THE PLACE OF LANGUAGE IN SCIENCE ASSESSMENT:
implications for validity

_EAA Conference, Sydney,_
13 October 2014

Presenter: Dr Saw-Choo TEO
PhD (Linguistics) University of Sydney
Consultant in Language and Education
1 The issues

1. Assessing Science in Education – the prerequisites of validity and reliability

2. Scientific literacy/the register of Science

3. The communicative efficiency of elements/channels used in test items

4. The assumption of test-takers’ competence in decoding the language and other media used to communicate the message (test item)

5. The need to avoid presenting too much information - cognitive load

…………………

2 Interactive/Communicative situations

➢ For building social relationships – phatic communion where content/information in message is not important – talk about ‘the weather’ to ‘break the ice’ and establish/cement relationships.

➢ For conveying information – test situation – where the delivery of the content of the message is most important.

➢ Most situations in life where both elements matter
3 The test situation

- Information conveyed in the message (test item) is crucial and **clarity** is of primary importance.

- Means by which information is delivered – form(s)/channel(s) used to deliver the message:

1. Language only?

2. Language plus one other form e.g. diagram? illustration? table? chart? map?

3. Language plus **two or more forms**

4 The prerequisite of validity in testing

**Validity** relates to the relevance or pertinence of a test to the subject/area tested.

Does the test actually test what it claims to test? For example:

- The purpose of a science test is to measure ability in science and **NOT**, say, literacy skills in English

- The purpose of an English test is to measure competence and performance in English, **NOT**, say knowledge of biology or geology
4.1 Content validity

**Content validity** – relates to the content of the test

- Test specifications – are they drawn from a curriculum framework/syllabus: is there such a framework/syllabus?

- A summative test should contain a representative sample of ALL the elements/skills of the field/subject it is meant to test

- A formative test is more restricted

- The broader the representation, the better the content validity, the more accurate or reliable the test.

4.2 Construct validity

**Construct validity** – relates to what specifically is being tested in each item/segment/question (e.g. understanding the process multiplication involving fractions).

- A test item has ‘construct validity’ when it is clear that it measures just that ability or skill (and NOT something else or something more than) it is supposed to measure.

  e.g. an item testing fractions must strictly test understanding of fractions and NOT something else or something more.
5 Test-taking as an interactive situation

- A sender of a message – the test developer
- A message – the test item
- A receiver/recipient who interprets and responds to the message (the test item) – the student/test-taker

6 The Sociolinguistics: Language, Intention and Meaning

In real-life communication, language is the means by which

- the intention of the speaker/language user is conveyed
- the content of the intended meaning is conveyed through a variety of language and other means
- In real-life communication, speaker/user intention is not necessarily always clear – there is often a need for subtlety/ambiguity, not just clarity.

In testing, however, CLARITY is the absolute priority.
6.1 Speaker Intention and Hearer Interpretation (Pragmatics)

J.L. Austin & J.R. Searle

✓ Locution – the words / forms chosen
✓ Illocution – the speaker’s / user’s intention
✓ Perlocution – the hearer’s / receiver’s understanding / interpretation

6.2 Effective communication: 1 providing a context

Providing contextual Clues: theories of communication

1 Communication as transmission or ‘Information theory’: a message-centred view of communication

\[ \text{SENDER} \rightarrow \text{MESSAGE} \rightarrow \text{RECEIVER} \]

➢ The sender is regarded as the ‘active’ party and the receiver ‘passive’
➢ The sender is solely responsible for the message
➢ Not reflective of real communication

2 Communication as transaction: The focus is on meaning and interaction

\[ \text{SENDER} \rightarrow \text{MESSAGE} \leftrightarrow \text{RECEIVER} \]

➢ Meaning is jointly-constructed by both sender and receiver.
Both/all participants negotiate meaning in communication – in a real-life situation.

Many test items are presented in a context.

Is the context familiar to the receiver?

**Effective communication: 2 context and clarity - pronouns**

- Use of PRONOUNS depends on clarity of context (anaphora and cataphora) can lead to misinterpretation.
- All languages use PRONOUNS e.g. ‘it’ ‘this’ ‘that’.
- Speaker/sender/test developer can make the mistake of assuming that pronominal references are ‘clear’.

**Effective communication: 3 context and clarity - word choice**

Levels of meaning in word choice

- Denotation = straight literal meaning, as *per* the dictionary (no frills)
- Connotation = meaning plus ‘suggestions’, affective meaning

**Effective communication: 4 idiomatic language**

- Language and culture intertwined
- Idiomatic language is culturally-learned (e.g. phrasal verbs)
- Best to avoid idiomatic language in the wording of science/non-language test items.
7 The psycholinguistics: the issue of cognitive load

Long-term memory: what we consign to memory for retrieval, as and when required – massive in capacity

Short-term memory or working memory: what we are able to retain in the short-term while working out a solution to a ‘problem’ – this is what is mainly at work in a test situation

8 Teaching and Testing

➢ The desire to ‘make things clear’ in a test item is understandable – but it could lead to language-dense items. Increased cognitive load?

➢ The need to ‘set out the problem’ is again understandable, i.e. provide the ‘situation’ but again leads to language-dense items.

➢ Given (old) information needlessly added? Distinguish teaching from testing

9 Language and cognitive load

➢ Complexity of language/density of text imposes enhanced cognitive load, especially if irrelevant to the task at hand

➢ Opacity – resulting from ambiguous or inappropriate lexical choice; oblique grammatical expression, idiomatic language, etc. unnecessarily increases cognitive load
Both complexity and opacity are very much decided by the language used in the test, esp. if it is NOT one’s first language (or language of thought).

10 Retrieving and presenting information – deciding what needs to be communicated

1. Deciding what is essential information – i.e. ‘new information’ and vital to the communication

2. Deciding what is given / old information – that which is redundant /already known / not important / may be deleted / unnecessary

11 What needs to be communicated in a test?

✔ What is essential?

➢ New information vital to the item, that needs to be processed to solve the problem and to arrive at the required answer

✔ What may be omitted?

➢ Parallel/complementary information

➢ Redundant/repeated (unnecessary – adds to cognitive load) information

➢ Confusing/padded information?

➢ Worse still, conflicting information?
12 Writing good Science items: a summary of language concerns

✓ Distinguish testing from teaching

✓ Prioritise clarity of communication

✓ Use clear and unambiguous language throughout – avoid opacity

✓ Enforce clear and unambiguous rubrics/instructions

✓ Unless testing language, avoid vagueness e.g. indirect or idiomatic language

✓ Avoid tripping up test-takers with ‘double entendres’ (by being ‘too clever’, by insinuating ‘Are you thinking what I’m thinking?’); repeated information and multiple negatives, etc.