Assessing the impact of culture on Theory of Mind performance in South African Adults

Alice Lazarus
ACSENT Laboratory
Department of Psychology
University of Cape Town

ABSTRACT

Theory of Mind (ToM) refers to the ability to recognize other people's mental states. Most ToM tests used in research have been developed in the UK and US and are designed for, and have been predominantly tested on, Western, English-speaking populations. As no studies have yet been conducted on ToM ability in normal South African adults it is still unclear to what extent culture, language and ethnicity play a role on ToM ability in normal adults. It is unknown whether these UK/US developed ToM tests can measure ToM ability equally across the South African population. Using the 3 most widely accepted ToM tests, *Strange Stories, Reading the Mind in the Eyes* and *Faux Pas,* the aims of this study were to ascertain whether ToM performance of South African Students from different cultural groups differed from each other and whether they differed from the UK/US populations. This was part of a larger study which tested ToM ability across all the SADC languages. This study compared two distinct South African cultural groups, White English first-language speakers and Black Xhosa/Zulu first-language speakers, in their performance on these 3 ToM tests to assess whether normal South African students from the two different ethnic/language backgrounds differ in their results. 37 English and 28 Xhosa/Zulu UCT undergraduates were tested. Results indicated that both groups did not attain the ceiling effects expected of normal UK and US adults. After adjusting for intelligence and executive functioning, ToM scores of the Xhosa/Zulu group were still significantly lower than those of the English group. Several cultural biases in the tests were then identified and discussed.

Keywords: Theory of Mind, cross-cultural assessment, language, South Africa, adults

INTRODUCTION

In the last three decades, interest and investigation in Theory of Mind (ToM) has increased at a tremendous rate. Research on ToM in a number of different populations has grown rapidly in the almost 30 years since the term was first used in Premack and Woodruff's seminal publication "*Does the chimpanzee have a theory of mind*?" (1985). The vast majority of research in this young field of study has been pioneered in Western countries with the United Kingdom and the United States at the forefront of ToM research. As with all neuropsychological performance, ToM is influenced by both organic and environmental variables – which includes cultural aspects (Horton Jr., Carrington, & Lewis-Jack, 2001). It is for this reason that ToM tests developed in Western countries should be used with caution on non-Western populations. As the field of ToM research expands in South Africa, these Western ToM tests need to be assessed to guarantee that they are accurately and fairly measuring ToM ability across our culturally and ethnically diverse population.

Theory of Mind (ToM)

Theory of Mind refers to the ability to infer other people's mental states – beliefs, desires, emotions, intentions, knowledge, etc. – with the understanding that these mental states may be different from one's own (Baron-Cohen, Leslie, & Frith, 1985). This realisation that other people can want, feel and believe things is crucial to the understanding that these inner experiences result in and are manifested in human action (Wellman, Cross, & Watson, 2001). ToM encompasses the ability to correctly determine the intentions, desires and beliefs of others and to thus predict and/or explain their behaviours based on those mental states (Marjoram et al., 2006). Furthermore, ToM explains one's ability to understand that others' mental representations of the world do not necessarily reflect reality and that people can believe something to be true when in reality it is not so (i.e. they may have false beliefs). It is this understanding that allows others to manipulate people through lies, sarcasm and exaggerations. ToM is an ability used mainly for socializing with others and is believed to be an outgrowth of social intelligence (Abu-Akel, 2003). It is the aspect of social cognition that sets humans apart from animals and even primates. ToM is believed to have developed in humans as a means for complex social interactions as well as for internal introspection.

ToM in Human Development

ToM is believed to be an innate normal human ability and the focus of the majority of ToM research has been on abnormal or atypical ToM development. ToM impairment describes a difficulty someone would have with mental perspective-taking, and is otherwise known as mind-blindness. This means that individuals with a ToM impairment have a hard time seeing things from any other perspective than their own. Mind-blindness has been used to explain deficits in autism. For this reason research in the field of autism constitutes a large section of all ToM Research (Baron-Cohen, Leslie, & Frith, 1985; Baron-Cohen et al. 1999). However, ToM impairment is not specific to autism. ToM deficits have been studied in both negative and positive symptom schizophrenia as well as in people with right hemisphere damage (Abu-Akel, 2003; Marjoram et al. 2006; Surian & Siegal, 2001). Although innate and automatic, ToM ability is thought to require social and other experiences over many years to bring successfully to full development in adolescence.

Traditionally, the definitive test of ToM ability is the false-belief task (Wimmer & Perner, 1983). One of the first and most important milestones in ToM development is gaining the ability to attribute false belief – i.e., recognising that others can have beliefs about the world that are wrong in that their mental states differ from reality. This ability is acquired by children at around 4 years of age. At this age explicit ToM processing becomes possible and children are able to identify and explain examples of false-belief (U. Frith & C. D. Frith, 2003). There are two levels of false belief understanding: first-order false belief tasks (e.g. Sally thinks x, when really it's y) or second-order belief tasks (e.g. Sally thinks that Mary thinks x, but both Sally and Mary are wrong). At age 6 a child should be able to pass second-order tasks (Baron-Cohen et al., 1999). The classic scenarios used for testing false-belief is the 'unexpected transfer' or 'changed location' task in which a protagonist thinks an object or person is in a location when it has in fact been moved (e.g., Sally puts chocolate in a cupboard. John moves it to the drawer. Where will Sally look for the chocolate?). Another classic scenario is the 'appearance-reality' task in which something which is expected is shown to be false (e.g., finding crayons in a Smarties box when one expected Smarties) (Abu-Akel, 2003). However, children do have some understanding of others' mental states before the age of 4. Pretend play and joint attention, which develop in infants at around 18 months of age, are viewed as precursors to ToM ability as they show

implicit attribution of mental states (Frith & Frith, 2003). By the age of 3 most children can refer to their own and others' mental states, at least in terms of desire.

The majority of ToM literature and studies have focused on the development of ToM in children, particularly the acquisition of false belief. This leaves the impression that ToM does not develop beyond the age of 6 years (Wellman et al., 2001). However, ToM continues to develop until at least 11 years of age (Baron-Cohen et al., 1999; Abu-Akel, 2003). More advanced tests of theory of mind have been developed for adults with the three most widely accepted being: Happé's (1994) *Strange Stories* task, which involves the detection of sarcasm, irony, bluff and double-bluff; Baron-Cohen et al.'s (1999) *Faux Pas* task, which involves detection of socially awkward scenarios; and Baron-Cohen et al.'s (1997) *Reading the Mind in the Eyes* task, which involves detecting both basic and complex mental states from the information provided by people's eyes.

Universality of ToM development and Cross-cultural Assessment

There is disagreement amongst researchers as to whether there is a universal developmental trajectory in ToM ability. Certain researchers believe that different people may develop more, or less, effective theories of mind, varying from very complete and accurate ones, through to barely functional (Wellman et al., 2001). Recently more and more studies have been conducted on non-western populations in order to confirm the claim of universality in ToM development. These studies have had mixed results (U. Frith & C.D. Frith, 2003; Liu et al. 2008). For instance, Callaghan et al. (2005) showed that children from Samoa, Canada, Thailand, Peru and India did not differ significantly in time of onset of ToM (at around 5 years of age). Delayed onset was attributed to low socio-economic status, rather than cultural factors. Other factors found to influence ToM onset include aspects of family background, especially parental occupational class and education level (Cutting & Dunn, 1999). Wellman *et al.* (2001), in their meta-analysis, concluded that age of onset as well as the developmental trajectory of ToM is universal and comparable across cultures. People from all cultures reach all the ToM milestones at the same rate until complete adult ToM ability is attained. However, on the performance on false belief tests the Western countries and cultures did better than the others. This could mean that the false belief tests being used to assess ToM ability/acquisition are better suited to some cultures than others.

There are many other studies which have found discrepancies in the universal age of onset for ToM. For instance, Chasiotis, Kiessling, Hofer, and Campos (2006) found that children from Cameroon showed delayed development of ToM in comparison to children from Germany, America and Costa Rica. Contextual variables such as socioeconomic status, number of siblings and parenting style were found to influence the development of ToM abilities. The later onset of ToM in Cameroon was attributed to Cameroon's characteristically authoritarian parenting style which Chasiotis et al (2006) claim may lead to increased inhibition and less conversation exposure in children. Contrary to this claim of inhibition being detrimental to ToM, Liu *et al.,* (2008) found that cultures in which impulse control skills were deemed admirable had better false belief test results as test takers inhibited the automatic urge to answer what is correct or true and were better able to focus on the false belief or what is not correct. These cultures are generally more collectivist and non-Western.

All the above mentioned studies used ToM tests developed either in the UK or the US on both the Western and Non-Western populations. A universalist conception of ToM would argue that the measurement of ToM should be achieved with the same instruments in different cultures (Shuttleworth et al., 2004). The variance in ToM performance between cultures could be accounted for by the Western designed tests which do not take Non-Western environmental and socio-cultural factors into account.

ToM assessment in multicultural South Africa

Unlike other cognitive constructs, such as intelligence, there is no standard test battery for determining ToM ability. Most ToM tests used in research have been developed in the UK and US and are designed for, and have been predominantly tested on, Western, middle-class, English-speaking populations. These tests adopt a dichotomous approach, where if one attains full marks on the test one has a normal ToM ability and if one fails the test one has a ToM impairment (Abu-Akel, 2003). Although many tests have been developed that test for the presence of various aspects of ToM, such as false belief, emotion recognition, faux pas recognition, etc., there is no consensus as to which tests definitively measure overall ToM (or, indeed, if any set of tests can provide such a definitive measurement). Furthermore, there are no set standards or norms for ToM scores as no test has yet been normed on any population. What typically happens is that individual scores on ToM tests are compared and contrasted to the

general trend observed in the development of ToM seen in studies of predominantly White, middle-class participants. It makes sense, therefore, that a South African who is not White or English speaking or middle-class would have a misleading ToM score in these Western-designed ToM tests. Furthermore, ToM tasks rely on social norms – especially the ones that use stories or vignettes. As social norms vary from culture to culture, this raises concerns about using these ToM tasks in South Africa where the population is so culturally diverse. Because tests developed using Eurocentric approaches cannot be indiscriminately used with individuals who differ from a normative population (Padilla, 2001), it is questionable whether the existing ToM tests are appropriate for use in a multi-cultural and multi-lingual country such as South Africa.

South African Society and Culture

Culture refers to the shared beliefs, customs, folkways, worldviews, values, traditions and behaviours held by a particular group of individuals (Horton Jr., Carrington, & Lewis-Jack, 2001; Wallis & Brit, 2003). South Africa has been referred to as the 'rainbow nation', a title that is meant to epitomize the country's cultural diversity. South Africa's population is one of the most complex and diverse in the world. There are eleven official languages in South Africa: nine African languages, English, and Afrikaans. Although Zulu and Xhosa have the highest population of native speakers, at 24% and 18% of the population respectively, English is the official language of administration and is spoken throughout the country (Elion & Strieman, 2001). Although South Africa's Black majority still has a substantial number of tribal, rural inhabitants who lead largely impoverished lives, Black South Africans are becoming increasingly urbanised and Westernised, and usually speak English or Afrikaans in addition to their native tongue (Suzuki e al. (2001); Wallis & Birt, 2003).

A further complexity in the South African context is that an ethnic group may, but equally may not be homogenous in terms of socio-cultural characteristics (Shuttleworth-Edwards et al., 2004). Many Black South Africans identify themselves as part of their distinct tribe/ethnic group but also identify themselves as urbanised, Westernised members of a greater South Africa. This can be seen in the rapid movement of previously disadvantaged people to urbanized Western conditions.

Additionally, there is a disparity between the standard of education offered to White individuals compared to Black individuals within South Africa. During Apartheid, segregated

and inferior education (to that offered to White people) was enforced upon the Black population. Although currently many socially advantaged Black individuals receive the same quality of education as white individuals, the majority of the Black population still receive substandard education. Thus, quality of education in South Africa differs substantially not only across the ethnic groups but also within the Black ethnic groups.

Issues in Culturally Appropriate Neurocognitive Assessment in South Africa

When testing ToM abilities in the South African population, it is important to ensure that the differences in scores attained from different cultural groups reflect differences in ToM and not differences in, for instance, cultural factors or language comprehension. Therefore, a critical question is this: If a ToM test is to be standardised for the South African population, should each cultural, ethnic, racial, or language group have its own separate norms, or should there be one normative dataset for the entire population? Separate standardisation data would probably ensure more accurate descriptions of the individual's abilities, with his/her results being compared to an appropriately narrow normative sample. However, having separate measures for different cultural, ethnic, racial, or language groups may promote misunderstanding and harmful misinterpretation. For instance, if one ethnic group were to attain lower scores than another group on a certain cognitive test, one might be tempted to make dangerous genetic/biological interpretations (i.e., one ethnic group has a higher ability than the other) instead of interpreting lower results for an ethnic group as a problem with the test (Manly, 2005).

Research has continuously indicated that scores on tests of all types are consistently lower for individuals who differ from the normative population (Padilla, 2001). Without proper standardization it is highly problematic to test cognitive ability from one ethnic group to another. The potential for unfair discrimination against non-native English speakers in South Africa cannot be overlooked when English neurocognitive assessments are used for diagnostic or placement processes (Shuttleworth-Edwards et al., 2004; Wallis & Birt, 2003).

RATIONALE FOR RESEARCH

The literature clearly indicates that in the relatively new research field of ToM there is as yet no single test battery used to universally establish a person's overall ToM ability. Furthermore no ToM test been standardized on any given population. For these reasons it is still unclear to what

extent culture, language and ethnicity impact on ToM ability in normal adults and whether these Western-developed ToM tests can measure ToM fairly across the South African population. Furthermore, no studies have yet been done on ToM ability in normal South African adults, and it is currently unknown whether typical South African adults across all 11 official languages will display the same level of ToM in these Westernised English tests. We currently do not even have general characteristics of ToM performance across the different South African cultures.

The three most commonly used Adult ToM tests are: *Strange Stories* (Happé, 1994), *Reading the Mind in the Eyes* (Baron-Cohen et al., 1997) and *Faux Pas* (Baron-Cohen et al., 1999). Like all ToM tests these tests were developed for research, as opposed to clinical assessments. All research assessing ToM in adults thus far has involved studies of impaired individuals (i.e. with deficits resulting from brain damage or psychiatric disorders) with healthy normals used as controls. A trend found in these Western studies has shown that normal adults (in UK and US populations) get ceiling effects on the *Faux Pas* and *Strange Stories* tests and that only the *Reading the Mind in Eyes* test is hard enough to avoid ceilings in normals. However, in similar South African studies it has been observed that the healthy control adults do poorly in all three tests (S. Malcolm-Smith, personal communication, April 12 2009). This trend needs to be further investigated to ascertain whether South Africans do, in fact, perform worse than the UK controls in these ToM tests and to determine whether different cultural groups in the South African population attain different scores in these three ToM tests. This study is the first part of a continuing research project that aims to address these issues. This initial study looks specifically at the difference in ToM results between White English speakers and Black Xhosa or Zulu speakers.

SPECIFIC AIMS/ HYPOTHESES

This study aims to look at the performance characteristics on the three ToM tests in healthy neurologically normal South African students who are either English first language speakers or Xhosa/Zulu first language speakers. In particular, I aim to ascertain:

(1) Whether normal South African students from different cultural backgrounds differ in their results on the *Strange Stories, Reading the Mind in the Eyes* and *Faux Pas* ToM tests.

(2) Which (if any) elements of these three tests are problematic for each culture group.

(3) How these tests can be modified so as to control for these problems so that they are unbiased and fair to all South African ethnic groups.

METHODS

Research design and setting

This comparison of different South African cultures' ToM ability focused primarily on two groups: Those who are first-language English speakers and those who are first-language Xhosa or Zulu speakers. As Xhosa and Zulu are the two most widely spoken first languages in South Africa and as these two African cultures and languages are very similar to each other they comprised a single group in the study of non-English speakers. Ultimately all the South African cultural groups will be studied and stratified by SES, quality of education, etc. This study is simply the first exploratory phase of the research. Language is being used as a proxy for cultural group. The study is a cross-sectional quasi-experimental design of two existing groups. The two groups were compared on level of ToM attained by comparing the score obtained on each of three UK/US developed ToM tests. In addition, a qualitative questionnaire will qualitatively assess which parts (words, scenarios, concepts, instructions) of the test each group found difficult, confusing or inappropriate, and how the participants feel the tests can be changed to the best advantage of their culture group. Testing took place at the UCT Department of Psychology. Participants were tested in a quiet room free of distractions in a single two hour session.

Participants

I recruited 65 undergraduate Psychology students through the Student Research Participation Program (SRPP) to participate in the study. The students were separated into two groups based on their first language - English first-language speakers (n=37) and Xhosa/Zulu first-language speakers (n=28). The two groups also consisted of different ethnic backgrounds – White English speakers and Black Xhosa/Zulu speakers. Recruiting by race was discouraged by the SRPP system as ethically problematic - hence the use of the home-language proxy. The English speakers served as a reference group as they are likely to have the most similar cultural backgrounds to the British and American populations the tests were designed for. The two groups of students from the two different South African cultures are homogenous to the extent

that they are fluent in English, are educated to a similar degree and have achieved the same level of education, and have exposure to western thinking and testing. There are also marked differences between the White Western South African culture and Black Nguni South African culture. The Xhosa/Zulu students who make up the Xhosa/Zulu group are ideal for this exploratory study as they should be able to articulate any cultural inappropriateness they see in the tests. Students from both language groups who underwent the ToM tests could thus fluently explain to me how they thought the tests could be changed to best reflect their cultural background. Although it is assumed that all students studying at a tertiary institute should have above average cognitive skills, the participants' IQ and executive function were empirically assessed to ensure that the two groups were relatively equal. Participants were all roughly matched on age, sex and level of education.

Table 1. *Demographic information for the English speaking and Xhosa/Zulu speaking groups.*

Demographic Information	English (n=37)	Xhosa/Zulu (n=28)
Age Range (Years)	18 - 25	18 -44
Age (Years)		
Mean (SD)	19.49 (1.56)	21.25 (5.45)
Sex		
Male: Female	2: 35	2: 26

The study followed the ethical guidelines set out by the Research Ethics Committee of UCT as well as the Health Professions Council of South Africa (HPCSA). Ethical approval for this study was obtained from the Ethics Committee of the UCT Department of Psychology (Reference #2008007). All participants gave their informed consent before being tested by signing a consent form which described the test battery they were to undergo, assured them of the anonymity and confidentiality of their results, and inform them of their right to withdraw from the study at any point.

Inclusion and exclusion criteria

Exclusion criteria included students under the age of 18, students whose first language is one other than English, Zulu or Xhosa, students who are English speaking but who have not lived in South Africa for most of their lives. All participants had to understand and speak English fairly

fluently, which is also an expected condition for studying at UCT. Due to the complexity of South African society, some Black South African's have English as a home language – these South Africans are fluent in both English as well as an African language and are very Westernised. For the purposes of this study, in order to attain two distinctly different cultural groups, such participants were excluded from participating. The English speakers group were all White participants as to avoid the different cultures of non-White participants (i.e., Coloured or Indian participants) that may have an effect on the ToM scores. Of all the participants actually tested only one was excluded from the study: She has an African American mother and a Xhosa father, spoke both Xhosa and English at home to both parents and had moved to South Africa from the States at age 4. After being tested as a Xhosa participant her results were excluded from the study as I felt that her culture and English fluency was not distinctly different enough from either of the two language groups. Interestingly, as the only partly American participant she was also the only participant that attained full marks on the ToM tests.

Measures

Cognitive ability

Because critics often claim that ToM is not a distinct ability, but is based either on general intelligence or general executive function, it is important to account for these when looking at ToM ability (Bibby & McDonald, 2005). The WASI (The Wechsler Abbreviated Scale of Intelligence; Wechsler, 1999) was used to measure general intelligence across the languages. It has been normed and standardized for people between the ages of 6 and 89. The WASI includes four subtests, two of which are verbal (vocabulary and similarities) while two test performance (Block Design and Matrix Reasoning), to ascertain verbal IQ (VIQ) and performance IQ (PIQ) scores respectively. The two verbal tests include defining words and describing similarities between two related concepts. The two performance tests include visio-spatial tasks such as copying designs and identifying patterns.

Executive Function was measured by two tasks from the Delis-Kaplan Executive Function Scale (D-KEFS; Delis, Kaplan, & Kramer, 2001) – The Colour-Word Interference Test and the Sorting Test. The Colour-Word Interference Test assesses cognitive flexibility by both requiring the participant to inhibit automatic reading of words denoting colours while naming the colours themselves, and then subsequently having the participant switch back and forth between

naming the discordant ink colour and reading the conflicting word. The Sorting Test assesses problem-solving behaviour as well as abstraction ability by having the participant sort cards into two groups, with 3 cards per group, according to as many different concepts or rules as possible.

All these tests, including the ToM tests, were developed in, and will be administered in English. Using the WASI and D-KEFS to measure intelligence and executive function across the culture groups can be seen as ethically problematic. This is because these tests have not been standardised on a South African population and, as with most verbal neuropsychological tests, may be biased against non-English speakers. However, I wanted to make sure that IQ and executive function and their effects on ToM performance were controlled for in the participants of the 2 groups. I could not solely rely on the assumption that all UCT students have average or above average IQs. As the WASI and the D-KEFS are both well accepted measures of intelligence and executive functioning these were the tests I chose to ascertain that there were no gross differences between the English speaking and Xhosa speaking groups. The results of these two tests will, however, be analysed with caution.

Theory of Mind

Three ToM tasks were administered: *Strange Stories* (Happé, 1994), *Reading the Mind in the Eyes* (Baron-Cohen et al., 1995) and *Faux Pas* (Baron-Cohen et al., 1999). These three tests are the most common tests used in ToM research in adult populations.

Strange Stories consists of 16 stories of which half involve mental states such as - lie, white lie, joke, pretend, double-bluff, persuasion, forgetting, misunderstanding, figure of speech, appearance-reality, irony and contrary emotions. Participants read and memorise each of the 16 stories separately after which they are asked a question. There are 8 ToM stories in which the participants have to infer and describe the character's mental state. The other 8 control stories involve making physical inferences (i.e., not mental). The control stories help to ascertain whether someone has a genuine ToM deficit. If a person performs poorly on the ToM questions but correctly answers the control questions it indicates a specific deficit in inferring mental states.

Reading the Mind in the Eyes, contains 36 images of eyes portraying different complex emotions or mental states. The participant must select the word, out of four options per picture, that best describes what the person in the picture is thinking or feeling. This task relies on

processing subtle difference around the eyes to infer mental states – ranging from more basic emotions (e.g., upset, friendly) to complex and subtle mental states (e.g., anticipating, confident, reflecting). A definition handout that describes all the words and gives examples of them in a sentence is given to every participant in case they are not familiar with certain words. The definition handout supposedly cancels out the effects of the language/culture bias as, with it, a participant can understand all the words presented whether or not they had encountered them before the task.

The *Faux Pas* test, contains 10 control scenarios depicting a normal social event, and 10 test scenarios wherein a character says something awkward or embarrassing. After a story is told the participant is asked, "Did anyone say something they shouldn't have said or something awkward?" If he/she responds 'yes' questions follow which ask the participant to infer both the characters beliefs/knowledge, their intentions and the emotional reactions of the other character. There are also 2 very simple control questions after each story to make sure that the participant has understood the situation presented. The test aims to see if a participant can spot a socially awkward interaction and whether s/he can infer to the mental states of others to explain the awkwardness. Unlike the *Strange Stories* test and the *Reading the Mind in the Eyes* test, the *Faux Pas* test requires the participant to infer more than one mental state (for each of the 2 or 3 characters in each story). Not only does the participant need to show an awareness that the character who said the faux pas did not intend and was not aware of the awkwardness caused but also that the person hearing the faux pas would be offended or embarrassed by what was said. Thus the task contains both a cognitive and an affective component - a participant should be able to spot a faux pas, explain the awkward situation, and be able to empathise with the characters to infer their emotions.

Procedure

The study was advertised on the SRPP notice board to recruit participants. An email account was created to keep track of and communicate with the participants (theorymind.uct@gmail.com). Each participant had one testing session of 2 to 2.5 hour duration, during which the IQ, Executive function and ToM tests were administered. The participant could indicate at any point during the tasks that they wanted a break if they felt fatigued.

The length of each session was dependent on the speed at which the participant read the stories in ToM tasks and understood them. Due to the very long duration of the session the testing order was counterbalanced with half the participants being tested for IQ and Executive function first and half being tested on ToM first to control for fatigue effects. This particular study only compared the results of White English first language speakers with Black Xhosa/Zulu first language speakers as part of a larger study testing all the SADC languages/cultures.

Once the participant was at the testing location s/he was given a consent form to read and sign and a chance to ask questions after which testing immediately commenced. The three ToM tests were administered exactly as they are without any modifications and they were marked according to the author guidelines (Baron-Cohen, Jolliffe, Mortimore, & Robertson,1997; Baron-Cohen et al., 1999; Happé, 1994). This enabled me to see exactly which aspects of each test need to be changed so as to be more applicable to the South African population. Furthermore, once the ToM tests had been administered the participants were asked a series of qualitative questions to ascertain which aspects of the tests were biased against them and their culture and how they believe the tests can be modified to best reflect their cultural background (for the list of qualitative questions see Appendix A).

Data Analysis

I looked at IQ, executive function and ToM performance of the two cultural groups using descriptive statistics. Along with the scores for the different ToM tests I created an overall ToM score for each participant by adding their scores on the Strange Stories ToM stories, Reading the Mind in the eyes score and Faux Pas ToM score. I compared IQ, executive functioning and ToM ability in White English speakers and Black Xhosa/Zulu speakers using independent t-tests with language as the independent variable and the scores as separate dependent variables. Correlations between total ToM score and IQ and executive function scores were established. Then analyses of covariance were done with VIQ, PIQ, FSIQ, Inhibition/Switching and Free sorting Confirmed Correct score as potential covariates to ToM to check whether 'test taking ability' effected ToM scores. Additionally, I conducted further t-tests on counterbalanced groups with 'order of tests' (i.e, WASI and D-KEFS first or ToM first) as the independent variable and ToM score as the dependent variable.

My two language groups were small and of unequal sample size. In addition, the assumptions on normality and heterogeneity of variance were not met for a few of the scores. However, using non-parametric analyses would reduce my statistical power. There was also a high possibility of making type-I errors due to the large number of tests done. To compensate between lack of statistical power and the risk of having less reliable results, I used an alpha level of 0.01 as the threshold for statistical significance in all my analyses. The statistical analyses were performed using STATISTICA version 8 (StatSoft, Inc, 2007).

RESULTS
ToM

T- tests were conducted to investigate the difference between ToM scores between the White English speaking grouping and the Black Xhosa/Zulu speaking group.

Table 2. *Descriptive statistics for performance on the basic ToM battery.*

Measure	English	Xhosa/Zulu	Group Differences		Effect Size
	(*n*=37)	(*n*=28)	*t*	*p*	*d*
Strange Stories ToM	14.05 (2.15)	11.86(2.46)	3.83616	0.000292	0.95
Strange Stories control	13.03 (1.95)	9.54 (3.31)	5.32319	0.000001	1.28
RME	27.78 (6.43)	21.25 (5.19)	5.04902	0.000004	1.26
Words looked up	2.22 (2.89)	10.14 (8.81)	-5.13176	0.000003	1.20
Faux Pas ToM stories	51.46 (6.16)	45.29 (6.70)	3.85383	0.000275	0.96
Faux Pas control stories	9.05 (1.18)	8.07 (2.32)	2.22552	0.029633	0.53
Faux Pas control questions	39.62 (0.64)	38.11 (2.54)	3.48729	0.000895	0.82
No. Faux Pas correct	9.68 (0.85)	9.28 (0.90)	1.78690	0.078765	0.44
ToM combined score	93.30 (9.18)	78.39 (11.50)	5.81331	0.000000	1.43

Note: Means are presented with standard deviations in parentheses.
RME= Reading the Mind in the Eyes

Strange Stories

Xhosa/Zulu speakers scored significantly lower than English speakers on the ToM questions in the Strange Stories task, ($t_{(63)}$= 3.84, *p*=0.0003, *d*=0.95). They also scored significantly lower

than the English speakers in the Control questions of the Strange Stories task, ($t_{(63)}$= 5.32, p<0.001, d=1.28) with an even larger significance.

Reading the Mind in the Eyes

In the Reading the Mind in the Eyes task Xhosa/Zulu speakers got significantly lower scores than English speakers, ($t_{(63)}$= 5.05, p=0.000004, d=1.26); along with a significantly higher number of words looked up in the definition handout, ($t_{(63)}$= -5.13, p=0.000003, d= 1.20). A test for correlation between number of correct emotion readings and number of words looked up yielded a moderate negative correlation, (r=-0.48, p=0.000064, r^2=0.23), showing how for both languages as Reading the Mind in the Eyes score increased the number of words looked up decreased.
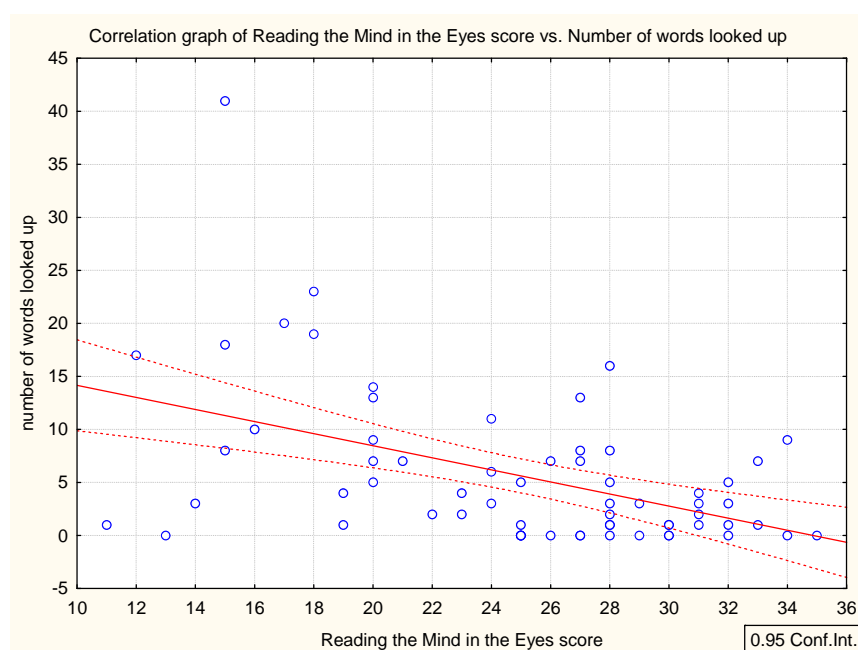
Figure 1: Correlation graph of Reading the Mind in the Eyes score vs. number of words looked up

Faux Pas

Xhosa/Zulu speakers got significantly lower scores on the total score for Faux Pas ToM stories than English speakers, ($t_{(63)}$= 3.85, p=0.0003, d= 0.96). However, the number of Faux Pas recognised by English speakers (m=9.68, SD=0.85) and Xhosa/Zulu speakers (m=9.28, SD=0.90) was not significantly different, ($t_{(63)}$= 1.79, p=0.08, d= 0.44), indicating that although individuals from both language groups are able to spot the faux pas the Xhosa/Zulu group have more

difficulty explaining the faux pas correctly. Similarly, there was no significant difference in the Faux Pas control stories, ($t_{(63)}$= 2.23, p=0.03, d= 53), indicating that both English speakers and Xhosa/Zulu speakers have equal capacity for spotting when no faux pas occurs. However, a comparison of overall number of faux pas recognised (m=9.51, SD=0.89) and overall number of correct control stories (m=8.63, SD=1.82) indicated that the number of faux pas recognised was significantly higher than the control examples, ($t_{(128)}$=3.50, p=0.0006, d=0.62). Xhosa/Zulu speakers had significantly lower scores for the Faux Pas control questions than the English speakers, ($t_{(63)}$= 3.49, p=0.0009, d= 82) indicating that Zulu/Xhosa people understood the stories less than the English speakers.

The overall ToM score was significantly higher for English speakers (m=93.30, SD=9.18) than for Xhosa/Zulu speakers (m=78.39, SD=11.50), ($t_{(63)}$= 5.81, p<0.001, d= 1.43). Overall ToM score had the largest effect size of 1.43 which is a very small.
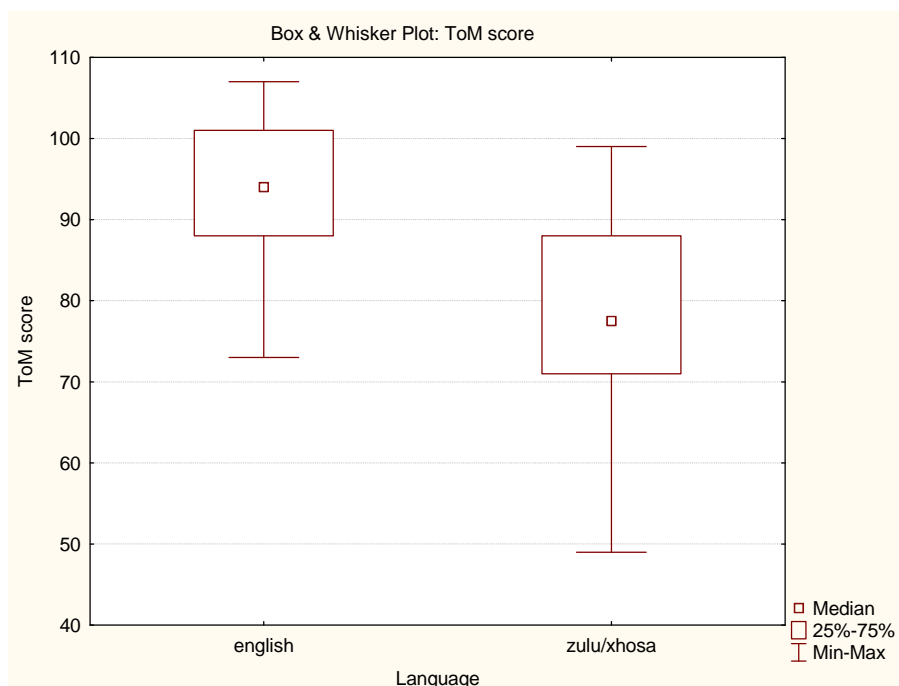

Figure 2: Box & Whisker plot of total ToM scores for the two language groups

The difference in ages between English speakers (M=19.49, SD=1.56) and Xhosa/Zulu speakers (M=21.25, SD=5.45) was not significant, ($t_{(63)}$= -1.8, p=0.07). Therefore age was not a factor leading to differences in ToM between the two groups.

WASI and D-KEFS

Black Xhosa/Zulu speakers had significantly lower scores for all VIQ, PIQ as well as FSIQ than White English speakers (see table 3) with VIQ, ($t_{(63)}$= 11.34, $p<0.001$, d= 2.85); PIQ, ($t_{(63)}$= 8.47, $p<0.001$, d= 2.03) and FSIQ, ($t_{(63)}$= 12.03, $p<0.001$, d= 2.94). The effect sizes for the IQ scores are small. However, it is important to note that the IQ scores of both language groups fall within average range, despite the statistical differences between the scores of the two groups.

In the D-KEFS Colour-Word Interference test The Xhosa/Zulu group attained significantly lower scores for both the inhibition task, ($t_{(63)}$= 5.76, $p<0.001$, d= 1.43), as well as the inhibition/switching task, ($t_{(63)}$= 4.69, p=0.000015, d= 1.14). They also had significantly lower scores in the Sorting test for the free sorting confirmed correct score, ($t_{(63)}$= 5.53, $p<0.001$, d= 1.40), the free sorting description score, ($t_{(63)}$= 5.57, $p<0.001$, d= 1.43), the sort recognition description score, ($t_{(63)}$= 4.12, p=0.000112, d= 1.05), and the combined description score, ($t_{(63)}$= 5.27, p=0.000002, d= 1.35). Overall, English speakers did significantly better in the IQ and executive function tests. These findings are consistent with the literature on Western psychometric tests on Non-Western populations.

Table 3. *Descriptive statistics for performance on the WASI and D-KEFS.*

Measure	English	Xhosa/Zulu	Group Differences		Effect Size
	(*n*=37)	(*n*=28)	*t*	*p*	*d*
VIQ	114.51 (9.32)	88.75 (8.72)	11.34480	0.000001	2.85
PIQ	109.41 (7.96)	84.86 (15.10)	8.46608	0.000001	2.03
FSIQ	113.89 (7.72)	85.68 (11.18)	12.03129	0.000001	2.94
D-KEFS inhibition	11.30 (2.00)	8.25 (2.25)	5.76064	0.000001	1.43
D-KEFS inhibition/switching	11.32 (1.83)	8.61 (2.83)	4.69175	0.000015	1.14
Free sorting confirmed correct score	10.62 (2.45)	7.36 (2.21)	5.53621	0.000001	1.40
Free sorting description score	10.27 (2.60)	7.07 (1.80)	5.57020	0.000001	1.43
Sort recognition description score	9.92 (3.24)	6.86 (2.55)	4.12135	0.000112	1.05
Combined description score	10.11 (2.96)	6.57 (2.25)	5.26885	0.000002	1.35

Note: Means are presented with standard deviations in parentheses.

Statistical Control of potentially confounding variables

Before deciding which subtest scores to run analysis of covariance on I first worked out the correlation between the IQ and executive function scores to the total ToM score. This was to ascertain how strong the relationship is between them as well as how strongly they predict ToM scores (See table 4).

Table *4. Correlations for test performance and ToM score*

Measure	*m* (*n=65*)	*r* (**X,Y**)	r^2	*t*	*p*
VIQ	103.42 (15.69)	0.59	0.35	5.77	< 0.000001
PIQ	98.83 (16.79)	0.57	0.32	5.46	0.000001
FSIQ	101.74 (16.87)	0.63	0.39	6.38	<0.000001
D-KEFS inhibition	9.98 (2.59)	0.43	0.18	3.78	0.000350
D-KEFS inhibition/switching	10.15 (2.66)	0.51	0.26	4.64	0.000018
Free sorting confirmed correct score	9.22 (2.85)	0.46	0.21	4.11	0.000117
Free sorting description score	8.89 (2.78)	0.44	0.20	3.92	0.000218
Sort recognition description score	8.60 (3.32)	0.34	0.12	2.91	0.005024
Combined description score	8.58 (3.19)	0.43	0.18	3.73	0.000412
ToM Score	86.88 (12.59)	1.00	1.00	-	-

Note: standard deviations are presented in parenthesis

Although all the scores were significantly positively correlated with ToM scores it was VIQ, PIQ and FSIQ scores as well as the D-KEFS inhibition/switching score and the Free sorting confirmed correct score that had the strongest significant correlation to ToM score. These five scores were chosen for analysis of covariance (ANCOVA) to determine whether there would still be a difference in ToM performance between the groups had both groups been equal in IQ and executive function scores (See Appendix C for ANCOVA tables).

Neither VIQ nor PIQ independently had a significant effect on ToM performance. FSIQ, however, did account for some of the variance in ToM performance ($F_{(1,62)=}6.15$, $p=0.02$). However, the effect of culture on ToM performance remained significant indicating that this result cannot be entirely accounted for by the impact of FSIQ ($F_{(1,62)=}74.48$, $p<0.01$, $\eta^2=.55$) Free Sorting Confirmed Correct score also accounted for some of the variance in ToM performance ($F_{(1,62)=}2.17$, $p=0.15$). The effect of culture on ToM performance remained

significant indicating that this result cannot be entirely accounted by the impact of that executive function ($F_{(1,62)}$=13.23, p<.01, η^2=.09). However, Inhibition/Switching accounted for all of the variance in ToM ($F_{(1,62)}$=5.88, p=0.02). The effect of culture on ToM performance was not significant indicating that this result can be accounted for in part by the impact of Inhibition/Switching($F_{(1,62)}$=6.23, p=0.02, η^2=.18). This indicates that the executive function of inhibition/switching may be a potential covariate of ToM performance and had both language groups attained the same score in the Inhibition/Switching task there would not have been a significant difference between the two groups.

Counterbalancing

To check for fatigue effects on ToM results I ran an independent t-test with pooled variance (Levene=0.86) with order of testing as the independent variable and ToM score as the dependent variable. Assumptions of normality and homogeneity of variance were upheld. Results indicate no significant interaction effect between order of testing and ToM score, $t_{(63)}$=0.03, p=0.98, signifying that there is no difference in ToM score whether an individual gets tested for ToM first (*m*=86.83, *SD*=13.30) or undergoes the hour long WASI and D-KEFS first (*m*=86.91, *SD*=12.14).

DISCUSSION

ToM

The ToM test battery comprising of the *Strange Stories* task, the *Reading the Mind in the Eyes* task and the *Faux Pas* task have all been developed for a Western US and UK population. Any normal English speaking adult from the UK or US should be scoring full marks for the *Strange Stories* and *Faux Pas* task and high scores for the *Reading the Mind in the Eyes* task. However this was not the case for the South African sample used in this study.

Strange Stories

On average the White English group got 14.05 out of 16 ToM stories correct whereas the Black Xhosa/Zulu group only got 11.86 correct. Importantly, both English and Xhosa/Zulu speakers scored below the expected 100%. The scoring criteria for the 8 ToM questions described that 2 points be given if the participant made reference to a mental state or intention, 1 point if the participant made reference to the outcome or facts without any reference to the protagonist's

mental state and intention and 0 points if the answer was incorrect or irrelevant. The authors' instruct that each story be read and memorised after which the page the story is on must be turned so that the participants may not refer back to it when answering the questions. In the English group errors were due to not explicitly mentioning a mental state. An example would be in the sausages story (Appendix D.1) where the participant is asked "Why does Brian say this?" they would generally answer that it is because he is greedy or because he wants more sausages without making reference to his intention to manipulate or mislead the lunch. These answers are correct but show no ToM. The Xhosa/Zulu group on the other hand would often get answers incorrect due to them not understanding the story or forgetting details or difficulty in explaining their answers in English.

Both the English group (m=13.03) and the Xhosa/Zulu group (m=9.54) did worse than UK/US controls in the Strange Stories control stories with a larger significant difference between the two South African groups (ToM p=.000178 vs. Control p=.000027). The problem both groups experienced was remembering the details of the stories once they turned over the page. Answering the ToM questions required an intuitive understanding of the protagonist's mental state. The control stories' answers involved remembering detailed facts and information. Often there was more than one potential answer to the question but only one answer which the author deemed worth 2 points, for example, in the story where John buys a multipack of light bulbs (Appendix D.2) the answer that it's cheaper to buy light bulbs in multipacks was worth 2 points. However, most participants answered that it is more convenient to have many light bulbs for use in the future, which is only worth 1 point. A particular control story that both English and Xhosa/Zulu speakers found difficult was the story of the new book in the Library (Appendix D.3). Many participants explained that she didn't shelve the book on medicinal plants in either the botany or medicine sections of the library because they were equally relevant to both so should be placed in a separate section. This answer, although reasonable, was not eligible for any points. The correct answer was that the book was in delicate condition due to age and was in a separate section of the library under controlled temperature to preserve it. When asked, the participants admitted to not knowing that libraries had these temperature-controlled rooms for special status books. This question may be biased for South Africans in general.

There were certain cultural biases in many of the stories against non-Western cultures and many words and constructs that most Xhosa/Zulu participants had never encountered before

(such as artillery, electronic detector beam and meringues) and could therefore not understand the stories they were in. Many participants confused long-sightedness with short-sightedness (Appendix D.4) and therefore could not understand a particular story. In the story in which Henry makes meringues (Appendix D.5) not only did almost all the Xhosa/Zulu participants not know what meringues were but they therefore also didn't know that they were made from egg whites and that mayonnaise was made from egg yolks. This question was the most biased against the Xhosa/Zulu group in the Strange Stories task.

Reading the Mind in the Eyes (RME)

Of the three ToM tasks RME showed the largest significant difference (p=.000007) between the two language groups with the English group getting on average 27.78 (SD=6.43) correctly identified emotions and the Xhosa/Zulu group getting 21.25 (SD=5.19) emotions correct. In addition, the Xhosa/Zulu group looked up a significantly larger number of words in the definition handout compared to the English group. Furthermore, the significant moderate negative correlation between RME scores and number of words looked up (r=-0.48) indicates that the Xhosa/Zulu group did far worse than the English group because they could not understand the English words used and thus were unable to identify the emotions they represented in the eyes. These results imply that RME is first and foremost a vocabulary test and only secondly and ToM test.

Most Xhosa/Zulu participants explained that often the emotions/mental states that were in the options to choose from did not exist in the Xhosa/Zulu vocabulary and in their culture. Occasionally there were emotions that existed in their culture and language that they could see in the eyes but a word option was not present as the concept did not exist in Western culture. Although the definition handout was meant to put everyone on equal footing with regards to understanding what all the words meant, it did not, describe or help with what that specific mental state looks like on a face. For example, to someone who has never encountered the word "aghast" before or even the concept of being "aghast" it is almost impossible to then try and spot it in a stranger's eyes.

Xhosa/Zulu participants explained that in Xhosa/Zulu culture it is rude and disrespectful to look a person directly in the eyes and that usually emotions and mental states are read through entire facial expressions and not through what can be seen in eyes alone. This cultural bias in the

task may explain their poor performance. However, Baron-Cohen *et al.*,(1997) created this task with only the eye region specifically as people with autism and Asperger syndrome were able to recognise emotions from the whole face. So although this test is a more appropriate measure for research in autism it is very biased against normal Xhosa/Zulu adults. Furthermore, the 36 eyes in the test belonged to white people with the exception of a few Hispanic eyes which could also have negatively influenced the Xhosa/Zulu results.

Many participants from both language groups admitted to guessing some of the examples as they did not feel any of the words matched the emotions they could see in the eyes. When asked what they thought could be changed in the task a majority replied that apart from making the words easier one should also be able to see the entire face and not just the eyes. Baron-Cohen et al., (1996) found that in a cross-cultural study in which participants had to recognise mental states and emotions in facial expressions (note: not just the eyes) found that a large range of subtle and complex mental states are universally read in the face. However, even in this study there were certain 'cognitive' mental states such as 'recognise' and 'distrust' that were deemed private and unobservable. Many of these purely cognitive mental states such as 'distrustful' are examples in the RME test (Baron-Cohen et al., 1997). RME is meant to be a pure measure of ToM ability as it is not language based or anecdotal. However, the results show that not only is RME culturally biased it is also very much based in language.

Faux Pas

The 10 Faux Pas ToM stories were each marked out of 6 with 1 point allocated for identifying that someone is saying something awkward or something that shouldn't be said and then 1 point for each of the subsequent 5 questions making up a total of 60 points. Overall, there was a statistical difference between the ability to correctly identify faux pas and non-faux pas situations with the participants identifying faux pas more easily than the control situations. The Xhosa/Zulu group (m=45.29) got significantly lower scores for the faux pas stories than the English group (51.46) however there was no significant difference between the number of faux pas that each language group identified correctly. This means that the Xhosa/Zulu participants are able to identify when someone is saying something awkward or something that shouldn't be said just as well as the English participants. However the Xhosa/Zulu participants are unable to describe and explain the faux pas as well the English participants. There was no significant difference

between both cultural groups in the control stories. The Black Xhosa/Zulu groups could identify non-Faux pas situations as well as the White English Group.

Of the 6 questions that followed each story the most problematic for both languages was question 4 – "Why do you think he/she said it?" this question leads on from question 3 – "Why shouldn't he have said it or why was it awkward?" Whereas in question 3 the participant's answer should make reference to the faux pas without necessarily referring to the mental states of the characters the answer for question 4 has to specifically indicate that one of the characters in the story didn't know something or realise something. Often participants would combine both answers in question 3 to explain the faux pas as not knowing/realising something is a crucial factor in a faux pas and thus answer question 4 with the character's motivation for their action/statement (Appendix E.1). Question 5 then goes on to ask directly if they thought the protagonist was aware of his/her faux pas. As the questions for all the stories are the same the participant becomes familiar with them. The participant may think that as question 5 is asking them directly about the protagonist's awareness of the faux pas that there is no need to mention it in previous questions.

This means that although the participant might understand the situation perfectly he/she is getting marked wrong for saying the correct thing in the incorrect place. In the future, for a South African population, I would recommend a new marking scheme which deviates from the rigid correct answer per question system that Baron-Cohen et al (1999) have specified. Each story should be worth 6 points with a point allocated for each aspect of explaining the faux pas no matter which question it gets mentioned in.

In the Faux Pas test all the stories involve an awkward situation/scenario but it is only in the ToM stories that the awkwardness is caused by something that one of the characters has naively said. However, often the participants would state that someone is saying something awkward in a control story and then proceed to explain the awkwardness either by confabulating malicious subtexts to a character's otherwise non-awkward statement or by just explaining why the scenario is awkward. The first story in the Faux Pas caused the highest amounts of misinterpretations in both groups (Appendix E. 2). Participants often claimed someone was saying something awkward in this control story. This could possibly be that because it is the first story, the participants have not yet been exposed to the more awkward ToM stories and feel a desire to explain away the awkwardness.

There were many stories that had distinctly American cultural references and words. Although the English group were familiar with these terms from watching American media many of the Xhosa/Zulu participants had never encountered them before. This made understanding the story extremely difficult, especially if the unfamiliar concept played a vital role. In one ToM story a little boy is in a 'stall' instead of 'cubicle' of a public toilet (Appendix E. 3). Many Xhosa/Zulu participants did not understand what a stall was and thus were not aware that the little boy was in the bathroom with the other boys who were gossiping about. The story with the most bias against the Xhosa/Zulu group was the petrol station control story (Appendix E. 4). The story refers to filling a car up with gas instead of petrol and paying in dollars instead of rands. In the story John fills his own car up with petrol, and then goes inside to pay the cashier, as is standard in the U.S. This is unlike the South African system in which a petrol attendant fills your car for you and then you pay them while remaining in your car throughout the process. However, most Xhosa/Zulu participants were unaware of the U.S. system and could not fathom what exactly was going on, but, being sure that something odd was happening they would pick at some aspect of the story as a Faux Pas, e.g. John shouldn't have gone indoors to the cashier and not paid the attendant, The cashier should not have filled John's car, the cashier put in the wrong petrol for the car, John knew he had no money on his card but was trying to waste the cashier's time etc.

Often, Xhosa/Zulu participants would confuse the names of the characters, especially if there were more than two of them. Even though they understood the story they would get scores deducted for giving wrong names. This could potentially be avoided by the inclusion of traditional Xhosa/Zulu names amongst the all English names. Furthermore, typically South African scenarios and places could be included so that South Africans can relate to them better and envision the situations more clearly.

IQ and executive function's influence on ToM
As expected the Xhosa/Zulu group did significantly worse for VIQ, PIQ and FSIQ than the English group. Although the WASI (Weschler, 1999) is a well accepted measure of intelligence it has not been standardised on a South African population. This U.S. designed test is thus also biased against non-Western populations. As ToM ability has been attributed to intelligence a measure of intelligence had to be conducted to ensure the groups did not vary significantly in this

ability, which could influence ToM performance. It is important to note that although most scores from both groups did fall within average range the IQ scores did differ significantly. Had the test been normed on the full South African population I assume that the difference in IQ scores between the two language groups would be far smaller. ToM is viewed to be most attributed to verbal intelligence due to the textual/narrative format of the ToM tests (Milligan, Astington, & Dack, 2007). However, the correlation between VIQ, PIQ and FSIQ and overall ToM score were very similar. Generally PIQ is seen to be the least culturally biased measure of intelligence due to its non-reliance on language. However, the difference in PIQ scores between the language groups although smaller than the difference between their VIQ scores was still very significantly different. Although the 3 IQ scores were moderately correlated with total ToM score separate ANCOVAs showed that ToM scores would have still been significantly different between the two groups had IQ been equivalent. Even when controlling for VIQ a significant difference in ToM ability remained indicating that VIQ is not the only factor affecting the significant difference in ToM score between the two culture groups. The Xhosa/Zulu group's English verbal ability was poorer than that of the English group due to English being a second language. As with the VIQ subtests all three of the ToM tasks have high language demands. Without good English linguistic skill it is more difficult for a Xhosa/Zulu speaker to describe higher-order representations and complex mental states (Foxcroft & Aston, 2006).

Inhibition/Switching, was found to account for some variability in ToM performance implying that this executive function may play a role in ToM ability. ToM processing in general involves high-level executive functioning. Lesion studies found that patients with damage to the medial frontal cortex had deficits in ToM ability (Gallagher & Frith, 2003). However, these patients also had difficulties with traditional executive function tasks and their ToM deficit may have been influenced by an overall executive function deficit. It is possible, that higher cognitive abilities such as inhibition/switching, remembering information, focusing attention, comprehending and answering questions are required to complete the ToM tasks and these processes involved in the ToM task may be correlating with the Inhibition/Switching score. Overall, it appears the ToM performance is not entirely attributable to intelligence or executive functioning and the differences in the ToM scores of the two groups are due to cultural bias and not cognitive differences

Limitations and future directions

Due to time constraints as an honours project this study had small sizes and unequal groups. Testing more participants to create larger and equal sample sizes will allow for parametric testing which would increase the statitistical power and reliability of the results. As many statistical tests were conducted, the possibility of a type-I error in the analysis is there even with my adjusted alpha. More males from both language groups need to be tested as there may be differences in ToM ability between sexes. My current participants are almost only female. None of the participants' socio-economic statuses were ascertained and if SES does effect ToM scores then the two groups could potentially be unmatched. As I was recruiting UCT students through the SRPP program matching groups on SES for this study was restricted as I would have been more likely to find more low SES Xhosa/Zulu participants than English participants within the UCT system. The three ToM tests should be adapted based on the problems observed in this study and then the study should be re-run with the adapted ToM tests. If both language groups improve and if there is no significance between the two groups' performances (or a drastic reduction in significance) it will prove the cultural bias of the Current three ToM tests. This exploratory study specifically looked at White English speakers and Black Xhosa/Zulu speakers from tertiary level education. The long term project seeks to test ToM performance on South Africans from all culture groups and stratified for factors found to influence neurocognitive testing such as SES, level of education, quality of education, ethnicity, sex, acculturation effects etc. Once ToM performance for all South African groups is ascertained a culturally unbiased South African ToM battery can be created.

CONCLUSION

ToM is believed to be an innate and universal human ability. However, South Africans perform poorly on ToM tests that are designed to identify the ToM impaired. As normal adults (without autism or brain damage) should be attaining ceilings on these tests, doing poorly signifies a problem with the test as opposed to a problem with the population. This study took two samples from separate and distinct South African cultures - the more Western English speaking participants and Xhosa/Zulu speaking participants from the non-Western Nguni culture - and tested them on these U.K. and U.S. designed ToM tests. The purpose was to uncover where and why South Africans were struggling in them. Both groups performed lower than the U.K. and

U.S. ceiling effects. The Xhosa/Zulu group did significantly worse than the English group in all the ToM tests and it was found that the tests were especially biased against them. The adaptation of the ToM tests is essential in a multicultural and multilingual society like South Africa if test results are to be valid and reliable for all test takers. ToM tests used in this country should accommodate the different languages and cultural backgrounds of the participants by being adapted in such a way that they retains the original principle while also featuring culturally-relevant contexts and examples. Creating South Africa appropriate ToM tests is critical for meaningful research that incorporates both clinical and normal populations.

REFERENCES

Abu-Akel, A. (2003). A neurobiological mapping of theory of mind. *Brain Research Reviews*, *43*, 29-40.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition, 21,* 37-46.

Baron-Cohen, S., Ring, H. A., Wheelwright, S., Bullmore, E.T., Brammer, M.J., Simmons, A. et al. (1999). Social intelligence in the normal and autistic brain: an fMRI study. *European Journal of Neuroscience*, 11, 1891-1898.

Baron-Cohen, S., O'Riordan, M., Stone, V., Jones, R., & Plaisted, K. (1999). Recognition of faux pas by normally developing children and children with Asperger syndrome or high-functioning autism. *Journal of Autism and Developmental Disorders*, *29*, 407-18.

Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind: Evidence from very high functioning adults with autism or Asperger Syndrome. *Journal of Child Psychology and Psychiatry, 38,* 813-822.

Baron-Cohen, S., Riviere, A., Fukushima, M., French, D., Hadwin, J., Cross, P., et al. (1996) Reading the mind in the face: a cross cultural and developmental study. *Visual Cognition*, *3*, 39-60.

Bibby, H., & McDonald, S. (2005). Theory of mind after traumatic brain injury. *Neuropsychologia ,43*, 99–114

Callaghan, T., Rochat, P., Lillard, A., Claux, M. L., Odden, H., Itakura, S., Tapanya, S., Singh, S. (2005). Synchrony in the onset of mental-state reasoning: evidence from five cultures. *Psychological Science*, *16*, 378-84.

Chasiotis, A., Kiessling, F., Hofer, J., & Campos, D. (2006). Theory of mind and inhibitory control in three cultures: Conflict inhibition predicts false belief understanding in Germany, Costa Rica and Cameroon. *International Journal of Behavioral Development, 30*, 249-260. doi: 10.1177/0165025406066759.

Cutting, A. L. & Dunn, J. (1999). Theory of Mind, emotion understanding, language, and family background: individual differences and interrelations. *Child Development, 70*, 853-865.

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan Executive Function system examiner's manual*. San Antonio: The Psychological Corporation.

Elion, B. & Strieman, M. (2001).*Clued up on culture – a practical guide for all South Africans*. Cape Town: Juta Gariep Publishing Company Ltd.

Foxcroft, C. D., & Aston, S. (2006). Critically examining language bias in the South African adaptation of the WAIS-III. *SA Journal of Industrial Psychology, 32*, 97-102.

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transaction of the Royal Society of London. B, 358*, 459-473.

Gallagher, H. L., & Frith, C. D. (2003). Function imaging of 'theory of mind'. *TRENDS in Cognitive Sciences, 7*, 77-83.

Greif, A. (1994). Cultural beliefs and the organization of society: A historical and theoretical reflection on collectivist and individualist societies. *Journal of Political Economy*, 1*02*, 912 - 925.

Happé, F. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders*, *24*, 129-154.

Horton Jr., A.M., Carrington, C.H., & Lewis-Jack, O. (2001)*. Handbook of multicultural assessment: clinical, psychological, and educational applications*. Chapter 5 – Neuropsychological Assessment in a Multicultural Context. San Francisco : Jossey-Bass

Liu, D., Wellman, H.M., Tardif, T., & Sabbagh, M.A. (2008). Theory of mind development in Chinese children: A meta-analysis of false – belief understanding across cultures and languages. *Developmental Psychology*, *44*, 523-531.

Manly, J. J. (2005). Advantages and disadvantages of separate norms for African Americans. The *Clinical Neuropsychologist, 19*, 270-275.

Marjoram, D., Miller, P., McIntosh, A.M., Cunningham Owens, D.G., Johnstone, E.C., & Lawrie, S. (2006). A neuropsychological investigation into 'theory of mind' and enhanced risk of schizophrenia. *Psychiatry Research*, *144*, 29-37.

Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development*, *78*, 622-646

Naito, M., & Koyama, K. (2006). The development of false-belief understanding in Japanese children: Delay and difference? *International Journal of Behavioral Development*, *30*, 290-304.

Nell, V. (1999). Standardising the WAIS-III and the WMS-III for South Africa: legislative, psychometric, and policy issues. *South African Journal of Psychology*, *29*, 128-137.

Padilla, A.M. (2001). *Handbook of multicultural assessment: clinical, psychological, and educational applications*. Chapter 1 – Issues in Culturally Appropriate Assessment. San Francisco: Jossey-Bass

Premack, D. G., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1,* 515-526

Shuttleworth-Edwards, A. B., Kemp, R. D., Rust, A. L., Muirhead, J. G. L., Hartman, N. P., & Radloff, S. E. (2004). Cross-cultural effects on IQ test performance: A review and preliminary normative indications on WAIS-III test performance. *Journal of Clinical and Experimental Neuropsychology, 26*, 903-920.

Surian. S., & Siegal, M. (2001). Sources of performance on theory of mind tasks in right hemisphere-damaged patients. *Brain and Language, 78***,** 224–232

Suzuki, L.A., Short, E.L., Pieterse, A., & Kugler, J. (2001). *Handbook of multicultural assessment: clinical, psychological, and educational applications*. San Francisco : Jossey-Bass

Wallis, T., & Birt, M. (2003). A comparison of native and non-native English speaking groups' understanding of the vocabulary contained within 16PF (SA 92). *South African Journal of Psychology, 33*, 182-190.

Wechsler, D. A. (1999). *Wechsler Abbreviated Scale of Intelligence manual*. San Antonio: The Psychological Corporation.

Wechsler, D. A. (1997). *Wechsler Adult Intelligence Scale-III.* San Antonio: The Psychological Corporation.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-Analysis of theory-of-mind development: The truth about false belief. *Child Development*, *72*, 655-684.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*, 41-68

APPENDIX A

Qualitative Questions

1) Did you find any of the instructions and/or questions confusing?

2) Where there any words that were new to you?

3) How difficult or easy were these tests? Which was the most difficult?

4) Were these stories easy to understand? Why/why not

5) Were there any situations in the stories that were new to you?

6) Do you think the test was relevant to your culture group?

7) Did the stories represent culturally appropriate social situations?

8) Were any of the interactions culturally inappropriate?

9) How could this test be changed to be made more appropriate for your culture group?

APPENDIX B

Consent Form

INFORMATION AND INFORMED CONSENT FORM:

INVESTIGATION OF THEORY OF MIND IN SOUTH AFRICAN POPULATIONS

PURPOSE:

This study is part of a research project investigating Theory of Mind in South Africa. Theory of Mind is regarded as a universal human ability to think about other people's mental states, eg how they feel, what they know, intend, desire, etc. Theory of Mind development has been studied in various cultures, but this is the first study being conducted in South Africa. Many of the standard tasks developed to measure Theory of Mind abilities were developed in the UK, so it is important to establish if South Africans think in similar ways, or if the tasks need to be modified for our context.

STUDY PROCEDURE:

If you decide to participate, we will ask you to attend a session in the ACSENT lab in the Psychology Dept. The tasks will take 120 to complete, so you will get 4 SRPP credits for participating in this study. You may be asked to return for a short focus group interview session with other students to discuss the tests and your perception of them (for extra SRPP points).We will ask you to provide some demographic information about yourself (gender; home language).

The Theory of Mind tasks will involve pen and paper or verbal responses. The tasks involve choosing which of 4 emotions best describes a displayed facial expression; deciding if a described social interaction is awkward; and deciding why people say or do things in short story segments. In addition, a few other short items will be administered – these will give a rough estimate of IQ and executive function.

CONFIDENTIALITY:

If you consent to participate in this study, your identity will be kept confidential. All research information will be safely stored and identified by code number and access will be limited to authorized scientific investigators. Any publications resulting from this study will not identify you by name.

VOLUNTARY PARTICIPATION:

Your participation in this study is voluntary and you may refuse to participate or withdraw from the study at any time.

RESEARCH QUESTIONS AND CONTACTS:

The researchers will answer any questions you might have about the procedures described above, or about the results of the study. If you have any questions, you may call Susan Malcolm-Smith on 021 650 4605. Or preferably e-mail Alice Lazarus or Margaret Mc Grath at theorymind.uct@gmail.com

INFORMED CONSENT:

I have read the above information, my questions have been answered, and I consent voluntarily to participate in this study.

Participant name: _____   Signature: _____

Date: _____

I have discussed the proposed research with this participant and, in my opinion; s/he understands the benefits, risks, and alternatives (including non-participation) and is capable of consenting to voluntary participation.

Print name: _____   Signature: _____
       Study Investigator

Date: _____

APPENDIX C

Analysis of Covariance tables

Table A. Analysis of covariance, with Verbal IQ as a possible covariant to ToM.

	SS	df	MS	F	p	Partial eta-squared	Non-centrality	Observed power (alpha=0.01)
VIQ	262.03	1	262.03	3.30	0.073929	0.05	3.30	0.43
Group	5389.79	1	5389.79	67.97	<0.000001	0.52	67.97	1.00
Error	4916.465	62	79.30					

Table B. Analysis of covariance, with performance IQ as a possible covariant to ToM

	SS	df	MS	F	p	Partial eta-squared	Non-centrality	Observed power (alpha=0.01)
PIQ	509.11	1	509.11	3.98	0.050478	0.06	3.98	0.501552
Group	4320.91	1	4320.91	33.77	<0.000001	0.35	33.77	0.999915
Error	7933.24	62	127.96					

Table C. Analysis of covariance, with FSIQ as a possible covariant to ToM

	SS	df	MS	F	p	Partial eta-squared	Non-centrality	Observed power (alpha=0.01)
FSIQ	498.39	1	498.39	6.15	0.01586	0.09	6.15	0.68
Group	6034.05	1	6034.05	74.48	<0.00001	0.55	74.48	1.00
Error	5023.28	62	81.02					

Table D. Analysis of covariance, with Inhibition/Switching as a possible covariant to ToM

	SS	df	MS	F	p	Partial eta-squared	Non-centrality	Observed power (alpha=0.01)
Inhibition/Switching	29.17	1	29.17	5.88	0.02	0.09	5.88	0.67
Group	30.92	1	30.92	6.23	0.02	0.09	6.23	0.69
Error	307.62	62	4.96					

Table E. Analysis of covariance, with Free Sorting Confirmed Correct score as a possible covariant to ToM

	SS	df	MS	F	p	Partial eta-squared	Non-centrality	Observed power (alpha=0.01)
Free Sorting Confirmed Correct Score	11.81	1	11.81	2.17	0.1456	0.03	2.17	0.31
Group	71.97	1	71.97	13.23	0.0005	0.18	13.23	0.95
Error	337.32	62	5.44					

APPENDIX D
Strange Stories examples

Brian is always hungry. Today at school it is his favorite (sic) meal - sausages and beans. He is a very greedy boy, and he would like to have more sausages than anybody else, even though his mother will have made him a lovely meal when he gets home! But everyone is allowed two sausages and no more. When it is Brian's turn to be served, he says, "Oh, please can I have four sausages, because I won't be having any dinner when I get home!"

Q: Why does Brian say this?

John is going shopping. He buys a nice new desk lamp, for his study. He needs a light bulb for his new lamp. He goes from the furniture department to the electrical department. In the electrical department he finds that there are two brands of light bulb of the right kind. Everbrite light bulbs cost less in single packs than Literite bulbs. However, only Literite bulbs come in multi-packs of six. John buys the multi-pack, even though he only needs one bulb.

Q: Why does John buy the Literite bulbs?

Mrs. Simpson, the librarian, receives a special book which she has to catalogue and find an appropriate place for. She has to decide which section to file it under. The library is very big, and has different sections on many different subjects. The new book is about plants and their medical uses, and is heavily illustrated. However, Mrs. Simpson does not put it on the shelf with the rest of the books on botany. Neither does she put it with the books on medicine. Instead, she carefully takes it into a separate room. In this room all the books are kept in special cases, and the temperature is kept constant.

Q: Why did she do this?

Henry is preparing for a big dinner party. He is famous for his excellent mayonnaise. He has bought lots of fresh eggs. The recipe says, "Carefully separate the yolks of six eggs and add oil very gradually". He has already bought easily enough dessert to feed everyone. However, he now looks up the recipe for meringues. Henry will not waste anything.

Q: Why does Henry make meringues?

APPENDIX E
Faux pas examples

1)Example of typical question 3 incorrect answer

Jeanette bought her friend, Anne, a crystal bowl for a wedding gift. Anne had a big wedding and there were a lot of presents to keep track of. About a year later, Jeanette was over one night at Anne's for dinner. Jeanette dropped a wine bottle by accident on the crystal bowl and the bowl shattered. "I'm really sorry. I've broken the bowl," said Jeanette. "Don't worry," said Anne. "I never liked it anyway. Someone gave it to me for my wedding."

Did anyone say something they shouldn't have said or something awkward? *yes*

If yes, ask:

Who said something they shouldn't have said or something awkward? *Anne*

Why shouldn't he/she have said it or why was it awkward?

She shouldn't have said it because Jeanette gave her the bowl as a gift and Anne had

forgotten that she had given it to her and said she didn't like it. It's offensive.

Why do you think he/she said it?

To comfort Jeanette

Did Anne remember that Jeannette had given her the bowl? *no*

How do you think Jeanette felt? *Insulted. Hurt*

2) Example of answer to the first control story in Faux Pas test

Vicky was at a party at her friend Oliver's house. She was talking to Oliver when another woman came up to them. She was one of Oliver's neighbors. The woman said, "Hello," then turned to Vicky and said, " I don't think we've met. I'm Maria, what's your name?" "I'm Vicky." "Would anyone like something to drink?" Oliver asked.

Did anyone say something they shouldn't have said or something awkward? *yes*
If yes, ask:

Who said something they shouldn't have said or something awkward? *Vicky*

Why shouldn't he/she have said it or why was it awkward?

It's not appropriate to introduce yourself so bluntly.

Why do you think he/she said it?

She is dominant. Maybe she is Oliver's girlfriend and does not like Maria talking to him so she's marking her territory.

Did Oliver know that Vicky and Maria did not know each other? *yes*

How do you think Vicky felt? *Down*

3) Mike, a nine-year-old boy, just started at a new school. He was in one of the stalls in the restroom at school. Joe and Peter, two other boys, came in and were standing at the sinks talking. Joe said, "You know that new guy in the class? His name's Mike. Doesn't he look weird? And he's so short!" Mike came out of the stall and Joe and Peter saw him. Peter said, "Oh hi, Mike! Are you going out to play football now?"

4) John stopped off at the gas station on the way home to fill up his car. He gave the cashier his credit card. The cashier ran it through the machine at the counter. "I'm sorry," she said, "the machine won't accept your card." "Hmmm, that's funny," John said. "Well, I'll just pay in cash." He gave her twenty dollars and said, "I filled up the tank with unleaded."