework for conceptualising writing test performance (adapted from Weir 2005b:47)

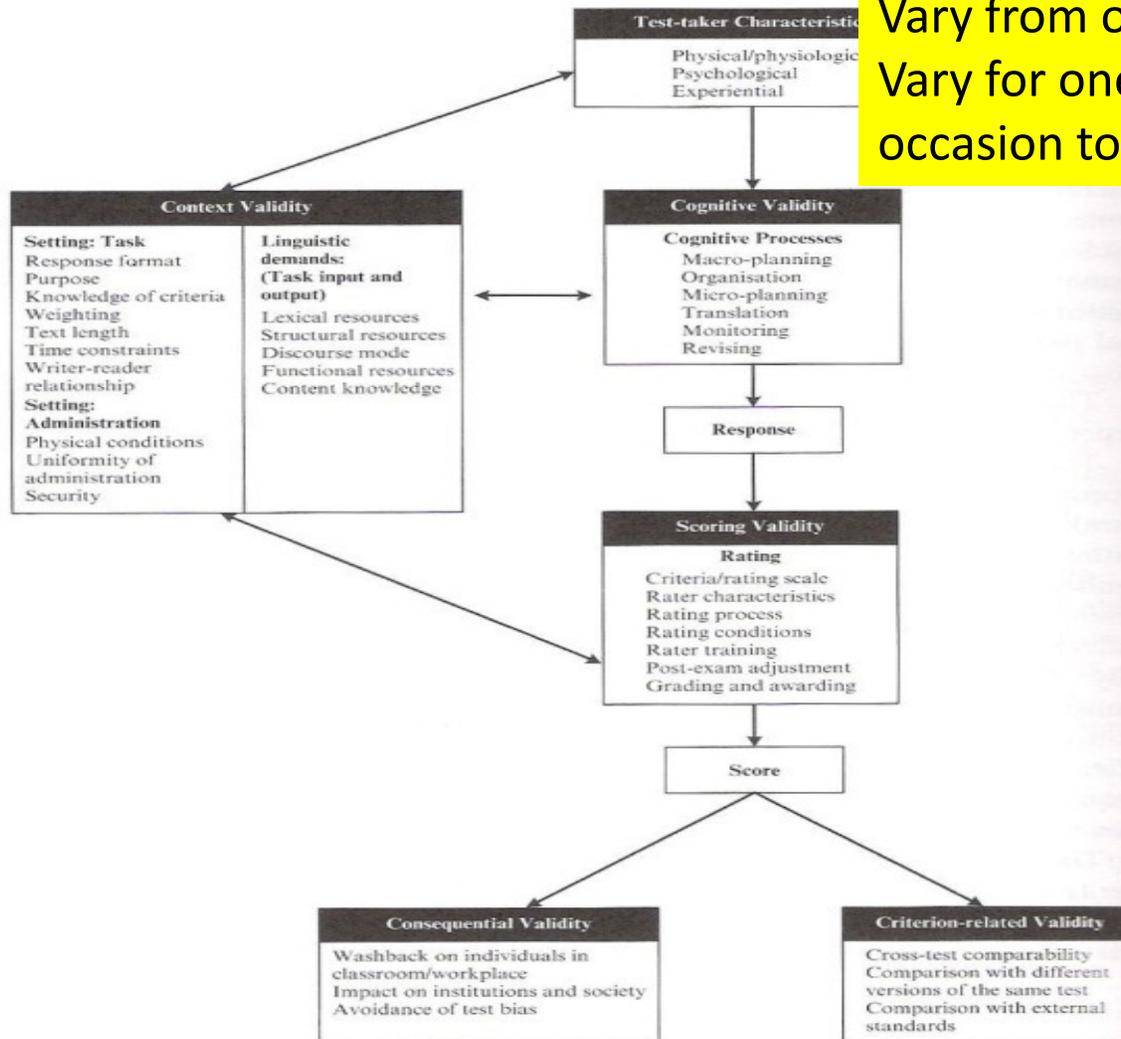


- How are the physical/physiological, psychological and experiential

Shaw & Weir (2007)

1 Introduction

Figure 1.1 A framework for conceptualising writing test performance (adapted from Weir 2005b:47)

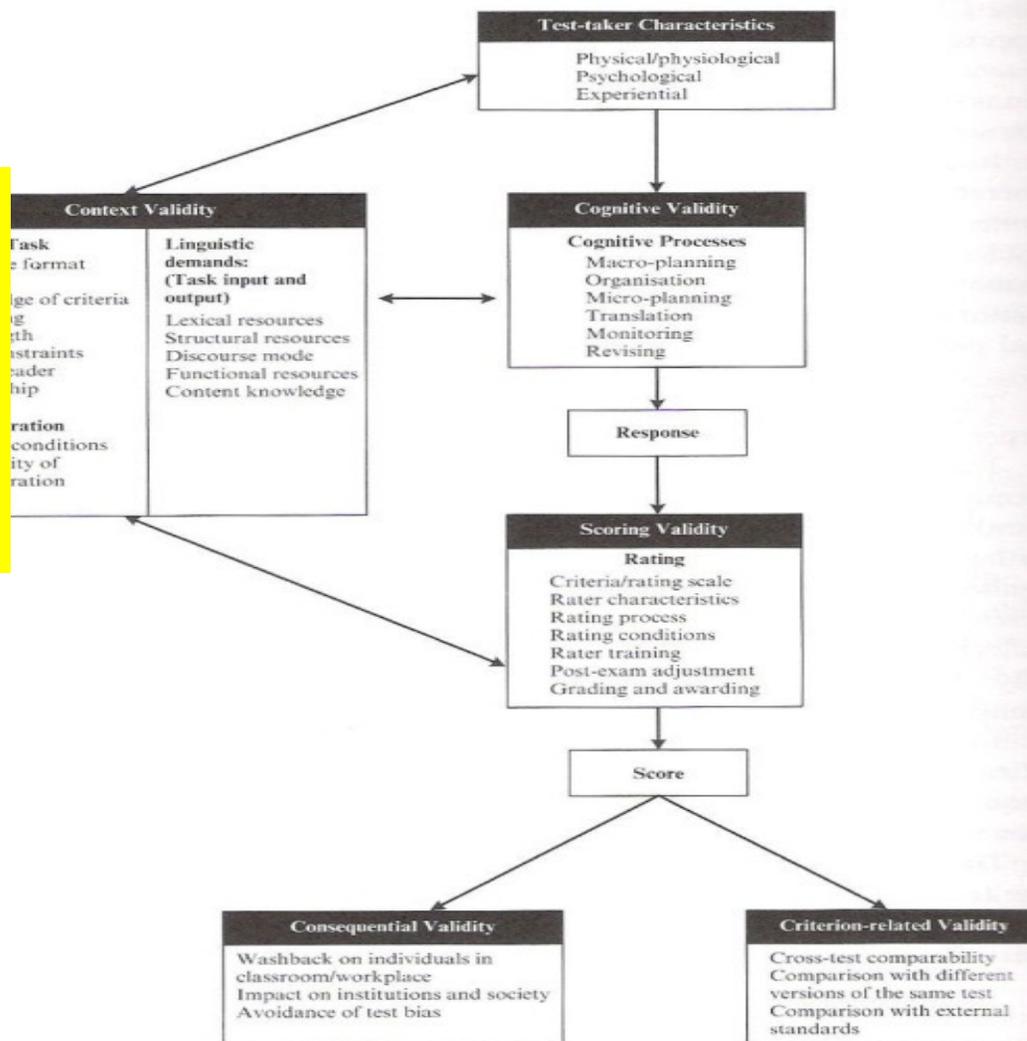


- How are the physical/physiological, psychological and experiential

Shaw & Weir (2007)

1 Introduction

Figure 1.1 A framework for conceptualising writing test performance (adapted from Weir 2005b:47)



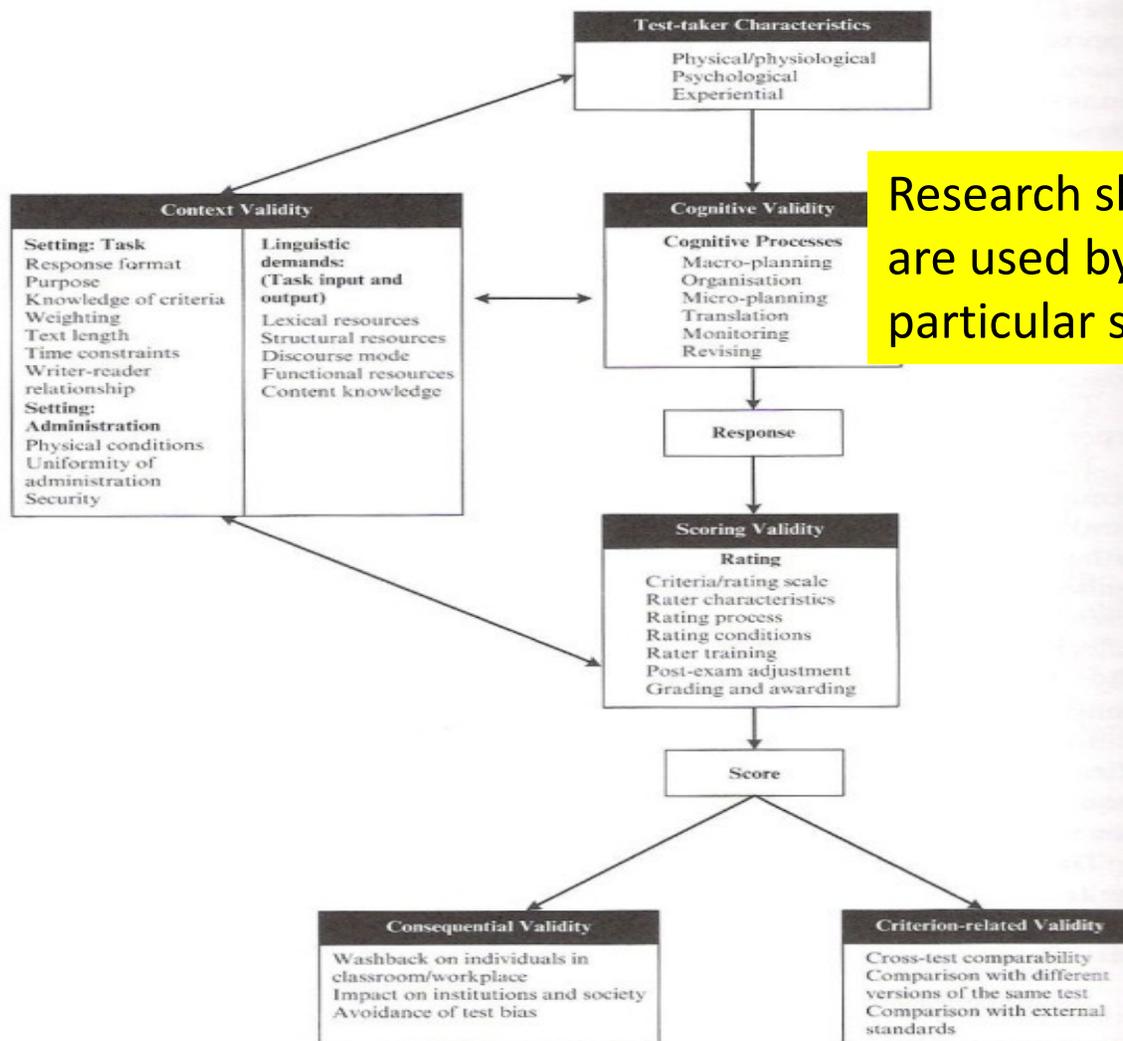
All elements within Task, Setting and Linguistic Demands not only vary, but vary by interaction with each other

• How are the physical/physiological psychological and experiential

Shaw & Weir (2007)

1 Introduction

Figure 1.1 A framework for conceptualising writing test performance (adapted from Weir 2005b:47)



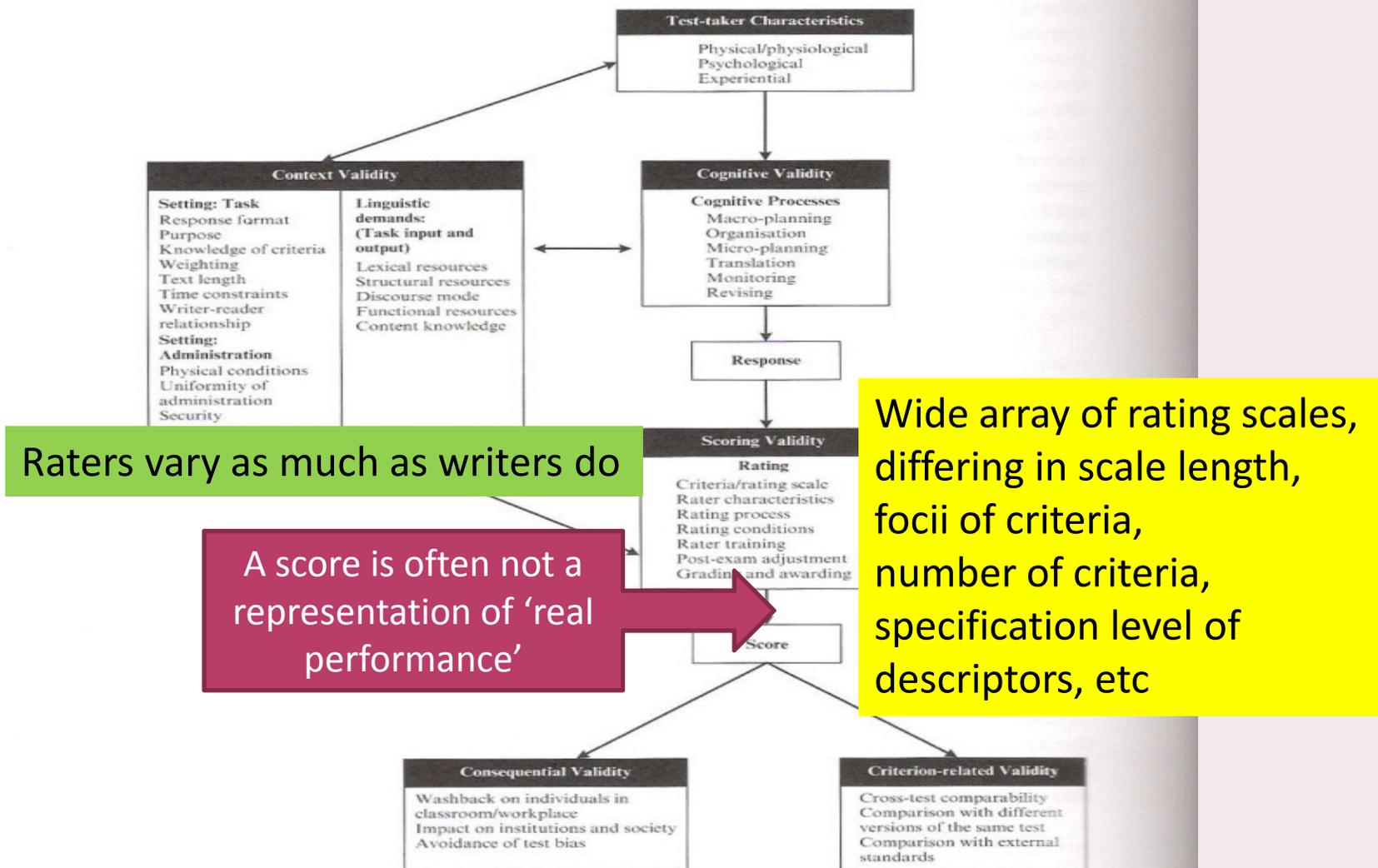
Research shows that not all of these are used by any one writer under any particular set of conditions

- How are the physical/physiological, psychological and experiential

Shaw & Weir (2007)

1 Introduction

Figure 1.1 A framework for conceptualising writing test performance (adapted from Weir 2005b:47)

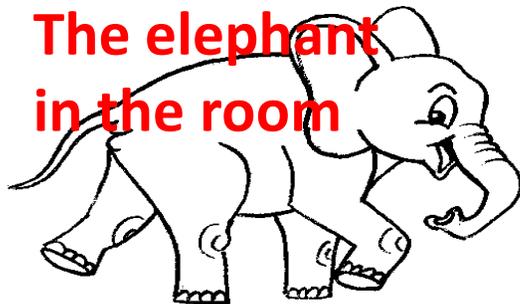
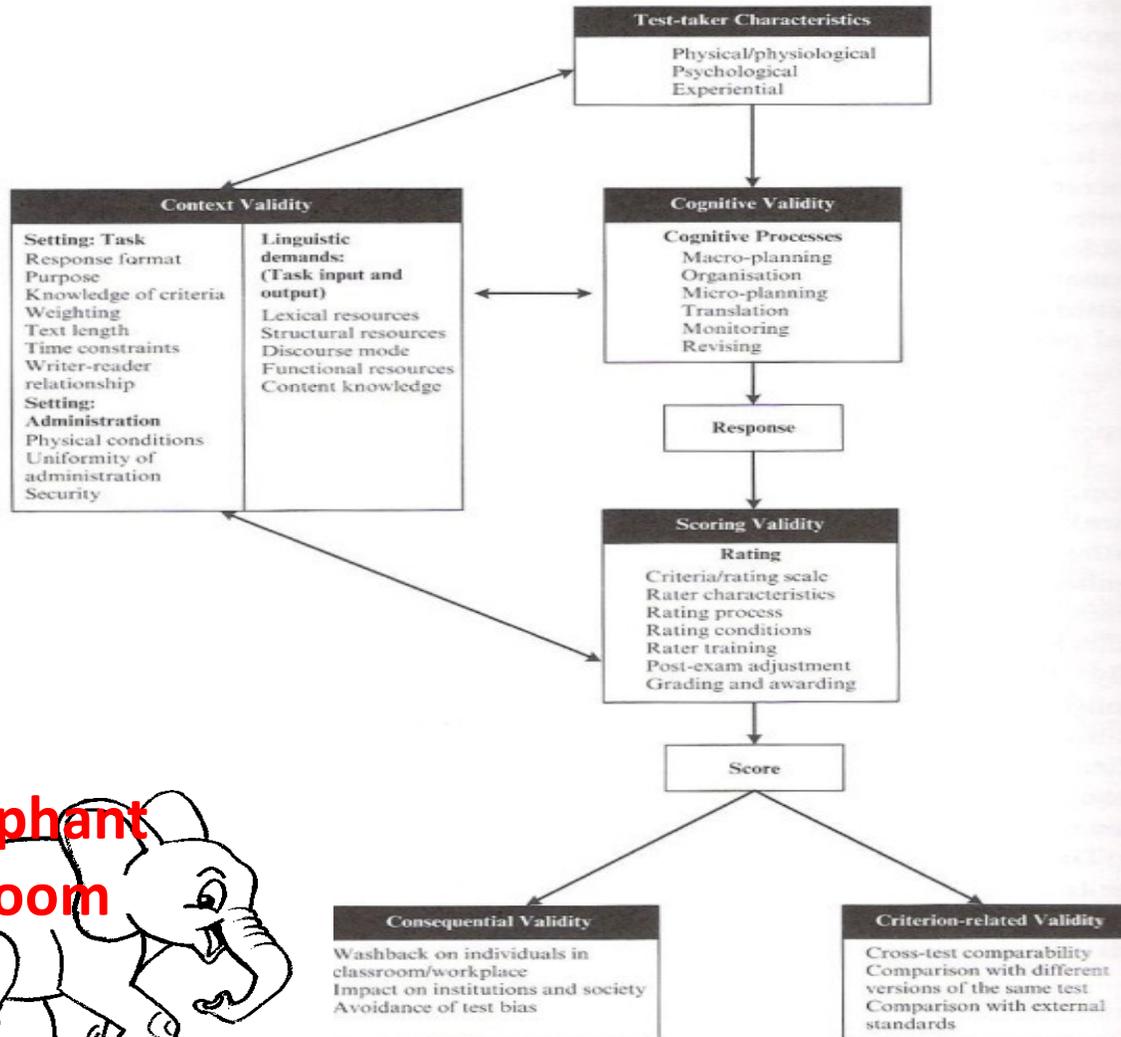


- How are the physical/physiological, psychological and experiential

Shaw & Weir (2007)

1 Introduction

Figure 1.1 A framework for conceptualising writing test performance (adapted from Weir 2005b:47)



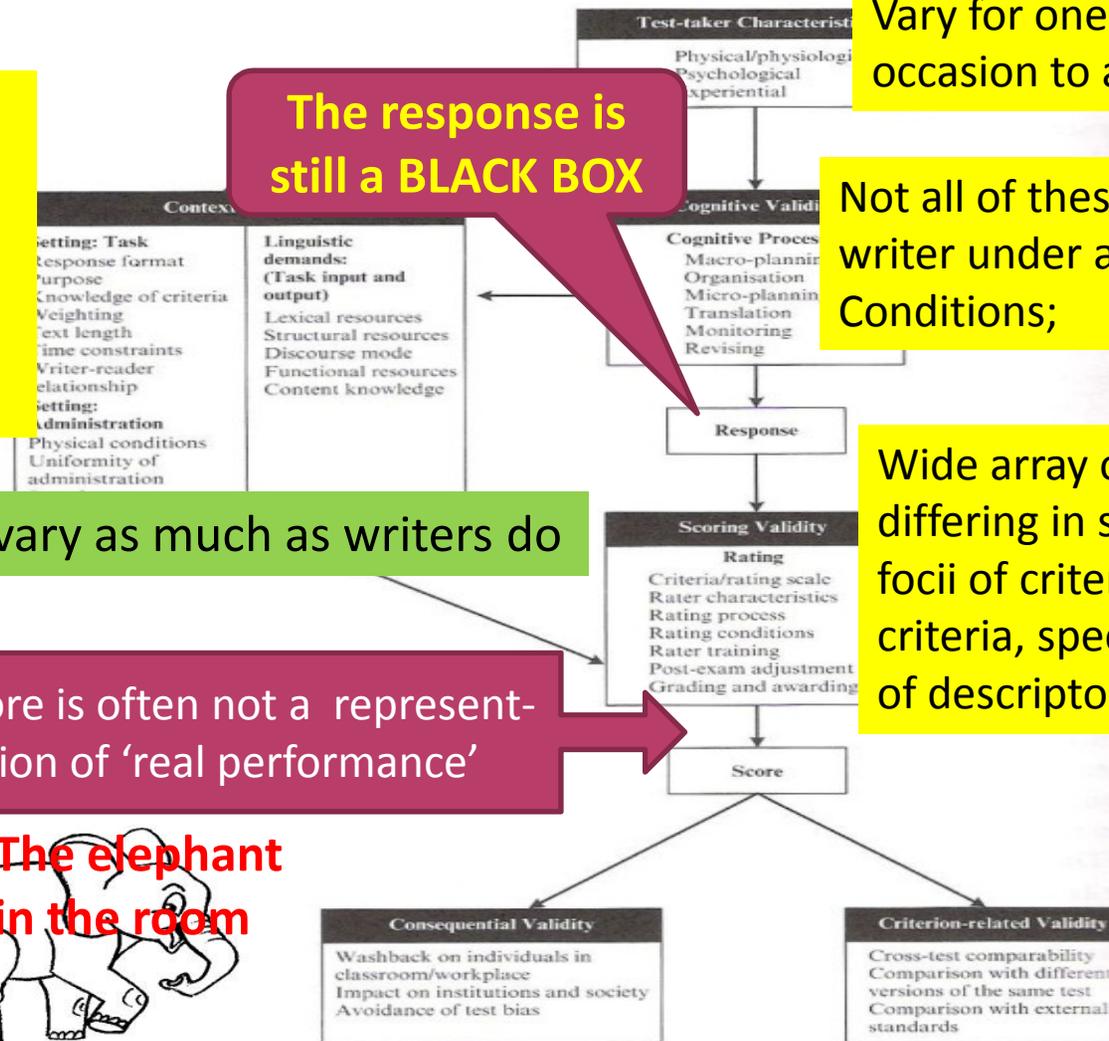
The elephant in the room

The easy bit – in theory!

Adapted from Shaw & Weir (2007)

1 Introduction

Figure 1.1 A framework for conceptualising writing test performance (adapted from Weir 2005b:47)



All elements within Task, Setting and Linguistic Demands not only vary, but vary by interaction with each other

The response is still a BLACK BOX

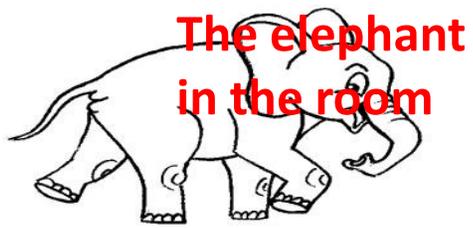
Vary from one writer to another
Vary for one writer from one occasion to another

Not all of these are used by any one writer under any particular set of Conditions;

Raters vary as much as writers do

Wide array of rating scales, differing in scale length, foci of criteria, number of criteria, specification level of descriptors, etc

A score is often not a representation of 'real performance'



The easy bit – in theory

- Not only is this very, very complex – it relates only to constructs of writing TEST performance – not to the majority of real world writing

Therefore -

- With writing assessment (as with all assessment of performance that is not 100% physical) there is a persisting validity issue.
- Writing assessment projects either need to begin with a sound understanding of the writing to be assessed in the specific context, or they need to be focused on gaining that understanding.

Examples

- One way to begin with a sound understanding is to (1) tightly define and limit the context for writing (e.g. end of term writing for an under-graduate History class (2) conduct rigorous needs analysis (3) write and consult on clear specifications.
- Then it should be possible to (4) design tasks and rating scale to fit the specs (5)select appropriate raters and train them on this specific scoring instrument (6) conduct validation to see how this has worked.

Examples

- A very different approach is to begin by seeking understanding through close engagement with real writers doing real writing in the actual context. This happens in Writing Centres in the US, Australia, Hong Kong. Instructors consult with learning writers on writing they are set for classes; they give feedback and teach writers how to use feedback, and how to give it to others.
- Experience with feedback is formative for writers, and also for teachers of writing: it can be informative to writing assessment developers and researchers too.

Why do we need a construct?

- While these are important to resolve, until we understand the act, art and purpose of writing better we will continue to ‘measure’ only trivial aspects of ‘writing’.
- If we do that, scores might be reliable—but can they be valid?

Progress

- Research into the constructs behind reading-into-writing, or “source-based”, writing tasks
- Research into task representation, linguistic features of tasks, and writers’ interpretations of tasks
- Research on rating scales, raters’ behaviours, and diagnostic uses of scales (the English Profile project can be seen as an example of this)
- Research into the nature of feedback