10th Western Cape Education Students' Regional Research Conference

Theme: 'Education as a lever for change'

Completed Master's Thesis
UWC-Department of Psychology

"The evaluation of group differences and item bias, across two language groups, of the English version of a standardised test of academic language proficiency."

By Ms Genevieve Haupt

Thesis Supervisor and Head Researcher
Prof. S.E. Koch

Background

Social science that makes a difference

HSRC
Human Sciences
Research Council

- ✓ Additive bi-or multilingual education is a form of education whereby the primary language of learners is used for cognitive and literacy development, while at the same time learners are being taught a second and/or third language.

- ✓ The South African Policy on language-in-education is one of additive multilingualism or bilingualism, however in reality most black children are being educated through the medium of English from Grade 4.

- ✓ This therefore affects the development of the proficiency in their primary language since these children are not engaging in cognitively demanding tasks, with the necessary academic language proficiency, in their primary language.

- ✓ With the implementation of additive bilingual education programs, came the necessity to evaluate the effectiveness of these programs, in order to establish whether learners are acquiring the necessary language proficiency in their primary language.

- ✓ One such program was started up in the Eastern Cape (Additive Bilingual Education-ABLE Project)-which had been largely formed to evaluate the effectiveness of additive bilingual education programs.
- ✓ The Woodcock Muñoz Language Survey (WMLS) is a test that is used extensively in the USA for this purpose.
- ✓ Following the International Guidelines for Test Adaptation, WMLS adapted for South African use and was done as one of the components of the ABLE Project.
- ✓ With the adaptation of tests for cross-cultural or cross-linguistic use of tests, there are two important psychometric issues to be considered, namely, equivalence and bias.

- ✓ It is important to note that this study is a sub-study which forms part of the ABLE Project.
- ✓ This study has focused on the adapted English version of the WMLS, where the researcher has evaluated the overall equivalence of this version to be used across the English-and Xhosa-first language groups.

Equivalence and Bias

Social science that makes a difference

HSRC
Human Sciences
Research Council

Equivalence

- ✓ Equivalence is associated with the measurement level issues, that is the extent to which the scores of different cultural and language groups are comparable.
- ✓ There are various levels of equivalence, construct, measurement unit and scalar or full comparability, where scalar is the highest form of equivalence.
- ✓ Equivalence of an instrument is a pre-requisite in any cross-cultural comparison.

Bias

- ✓ Bias, on the other hand, indicates the presence of nuisance factors that challenge the validity of cross-cultural comparisons and therefore threatens equivalence.
- ✓ Three types of bias, namely, construct, item and method bias.
- ✓ Therefore if any one or all, of the three types of bias is found, the scalar equivalence is compromised.
- ✓ Essentially, equivalence cannot exists when bias is found.

Framework of Equivalence

- ✓ Theorists such as Poortinga and Van de Vijver have made a clear theoretical link between test bias and test equivalence.
- ✓ Therefore this can be translated into the Framework of Equivalence.
- ✓ This framework assists with the conceptualisation of a comprehensive evaluation of tests for the use across groups, regardless of whether the tests are monolingual or available in more than one language.

Aim and Objectives of the Study

In order to obtain the overall aim the researcher has outlined four specific objectives.

Overall Aim and Specific Objectives

✓The overall aim of this study was to evaluate the equivalence of the adapted English version of the WMLS for use across English first-language and Xhosa first-language learners.

1.To evaluate group differences in terms of the mean scores of the test between English-and Xhosa-first language speakers on the English version of the WMLS.
2.To evaluate group differences in terms of reliability of the test between English-and Xhosa-first language speakers on the English version of the WMLS
3.To evaluate group differences in terms of the item characteristics of the test between English-and Xhosa-first language speakers on the English version of the WMLS.
4.To evaluate the item bias, across English-and Xhosa-first language speakers on the English version of the WMLS.

Design of the Study

Social science that makes a difference

HSRC
Human Sciences
Research Council

Design of the study

- Exploratory, Quantitative Study.
- Secondary Data Analysis

✓ **Procedure**

- The researchers of the main study, received permission from the Eastern Cape Dept. of Education.
- Made contact with principals, distributed consent forms.
- Ethical Clearance was received from the Ethics committee of the NMMU (Old UPE) in 2005.
- Honours Students
- Data Collected 2006/2007

✓ Sampling

As outlined by the researchers of the main study, sampling consisted of convenience purposive sampling.

Purposive sampling

- Allowed the researcher to control for confounding factors:
  - an equal number of participants,
  - Equal, Males and Females,
  - Educational background (grade 6 and 7).

Convenience Sampling

- Evaluation of the sample.

✓ Participants

- Consisted of 198 English first language learners and 197 Xhosa language learners.
- Both groups were selected from ex Model C and Previously Disadvantaged Schools.

✓ Measurement Tool

- WMLS was used
- Consisted of 4 subtests, Picture Vocabulary, Verbal Analogies, Letter-word Identification, Dictation.
- Psychometric Properties

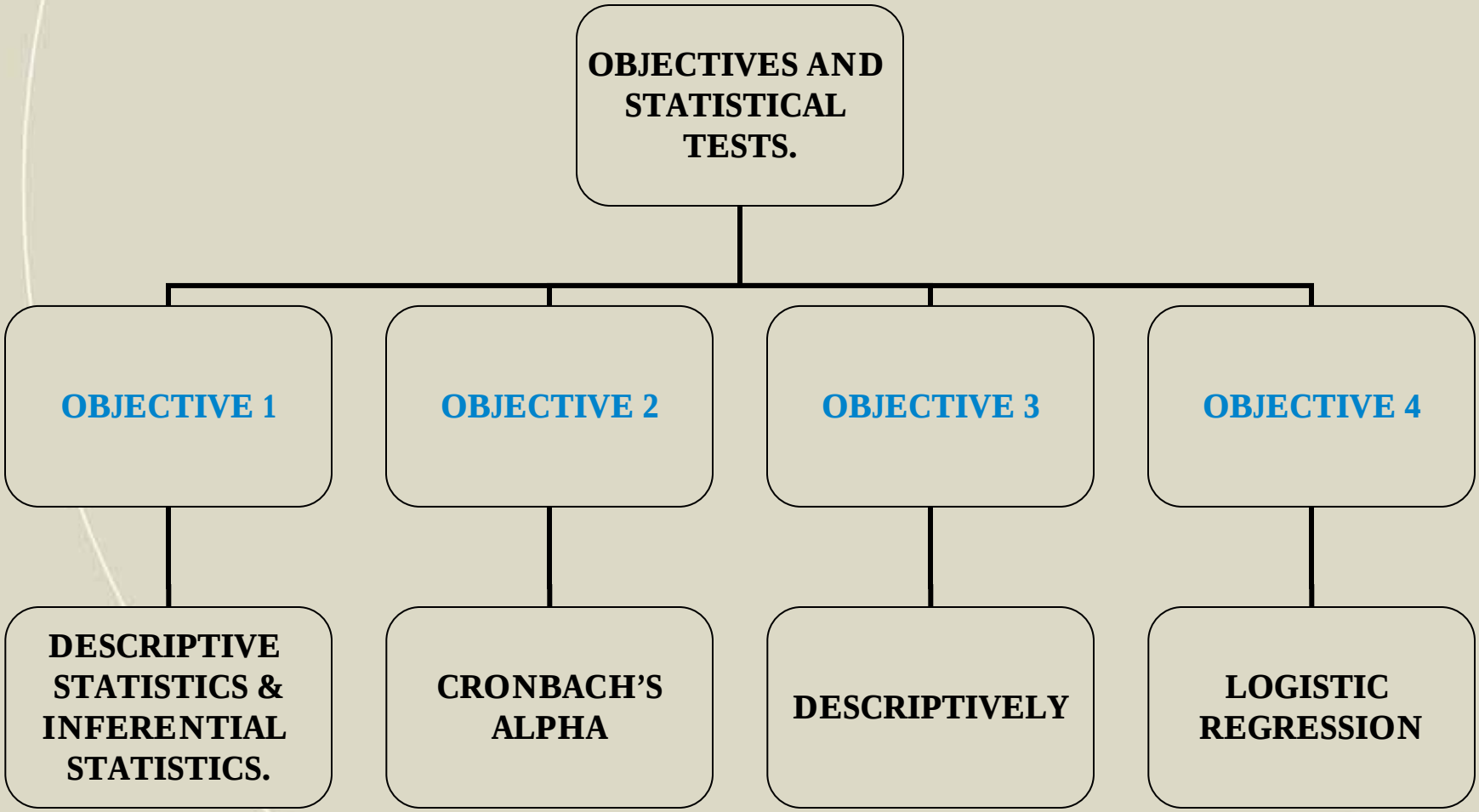
  Two important issues to be considered when selecting a test for research purposes, Reliability and Validity.

Reliability-for the USA version, internal consistency reliability coefficients and standard errors of the mean were calculated for all the English forms and clusters across their range of untended use.

Validity-for the USA version, was evaluated on content, concurrent as well as construct validity.

✓ The researcher of the main study made all the provisions of the necessary ethical considerations required by the NMMU (old UPE).

Data Analysis

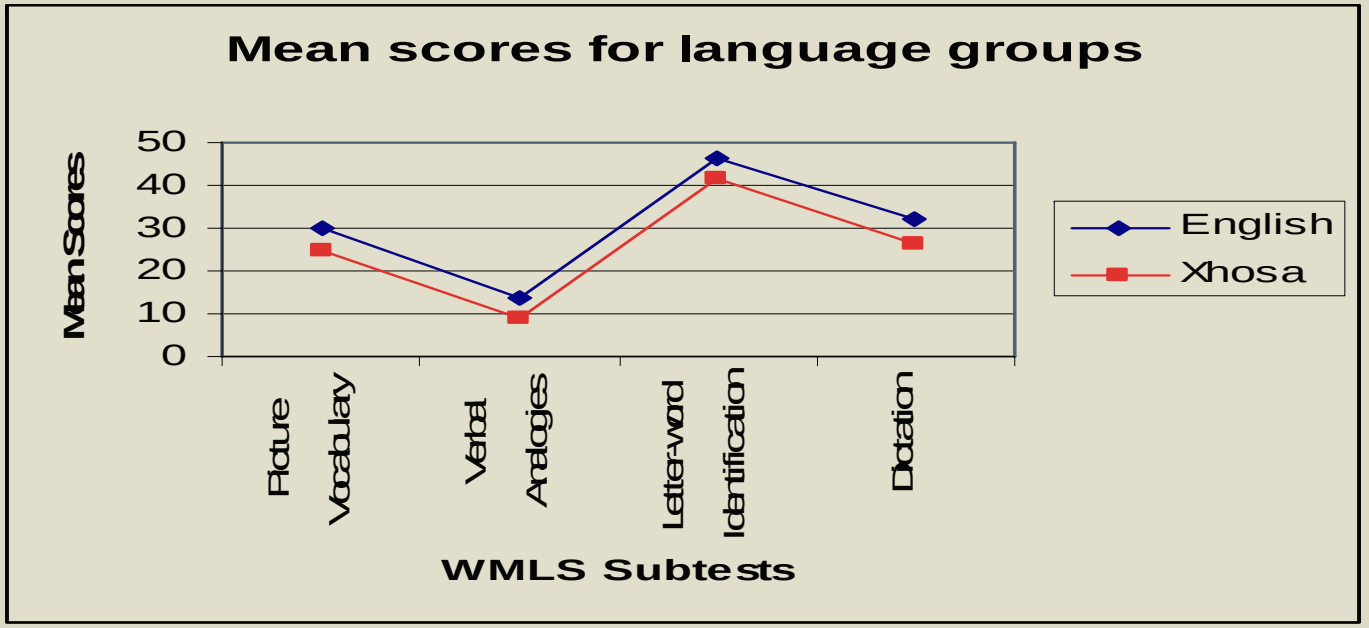


Results and Discussion

HSRC
Human Sciences
Research Council

Objective 1

Graph 1: Mean scores for each language group on each subtest



✓The above table indicates that the overall performance of the English first language group was higher than that of the Xhosa first language group.

✓A Hotellings T^2 Post-hoc tests was conducted to determine the degree of the difference and this test indicated that the overall differences between the groups are statistically significant.

Objective 2

Table 1: Cronbach's Alpha for each language group on each subtest

Language Group	Cronbach's Alpha			
	Picture Vocabulary	Verbal Analogies	Letter-word Identification	Dictation
English	0.73	0.83	0.89	0.82
Xhosa	0.81	0.86	0.90	0.90

✓Cronbach's Alpha used to measure the internal consistency (Reliability) of each subtest of the WMLS. 0.80s to 0.90s-good reliability

✓Xhosa first language speakers- overall higher reliability-English first language speakers

✓In addition were compared on their equality of reliability by using the following statistic (1-alpha1)/(1-alpha2).

✓These differences needed further investigation, as at this stage the findings do not conclude that the WMLS is equivalent for both the English first language speakers and the Xhosa first language speakers.

Objective 3

Table 2: Mean Item characteristics' (item difficulty with standard deviations and item discrimination)

Language Group	Picture Vocabulary			Verbal Analogies			Letter-word Identification			Dictation		
	Item dif.	Std Dev.	Item Dis.	Item dif.	Std Dev.	Item Dis.	Item dif.	Std Dev.	Item Dis.	Item dif.	Std Dev.	Item Dis.
English	0.53	0.14	0.13	0.39	0.31	0.32	0.82	0.20	0.26	0.58	0.20	0.19
Xhosa	0.43	0.15	0.15	0.25	0.26	0.34	0.73	0.29	0.34	0.46	0.26	0.30

✓indicate various differences between the English first language learners and the Xhosa first language learners.

✓These differences can be attributed to various reasons, such as differences in ability levels between the two language groups, possible cultural differences or various other bias (such as item bias).

✓Therefore further analysis was conducted to determine where these differences arise from.

Objective 4

- ✓ Objective 4 was evaluated by means of logistic regression. Defined as the prediction of the probability of occurrence of an event by fitting data to a logistic curve. In this case it was the probability of scoring 1 on an item which was due to ability only.

- ✓ With regard to the results, it was observed that Item bias or Differential Item Functioning (DIF) exists for all of the subtests (certain items) and within each language group.

- ✓ In other words the probability of scoring 1 was not due to ability only.

Major Findings and Conclusion

- ✓ The major findings indicates that there were overall differences between the English first language group and Xhosa language group on their mean scores (objective 1), their reliability displayed on each subtests for both groups (objective 2), there are significant differences between the language groups on the item characteristics of each subtests (objective 3) and Differential Item Functioning (item bias) was displayed on all of the subtest .

- ✓ Due to these significant findings it can be concluded that the scores of the adapted English version of the WMLS *cannot* be used for comparison across the two languages groups with all the items included in the scores.

- ✓ The scores cannot be regarded as displaying scalar equivalence across the for English first language speakers and Xhosa first language speakers.

- ✓ The DIF items with acceptable item characteristic within group, may be used in making statements about the construct within that group.

The importance of this study is that it demonstrates the need to evaluate the equivalence of monolingual language tests for use across language groups!!!


Thank you
Social science that makes a difference
HSRC
Human Sciences
Research Council