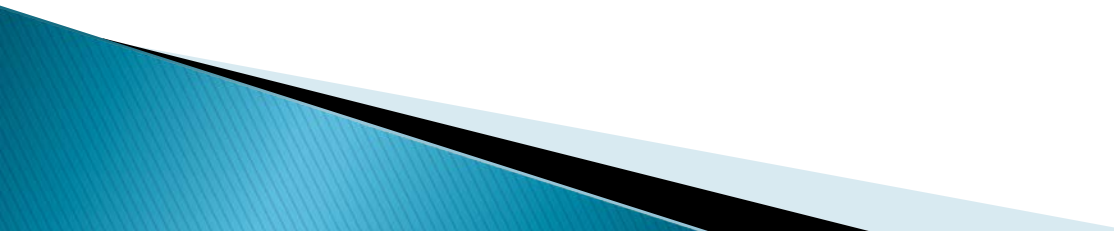


TESTING IN BILINGUAL EDUCATION PROJECTS: LESSONS LEARNT IN THE ABLE PROJECT

Elize Koch
DBE Research Indaba,
Pretoria

Overview

- ▶ The two stages in research and testing in the ABLE project
 - ▶ Lessons learnt about macro-contextual processes and “forces”
 - ▶ Lessons learnt about translation/adaptation/bilingual testing
 - ▶ Recommendations beyond ABLE
- 

The two stages in ABLE research: implications for testing in the project

First stage: Experimental quantitative focus

- ▶ ABLE Project: Homelanguage based bilingual education – till grade 6 (model of late exit bilingual education)
- ▶ Three research aims: how and if improve learning if isiXhosa used as LOLT for longer
- ▶ Experimental design: compare our school – two others
 - cognitive development,
 - language development, and
 - academic performance
- ▶ Testing:
 - Grade 3, 7 (and 9)
 - KABC (cognitive); Woodcock Muñoz Language Survey (WMLS); Imbewu tests (grade 3); JET tests
 - Needed tests in English and isiXhosa

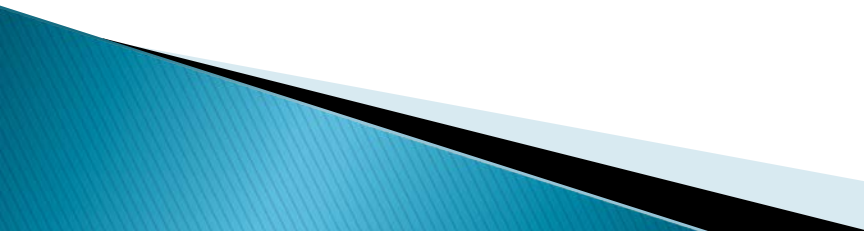
Adapted the English WMLS ->: isiXhosa

Second stage: action research and developmental focus

- ▶ ABLE symposium with EC DoE in 2008
 - Policy and model development, bi-literacy development, teacher development, terminology and materials and assessment
- ▶ Move away from testing to participatory action research
- ▶ Language Policy and model development: IsiXhosa only till end of grade 6
- ▶ Workshops, consultations
- ▶ 2010: ABLE children better on common tests than comparable schools
- ▶ Many problems -> no interventions during 2011

Testing lessons learnt: macro-contextual processes and “forces”

Current events at the school and project

- ▶ Drop in numbers and redeployment of teachers;
 - ▶ Phasing out isiXhosa LOLT at school
- 

Lessons learnt about macro-contextual processes and “forces” in testing

Processes impacting on project

- ▶ **Testing in project continued in any case but driven by external forces**
- ▶ 2009: systemic tests –> forerunner of the Annual National Assessments (ANAs)
- ▶ From 2009: common tests in the EC – twice yearly
- ▶ From 2011: ANAs
- ▶ Contradictions between LiEP of DoE and language in tests: only grades 1 – 3 in isiXhosa; from grade 4 only English or Afrikaans (see doc on ANA)
- ▶ **NO TRANSLATION OF THE TESTS FORTHCOMING FROM GRADE 4 ONWARDS**
- ▶ 2011: ANAs and common tests: Sosebenza learners underperform

Forces

- ▶ Social constructionism as paradigm and post colonial theorising framework –> neo-Fanonian
- ▶ Shohamy, and USA activists: explore political agendas of language in tests + contest testing of bilingual children
- ▶ Bulhan (1985) in Hook (2003): 3 stages in post colonial identity
 - Capitulation: transition into only English
 - Revitalisation: MT?
 - Radicalisation: the creation of third spaces and “languaging”: bilingual education and assessment
- ▶ Hypothesis: contradictions currently caught up/trapped in the first two stages
- ▶ Need to move into the third stage: bilingual tests to support

Bilingual testing

Distinctions and clarification of terms

- ▶ **Assessment:** a broad process of gathering information about a child (e.g. progress in a learning area);
 - tests form part of assessment and produce scores that must be **valid** (and interpretable)
- ▶ **Cross linguistic testing:** testing that takes place across language groups
 - ✗ Monolingual tests
 - ✗ Bilingual tests
- ▶ **Bilingual tests:** tests that are available in more than one language
 - tests that are available in two or more languages (two versions of the same tests)
 - Two languages in one test



Equivalence and bias as part of validity

Equivalence:

- The scores of the different language groups must mean the same

Bias:

- Items: when members of different groups with the same ability perform differently on an item
- the whole test: different constructs
- Method of administration

If bias is present: the scores do not mean the same thing

All tests in bilingual testing must be evaluated for bias: monolingual and bilingual tests

The Woodcock Munoz Language survey: an example a translated test

Sub-tests	Linguistic and curriculum areas	Stimuli	Test requirement	Response
Picture Vocabulary (PV)	<ul style="list-style-type: none"> •Oral expression •Language development •Expressive vocabulary 	Visual (Pictures)	Identify objects	Oral (Word) Total=57
Verbal Analogies (VA)	<ul style="list-style-type: none"> •Receptive-expressive vocabulary 	Auditory (Phrases)	Stating a word to complete and analogy	Oral (word) Total = 35
Letter Word Recognition (LWR)	<ul style="list-style-type: none"> •Reading •Reading-decoding 	Visual (text)	Identifying printed letters and words	Oral (letter name, word) Total= 56
Dictation (Dict)	<ul style="list-style-type: none"> • Spelling, writing language development 	Auditory (Words)	Writing skills and grammar	Motor (writing) Total=56

Bilingual testing: practice and research in ABLE

Practice of adaptation WMLS into isiXhosa

- Adapted into isiXhosa not translated
- Two workshops with multilingual and multidisciplinary team
- Linguistic and cognitive processes:
 - grading of difficulty of items;
 - underlying cognitive processes; relexification → loan words, roots;
 - reformulation of items

Results on the WMLS

English monolingual test across EL1 and XL1 groups:

- ▶ All subtests have biased items; some up to 40% of items (LWI)
- ▶ VA: measuring different constructs in the two groups

isiXhosa monolingual test across XR and XU groups:

- ▶ The subtests have biased items but far fewer than English monolingual test
- ▶ PV: equivalent constructs, but scores need to be interpreted with caution; better to assess Vocabulary in context

English (EL1) and isiXhosa (XL1) versions:

- ▶ All subtests have biased items but mostly fewer than on English test
- ▶ Rasch modelling on VA: same

Lessons learnt about bilingual testing

- ▶ Equivalence – always an issue in both monolingual tests and bilingual tests
- ▶ It is more valid to use the two–languages than the one language approach
- ▶ Propose to use tests in a criterion referenced manner:
 - What score indicates “proficiency” in a group? It may differ across groups
- ▶ Dialect differences do not impact that much on test scores of this nature: slight bias necessitate approaches that are more holistic
- ▶ Two languages in one test – in line with SIOP approach

Recommendations beyond ABLE

- ▶ Discourses around tests are powerful
- ▶ **Might** be useful in large scale programme evaluations
- ▶ Engage with the discourse around the language of tests and the purpose of tests
- ▶ To engage:
 - Purpose of **testing** and the purpose of **bilingual testing**: transition or developmental maintenance bilingualism and bi-literacy – in line with the model
 - Then:
 - Language of test: 1) one language, 2) two different language versions or 3) two languages in one test
 - Content and format: in line with underlying processes e.g. Reading of bi-literate learners + assessment principles → improvement in instruction
 - An example of reading:
 - For research: Combine large scale test scores with samples using over the shoulder miscue analysis and running records to improve tests and interpret results
 - Feedback loop between test development and findings and practice
 - For instruction and evaluation: combine the test scores with holistic assessments (as above) for better understanding of where to go.