<u>Plagiarism Declaration</u>

1. I know that plagiarism is wrong. Plagiarism is to use another's work and to pretend that it is one's own.
2. I have used the <u>APA</u> convention for citation and referencing. Each significant contribution to, and quotation in, this essay from the work, or works, of other people has been acknowledged through citation and reference.
3. This work-piece is my own work.
4. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.

___N. V. Mlandu______________ _23/10/19___

Signature Date

Detecting Cultural Influences on Social Cognition: The South African-Adapted NEmo Test Battery

Nwabisa Vuyolwethu Mlandu

ACSENT Laboratory
Department of Psychology
University of Cape Town

Primary supervisor: Dr Kevin Thomas
Co-supervisor: Dr Victoria Ives-Deliperi
Word count:
Abstract: 235
Main body: 7456

Abstract

The psychological construct of social cognition comprises several distinct forms of mental processing that are essential for healthy interpersonal relations. Two separate and hierarchically inferior constructs, emotion recognition and theory of mind (ToM), are central to social cognition and have been the subject of intense neuroscientific study. Although these constructs are universal, numerous studies have shown that cultural-linguistic influences might affect expression of social cognitive abilities in these domains. Given the importance of intact emotion recognition and ToM for adaptive functioning, it is imperative for local research to describe possible cultural and linguistic influences on their expression and to ensure that tools used to assess them are contextually appropriate. In this study, I evaluated a South African-adapted version of the NEmo battery, a newly developed battery by the Swiss Epilepsy Centre. The tasks on the NEmo test battery were translated from their original German into English, and South African faces and voices replaced the original Swiss faces and voices as stimuli for the emotion recognition tasks. South African university students (N = 40; age M = 22.14 ± 3.84 years) completed the NEmo test battery and other standardized cognitive tests. English-speakers and Afrikaans-speakers performed significantly better on tasks assessing ToM than Xhosa-speakers. This suggests that culture influences the expression of this construct. Future research needs to ensure these tasks are adapted to Xhosa to ensure that social cognition is reliably measured in this language group.

Keywords**:** emotion recognition; theory of mind; cross-cultural; language; NEmo battery; social cognition.

Introduction

Healthy social interactions are an integral component of adaptive human functioning. The mental processes that govern these interactions are referred to by the umbrella term *social cognition*. This psychological construct is a discrete entity, dissociable even from closely related cognitive processes. These cognitive processes include executive functions, which are directly influenced by social cognition and which have a direct influence on it (Cohn, St-Laurent, Barnett, & McAndrews, 2015). Humans with intact social cognition are able to use these abilities to facilitate better interpersonal communication, to interpret and deliver social signals appropriately, to adjust to novel social demands, and, thereby, to contribute to overall wellbeing and a good quality of life (Jokeit, 2010). Those with impaired social cognition tend to struggle with decoding social cues and find it difficult to consider situations from different perspectives and to respond appropriately to novel social demands.

Although social cognition is a universal feature of human mental processing, numerous studies describe cross-cultural differences in its development and expression. Often, these differences are attributed to the languages associated with different cultures. Hence, it is of critical importance that researchers and clinicians who need to assess social cognitive functioning are armed with tools appropriate to their local linguistic and cultural contexts. The aim of this pilot study was to evaluate a South African adaptation of the NEmo test battery (Bauer, Steiger, Kegel, Spirig, & Jokeit, 2018), a newly developed assessment tool that measures two major components of social cognition (emotion recognition and theory of mind).

Social Cognition: Definition of construct and delineation of sub-domains

In healthy adults, the domain of social cognition comprises several distinct and well-developed forms of mental processing. These include emotion recognition, theory of mind, prosody, and body language interpretation (Henry, Von Hippel, Molenberghs, Lee, & Sachdev, 2016). Neuroscientific research has largely focused on the first two in that list, largely because (a) they are regarded as central to the construct, (b) deficits in these areas are frequently part of the behavioural manifestation of several neurological disorders, and (c) laboratory stimuli suited to their investigation are readily available.

Emotion recognition: Definition and cultural influences. The psychological construct of *emotion recognition* describes the ability to decipher the emotions other people are experiencing, expressing, and/or signalling. Recognising others' emotions enables one to learn about their emotional states and their intentions, and to navigate social spaces accordingly. Deficits in emotion recognition (i.e., an inability to recognize and understand social cues) have detrimental effects on mental health and overall wellbeing. This is primarily because such deficits hinder the affected person's capacity to establish the healthy interpersonal relationships that are often the basis of psychological, social, and financial security (Cohn et al., 2015; Henry et al., 2016; Jokeit, 2010; Nomura et al., 2005).

In social interactions, humans extract emotional information from several different sources, including facial expressions, prosody in verbal communication, and body language (Paulmann & Uskul, 2014). Most studies investigating aspects of emotion recognition use facial expressions as stimuli. Broadly speaking, these studies suggest that people identify basic emotions (e.g., fear, sadness, happiness) accurately, but experience more difficulty in recognizing more complex emotions such as disgust and guilt (see, e.g., Adolphs, 2010; Gobbini & Haxby, 2007; Laukka et al., 2013). One critique of these studies revolves around ecological validity: Emotions, especially as expressed via static facial expressions, are hardly ever observed in pure isolation in nature.

Relatively few studies examine acoustic aspects of emotion recognition (e.g., ways in which individuals use prosody to gather information about emotional cues; Paulmann & Uskul, 2014; Stewart et al., 2013). Typically, these studies have focused on verbal speech as a mode of communicating emotion. They have found that, in contrast to what is observed for facial emotion recognition, complex affective states are better identified than basic emotions through the vocal modality (App et al., 2011; Liebenthal et al., 2016). For instance, subtle cues in intonation or intensity of the voice appear to convey information about the speaker's attitudes and intentions (e.g., boredom, tenderness, sarcasm) that is difficult to detect from facial expressions (Morningstar, Nelson & Dirks, 2018).

Moreover, studies on acoustic aspects of emotion recognition have found that some short, discrete non-linguistic verbal expressions, termed *affective bursts,* are highly recognizable across cultures even in the absence of context (Laukka et al., 2013; Sauter et al., 2010). For instance, Koeda et al. (2013) investigated differences between Japanese and Canadian participants in perceiving affective bursts using the Montreal Affective Voices (MAVs) task. Their analyses detected no significant between-group differences in ability to accurately decode affective bursts conveying pain, surprise, and happiness. The implication of these results is that there are no cross-cultural differences in ability to recognise these emotions via the auditory modality.

However, other cross-cultural studies of emotion recognition, regardless of whether they use visual or auditory stimuli, tend to emphasize in-group advantages. That is to say, people seem better able to accurately recognize others' emotions when they share the same culture versus when they are from different cultural groups (Laukka et al., 2013; Paulmann & Uskul, 2014). For instance, Sauter et al. (2010) presented two groups of participants (native English speakers of European origin and people from the Himba culture, a remote settlement in Namibia) with a series of non-lingual vocalization sounds (embedded within an emotional story) that had been pre-recorded by members of both cultures. Participants were asked to identify the emotion conveyed by the different vocalization sounds. Results indicated that recognition of the emotional vocalizations was better when the stimulus producer and the listener were of the same culture.

This pattern of data may reflect variation in socialization across cultures. Different cultures place different value on the expression of particular emotions, depending largely on the mores and goals of that culture (Boiger, De Deyne, & Mesquita, 2013). Several studies comparing collectivist societies, such as Japan, and individualistic societies, such as the United States, have reported significant differences in perceived importance of certain emotions (An et al., 2017; Boiger et al., 2013; Gardner, Gabriel, & Lee, 1999; Kitayama, Mesquita, & Karasawa, 2006). Specifically, these studies find there is greater valuation of emotions expressing self-reflection and independence (e.g., pride, anger, guilt) in individualistic societies, whereas collectivist societies place greater value on emotions promoting harmony and interdependence (e.g., shame and modesty). Thus, these studies suggest that the cultural context in which emotions are experienced and perceived might affect the ways in which they are appraised.

Theory of Mind: Definition and cultural influences. Alongside emotion recognition, theory of mind (ToM) is the most intensely investigated domain of social cognition. The psychological construct of ToM describes the ability to attribute thoughts, feelings, and beliefs to other people, and to understand that each person is unique in these aspects (Premack & Woodruff, 1978). Intact ToM is an integral part of everyday social functioning because effective interpersonal communication requires an understanding of others' perspectives and the thoughts behind their responses (Bradford et al., 2018).

Most ToM studies investigate the development of this mental capacity in children (see, e.g., Baron-Cohen et al., 1985; Hoogenhout & Malcolm-Smith, 2017; Liu et al., 2008; Naito & Koyama, 2006; Wellman et al., 2001; Wimmer & Perner, 1983). Most of these studies use performance on false-belief tasks as the core index of ToM development. Findings suggest that there is, when measured using those tasks, relatively delayed development of ToM in children from outside of Western cultures compared to those raised within those cultures (Mayer & Träuble, 2013). However, these tasks do not measure all aspects of ToM, and they may contain cultural and linguistic biases that render them unsuited to assessing the construct in certain contexts (Hamilton, Hoogenhout, & Malcolm-Smith, 2016).

Relatively few studies examine cross-cultural differences in adult ToM. Bradford et al. (2018) reported that cultural influences on ToM ability were evident when comparing adults from individualistic and collectivist societies. Their study measured ToM in native English speakers and Chinese individuals using a belief-attribution task where participants had to attribute a belief to either themselves (self-oriented) or other people (other-oriented). Results indicated that Chinese participants performed better in other-oriented belief attribution, whereas native English-speaking participants performed better in self-oriented belief attribution. However, such cross-cultural differences are only present in a small number of similar studies, suggesting that ToM task performance of healthy adults does not vary significantly across cultures (Kobayashi, Glover, & Temple, 2006; Shulruf et al., 2011).

Rationale and Specific Aims

Social cognition involves a distinct network of mental processes underlying social interactions. Impairments in these domains hinder the ability to function effectively in social situations and to communicate interpersonally, thus affecting quality of life and overall wellbeing. Most studies investigating this psychological construct have focused on the discrete sub-domains of emotion recognition and theory of mind, and most have emerged from high-income countries in the global north (e.g., Canada, the United States, and England). Little research on emotion recognition and ToM has been conducted in low- and middle-income countries (LAMICs) such as South Africa, despite some preliminary suggestions that there are cross-cultural differences in patterns of emotion recognition and in ToM capacities. The research that has been done in South Africa has focused mostly on how these domains are affected in neurological disorders of childhood (e.g., autism spectrum disorder). Given the universal importance of intact social cognition for adaptive functioning in everyday environments, and the many potential benefits in describing how cultural influences might affect expression of social cognitive abilities, it is imperative that more local research be conducted. It is equally imperative that this research use culturally and linguistically appropriate measures of social cognition.

This pilot study aimed to evaluate, using participants from three different language groups, a South African adaptation of the NEmo social cognition battery. The NEmo battery is comprised of multimodal tasks of emotion recognition and ToM. Preliminary data suggest that performance on the battery is sensitive to socio-cognitive deficits in patients with mesial temporal lobe epilepsy (MTLE; Bauer et al., 2018). This study is the first step in a larger research project that aims to compare the NEmo performance of healthy Swiss and South African adults.

Methods

Design and Setting

This study adopted a cross-sectional design, comparing the performance of three groups of healthy South African adults (native English-speakers, native Xhosa-speakers, and native Afrikaans-speakers; the three dominant languages of the Western Cape province) on the NEmo facial emotion recognition (FER) and ToM subtests. These subtests were adapted for local use by employing stimuli (faces and voices) more suited to the South African context than the originals. The study also assessed whether performance on the NEmo FER and ToM tasks was dissociable from performance on standardized tasks assessing other domains of cognition.

Study procedures were conducted in a research laboratory located within the University of Cape Town (UCT) Department of Psychology. Ethical approval for study procedures was granted by the UCT Psychology Department’s Research Ethics Committee (reference: PSY2019-31).

Participants

Recruitment. I used convenience sampling (the UCT Department of Student Affairs Research Invitation email listserv) to recruit 40 healthy UCT students. Thirteen were native English-speakers, 19 were native Xhosa-speakers, and 8 were native Afrikaans-speakers. They were compensated R100 upon completion of the full study.

Eligibility criteria. Participants were required to be aged between 18–40 years and to be fluent in English. The age criterion was set in place because social cognitive abilities are not mature in children and young adolescents, and certain cognitive abilities tend to erode in older adults (Killgore et al., 2001; Nelson et al., 2005; Yurgelun-Todd & Killgore, 2006; Choudhury, Blakemore, & Charman, 2006). The language criterion was set in place because the cognitive tests, including the NEmo battery’s tasks, were administered in English. Moreover, participants were excluded if they had a history of serious head injury, neurological or psychiatric disorder, or substance abuse. These factors are all implicated in impaired social cognition, and so including participants with such histories might have introduced potential confounds to data interpretation (Cacioppo & Decety, 2011; Giovagnoli, Parente, Villani, Franceschetti, & Spreafico, 2013; Lee et al., 2013).

Materials

Screening measures. A study-specific online *sociodemographic and medical questionnaire* (https://www.psytoolkit.org/cgi-bin/psy2.5.4/survey?s=KtKvN) was used to assist in determining group assignment and eligibility to participate. Regarding the former, the questions about mother's first language, father's first language, and language in which the participant was raised determined the language group to which the participant would be assigned. Only those with consistent answers to all three questions (either English, Xhosa, or Afrikaans as a response) were eligible to participate. This questionnaire also helped ensure the groups were well matched by gathering self-reported information regarding age, sex, highest level of education, past and current physical health, and past and current substance use.

The *Beck Depression Inventory-II* (*BDI-II*; Beck, Steer, & Brown, 1996) is a 21-item self-report measure that was used to measure the severity of depression in participants (see Appendix A). The BDI-II has a high level of internal consistency (α = .91), good test-retest reliability (α = .93), and adequate factorial and content validity (Beck et al., 1996). The BDI-II has been used in many South African research studies and has demonstrated excellent psychometric properties in South African samples (see, e.g., Steele & Edwards, 2008; Ward, Flisher, Chrysostomos, Muller & Lombard, 2009; Kagee, Nel, & Saal, 2014). Each item presents the respondent with four statements related to some aspect of their mood over the past 2 weeks. The respondent then has to choose the statement that characterizes them best. Each item is scored on a scale ranging from 0 to 3, with lower scores representing absence or milder presence of depressive symptomatology. Individuals with total scores ≥ 29 were excluded from participation as scores in this range suggest the presence of severe depression. Individuals with severe depression have been shown to process emotions differently from healthy individuals and to struggle with inferring others' mental states (Weightman, Air, & Baune, 2014).

A *verbal fluency* test (Butters et al., 1987) measured ability to generate English words under time pressure. First, the potential participant was instructed to say as many words as possible within a 60-s time limit beginning with the letter B, then L, then S (letter fluency). Then, they were instructed to name as many animals as possible within 60-s, then as many fruits and vegetables as possible within 60-s, and finally as many verbs as possible within 60-s (semantic and action fluency). Only individuals who averaged > 10 words on both the letter and semantic/action fluency tasks were eligible to participate.

NEmo test battery. This battery¸ recently developed by researchers at the Swiss Epilepsy Centre, assesses social cognition in adults (Bauer et al., 2018). It is comprised of a set of computerised tasks assessing FER and ToM. Most tests within the battery were administered using Paradigm software (Perception Research Systems, 2007) on an ordinary laptop computer (Acer brand, 15.6-in HD display featuring an Intel CORE i5-7200U processor and using 4GB of RAM) and took roughly 90 minutes to complete. For each task, the number of correct answers as well as the average reaction time was recorded.

Five independent multimodal tasks provided measures of FER. The *Static FER* task required the test-taker to name the emotion depicted in a static photograph of a facial expression. Each of the photographs was presented individually on the computer screen and the respondent had to choose the emotion depicted from seven options (happy, sad, angry, afraid, disgusted, surprised and neutral). There were 3 practice runs and 56 recorded runs. The *Dynamic FER* task required the test-taker to name the emotion depicted in a video of a moving facial expression. Each video was presented individually on the computer screen for 500 milliseconds. There were 6 practice runs and 48 recorded runs. The *Simultaneous FER* task required the test-taker to view the emotions depicted in two videos of two facial expressions and state whether these were the same or different. The videos were presented simultaneously, in split screen, on the computer monitor for 500 milliseconds. There were 4 practice runs and 60 recorded runs. The *Prosodic FER* task required the test-taker to compare the emotion depicted in a photograph of a facial expression to that depicted by an emotionally loaded voice recording and state whether these were the same or different. These visual and auditory stimuli were presented simultaneously. There were 4 practice runs and 40 recorded runs. The *Emotional Stroop Test (EST)* assessed inhibition of interfering emotional signals by requiring the test-taker to name, as quickly and accurately as possible, the colours on three different sets of stimuli (first flowers, then the faces with a neutral expression, and then the faces expressing anxiety). This task was designed to be analogous to the classic Stroop Color-Word Test (i.e., the expectation is that time taken to name the colour will be slower in Part C, where emotion interferes with naming, than in Part A, where the task simply involves recognizing and naming a colour). Each part of the task comprised 40 stimuli. Time taken to name all the colours on each stimulus sheet were recorded by the test administrator.

Three independent tasks provided measures of ToM. The *Movie to Assess Social Cognition (MASC)* subtest required the test-taker to view a 15-minute film and, during short pauses, answer multiple-choice questions that involved inferring the mental states of characters in the current scene. A standard scoring sheet (see Appendix B) was used to record performance on this task. The *Irony* task required the test-taker to differentiate between irony, friendliness, and neutrality in spoken expressions. There were 3 practice runs and 18 recorded runs. The *Faux Pas* task required the test-taker to read a series of five short vignettes and to answer a few questions after each one. These questions were designed to assess whether the test-taker could (a) detect the social faux pas committed in the vignette, and (b) infer the mental states of characters in the scenes. Administration time was approximately 15 minutes. For this task, the test administrator recorded the test-taker's responses.

Standardized cognitive tests. These paper-and-pencil tasks were used to assess performance in domains purportedly independent of social cognition (viz., processing speed, set-shifting, response inhibition, and Performance IQ). They included the *Trail Making Test* (TMT; Reitan, 1992), the *Stroop Color-Word Test* (SCWT; Golden, 1978), and the *WAIS-IV Block Design*, *Picture Completion*, and *Digit Symbol-Coding* subtests (Wechsler, 2008).

Procedure

I sent individuals who had expressed interest in participating a link to the survey site (https://www.psytoolkit.org/cgi-bin/psy2.5.4/survey?s=KtKvN) that hosted the screening questionnaires and a randomized code that would grant them access to the site. When they landed on the site, they entered the code and completed the questionnaires anonymously. Individuals who met the eligibility criteria related to age, medical history, neurological history, and substance abuse history were then invited to participate in the laboratory-based portion of the study. After agreeing on a time and date for the test session, I gave the participant directions to the research laboratory.

All measures were administered by either me or Rebecca Johannesen, a Master's student in Psychology from the University of Zurich. They were administered individually in a quiet, private room. First, the participant read a consent form outlining details about the study (see Appendix C). They were also informed that they could withdraw from the study at any point and that they could ask questions should they need clarity about any part of the study. After signing the consent form, they were administered the verbal fluency task. If they remained eligible, they then completed the other cognitive tests in this order: Picture Completion, Faux Pas, SCWT, TMT, Prosodic FER, Simultaneous FER, Static FER, Block Design, Digit-Symbol Coding, EST, Dynamic FER, Irony, MASC. After the MASC subtest, the researcher conducted a structured interview asking about the participant's experience of that subtest (see Appendix D). Overall, test administration took roughly 2 hours. Upon completion of all tasks, the researcher debriefed the participant, thanked them, and awarded the R100 compensation.

Data Management and Statistical Analyses

I used R (version 3.6.1) and SPSS (version 25.0) to analyse the quantitative data. First, I used R to compute the aggregate scores on the NEmo FER and Irony subtests. Thereafter, I used SPSS to generate a full set of descriptive statistics to check whether the assumptions underlying the planned parametric tests (e.g., those of normality, homogeneity, and independence) were upheld. There were no significant outliers observed in the data, but the assumption of homogeneity was not upheld due to unequal group sizes. Therefore, I ran Games-Howell post-hoc pairwise tests. For all analyses, the value of α was set at .05 unless required otherwise (e.g., if there was a danger of inflated familywise error due to multiple comparisons). Effect sizes were interpreted follow convention: $\leq .05$ = small, .06 – .24 = medium, and ≥ 0.25 = large (Cohen, 1988).

Inferential statistical analyses proceeded as follows. First, a series of one-way ANOVAs evaluated between-group differences in performance on the NEmo subtests and on the standardized cognitive tests. Second, a series of bivariate correlational analyses (using Pearson's coefficient) assessed the dissociability of NEmo-measured FER and ToM from the cognitive constructs measured by the standardized cognitive tests. For these analyses, I set the alpha value at $p = .01$ as there was a danger of inflated familywise error due to multiple comparisons (scores on eight NEmo subtests were correlated with scores on six standardized cognitive tests).

Results

Sample Characteristics

Although their individual sample sizes were unequal, the groups were well matched demographically. Analyses detected no significant between-group differences in terms of age, sex distribution, and education (as indexed by the highest qualification obtained; see Table 1). Analyses also detected no significant between group differences in terms of previous schooling and employment status. So, these factors were unlikely to be confounds in subsequent data interpretation. Clinically the groups were also well matched, with analyses detecting no significant between-group differences on the BDI-II. On average, participants in all three groups self-reported depressive symptomatology in the range conventionally described as "minimal depression" (Beck et al., 1996). Analyses also detected no significant between-group differences in verbal fluency performance, indicating that fluency in English was unlikely to be a confound in subsequent data interpretation.

Table 1
Sample Characteristics (N = 40)

Variable	Group: English (n = 13)	Xhosa (n = 19)	Afrikaans (n = 8)	F / χ^2	p	ESE
Age (years)	21.00 (1.00)	21.89 (2.96)	23.75 (6.80)	1.41	.256	0.18
Sex				0.13	.881	< 0.001
Male	5 (38%)	8 (42%)	4 (50%)			
Female	8 (62%)	11 (58%)	4 (50%)			
BDI-II	5.31 (4.90)	6.32 (7.20)	7.25 (9.19)	0.20	.821	0.01
Verbal Fluency[a]	15.56 (2.92)	14.61 (3.03)	17.15 (2.52)	2.18	.128	0.11
HQO				0.13	.877	.01
Matric	6	13	5			
Undergraduate	7	4	2			
Postgraduate	0	2	1			
Employment				0.29	.792	.02
Student	12	18	8			
Part-time	1	1	0			
Primary School				2.95	.065	.14
Rural	0	6	1			
Urban	13	13	7			
High School				2.19	.126	.11
Rural	0	6	1			
Urban	13	13	7			

Note. For the variables *Age*, *BDI-II*, and *Verbal Fluency*, means are presented with standard deviations in parentheses. For the variables *Sex*, raw frequency is presented with percentage in parentheses. For the other variables, only raw frequencies are presented. ESE = effect size estimate (in this case, η_p^2 for ANOVA and Cramer's *V* for chi-square); HQO = highest qualification obtained; BDI-II= Beck Depression Inventory - Second Edition. [a]Data are the average performance across the three letter fluency tasks and the three semantic/action fluency tasks.

Between-Group Comparisons

NEmo emotion recognition subtests. Analyses detected significant between-group differences, both associated with medium effect sizes, on two of the five subtests: Prosodic FER and the inhibition trial of the EST (see Table 2). On both of those subtests, post-hoc testing suggested that (a) English-speakers performed significantly better than Xhosa-speakers, p = .023 (Prosodic FER) and .013 (EST-C), and (b) there were no other significant pairwise differences.

Table 2

Cognitive Performance: Descriptive statistics and between-group comparisons (N = 40)

Variable	Language Group					
	English (n = 13)	Xhosa (n = 19)	Afrikaans (n = 8)	F	p	η^2
NEmo Emotion Recognition						
EST-C	23.62 (4.17)	28.89 (5.60)	26.00 (4.60)	4.39	.019*	.192
Simultaneous FER	74.87 (7.80)	69.82 (8.44)	76.67 (6.55)	2.79	.077	.129
Dynamic FER	88.30 (6.34)	83.11 (10.51)	91.15 (6.67)	2.88	.069	.135
Static FER	84.34 (7.00)	82.81 (8.71)	84.38 (6.02)	0.20	.818	.011
Prosodic FER	77.89 (5.85)	71.18 (7.65)	76.88 (5.79)	4.39	.020*	.192
NEmo Theory of Mind						
Faux Pas	16.55 (2.30)	13.84 (3.22)	16.25 (2.82)	4.06	.025*	.180
MASC	35.08 (2.57)	30.58 (4.19)	34.75 (3.37)	7.45	.002**	.287
Irony	70.94 (9.12)	56.73(15.45)	85.42 (13.59)	13.83	< .001***	.428
WAIS-IV						
Picture Completion	22.15 (1.14)	19.84 (1.92)	22.38 (1.19)	11.56	< .001***	.384
Block Design	50.85 (11.76)	40.32 (10.85)	55.25 (5.12)	7.39	.002**	.285
Digit Symbol-Coding	87.85 (10.22)	78.16 (11.42)	83.38 (18.16)	2.30	.114	.111
TMT Part A (*s*)	26.92 (5.53)	26.00 (8.36)	25.13 (6.88)	0.16	.855	.008
TMT Part B (*s*)	68.00 (43.21)	62.21 (17.28)	55.13 (12.65)	0.53	.593	.028
SCWT	19.38 (3.02)	22.89 (5.85)	21.00 (3.74)	2.17	.129	.105

Note. In columns 2–4, means are presented with standard deviations in parentheses. EST-C = Emotional Stroop Test, Part C (inhibition trial); FER = facial emotion recognition; MASC = Movie to Assess Social Cognition; WAIS-IV = Wechsler Adult Intelligence Scale – Fourth Edition; TMT = Trail Making Test; SCWT = Stroop Color-Word Test (inhibition trial). Performance on all NEmo Emotion Recognition subtests is measured by percent correct responses, except EST-C (seconds to completion). Performance on all NEmo Theory of Mind subtests is measured by raw scores, except for Irony (percentage). Performance on all WAIS-IV subtests is measured by raw scores. Performance on the TMT and SCWT is measured by seconds to completion.

*p < .05. **p < .01. ***p < .001. Bonferroni-corrected p = 0.01.

NEmo Theory of Mind tests. Analyses detected significant between-group differences, associated with medium-to-large effect sizes, for all three subtests (see Table 2). Post-hoc testing suggested that (a) on the Faux Pas subtest, English-speakers performed significantly better than Xhosa-speakers, p = .026, (b) on the MASC subtest, both English-speakers and Afrikaans-speakers performed significantly better than Xhosa-speakers, p = .002 and .037, respectively, and (c) on the Irony subtest, English-speakers performed significantly better than Xhosa-speakers, p = .008, and (d) there were no other significant pairwise differences.

Standardized cognitive tests. Analyses detected significant between-group differences, both associated with large effect sizes, on two of the six subtests: WAIS-IV Picture Completion and Block Design (see Table 2). On both of those subtests, post-hoc testing suggested that (a) English-speakers performed significantly better than Xhosa-speakers, p = .001 (Picture Completion) and .043 (Block Design), (b) Afrikaans-speakers performed significantly better than Xhosa-speakers, p = .001 (Picture Completion) and < .001 (Block Design), and (c) there were no other significant pairwise differences.

Correlational Analyses

As might have been expected because these tasks all measure the same underlying construct, analyses detected significant positive correlations, of varying strength, between scores on the different NEmo facial emotion recognition tests (see Table 3). Performance on the Static FER test was associated with that on the Simultaneous FER (p = .001), Dynamic FER (p = .001), and Prosodic FER (p = .040) tests; performance on the Prosodic FER test was associated with that on the Simultaneous FER test (p = .012). Moreover, performance on the Simultaneous FER test was associated with that on the Dynamic FER (p < .001).

Similarly, analyses detected significant moderate positive correlations between (a) scores on the Faux Pas test and those on the MASC test, p = .007, (b) scores on the Faux Pas test and those the Irony test, p = .001, and (c) scores on the Irony test those on the and the MASC test, p = .005. Hence, better performance on one ToM test predicted better performance on the others, suggesting these tasks are measuring the same construct.

Of further interest is that many of the NEmo measures correlated significantly with one another across the boundaries of emotion recognition and ToM. For instance, analyses detected a negative moderate correlation between the Faux Pas test and the EST-C (p = .008). Moreover, there were moderate positive correlations between (a) the Faux Pas test and the Dynamic FER (p = .011), and Prosodic FER (p < .001) tests, (b) the MASC test and the Dynamic FER (p = .009) and Prosodic FER (p = .009) tests, and (c) the Irony test and the

EST-C (p = .006), Simultaneous FER (p = .006), and Dynamic FER (p < .001) tests. Hence, better performance on one measures of emotion recognition tended to predict better performance on tests of theory of mind, suggesting these tasks might have been measuring the same underlying construct (viz., social cognition).

Regarding associations between NEmo test performance and performance on the other cognitive tests, analyses detected a number of significant correlations (see Table 3). These included:

(a) a moderate negative correlation between the measure of processing speed (viz., time to complete the TMT-A) and Prosodic FER, p =.001;

(b) a moderate positive correlation between the measure of response inhibition (SCWT) and the NEmo task modelled on it, the EST-C, p = .001;

(c) moderate positive correlations between one of the components of WAIS-IV Performance IQ (Picture Completion) and the Faux Pas (p = .003), MASC (p =.005), Dynamic FER (p = .007) and Irony (p < .001) tests;

(d) moderate positive correlations between another of the components of WAIS-IV Performance IQ (Block Design) and the MASC (p = .001) and Irony (p < .007) tests; and

(e) a weak positive correlation between the third component of WAIS-IV Performance IQ (Digit-Symbol Coding) and Simultaneous FER, p =.008.

(f) a moderate negative correlation between the third component of WAIS-IV Performance IQ (Digit-Symbol Coding) and the EST-C, p =.004.

Together, these correlational data suggest that performance on the NEmo social cognition tests is not as dissociable from performance on the other cognitive tests as one might have predicted.

Table 3

Correlation Matrix: Associations between NEmo scores and other cognitive test scores (N = 40)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. NEmo EST-C	1.00													
2. NEmo Simultaneous FER	-.34**	1.00												
3. NEmo Dynamic FER	-.35**	.62**	1.00											
4. NEmo Static FER	.040	.49**	.52**	1.00										
5. NEmo Prosodic FER	-.44**	.39**	.27**	.33**	1.00									
6. NEmo Faux Pas	-.42**	.40**	.51**	.26	.54**	1.00								
7. NEmo MASC	-.12	.38**	.41**	.24	.41**	.42**	1.00							
8. NEmo Irony	-.43**	.42**	.63**	.11	.37**	.51**	.43**	1.00						
9. WAIS-IV PC	-.35**	.38**	.42**	.13	.22	.46**	.44**	.57**	1.00					
10. WAIS-IV BD	-.17	.27	.21	-.04	.21	.19	.51**	.42**	.58**	1.00				
11. WAIS-IV DS-C	-.45**	.42**	.26	.12	.24	.21	.35**	.25	.27	.37**	1.00			
12. TMT Part A	.36**	-.09	-.02	-.18	-.51**	-.15	.07	-.12	.17	.01	-.25	1.00		
13. TMT Part B	.28	-.20	-.18	-.00	-.17	.12	.10	-.23	-.08	-.07	-.21	.27	1.00	
14. SCWT	.49**	-.20	-.03	-.10	-.13	-.11	.01	-.17	-.32**	-.19	-.31**	.24	.26	1.00

Note. Data presented are Pearson's *r* correlation coefficients. Statistically significant correlations ($p < .01$) are marked in boldface font. EST-C = Emotional Stroop Test, Part C (inhibition trial); FER = facial emotion recognition; MASC = Movie to Assess Social Cognition; WAIS-IV = Wechsler Adult Intelligence Scale - Fourth Edition; PC = Picture Completion; BD = Block Design; DS-C = Digit Symbol-Coding; TMT = Trail Making Test; SCWT = Stroop Color-Word Test (inhibition trial).

**$p < .01$. ***$p < .001$.

Discussion

This pilot study's main aim was to evaluate a South African adaptation of the NEmo battery of social cognition, which measures of emotion recognition and theory of mind (ToM). Using a cross-sectional design, I compared the performance of three groups of healthy South African adults (native English-speakers, native Xhosa-speakers, and native Afrikaans-speakers) on South African-adapted NEmo facial emotion recognition (FER) and ToM subtests. A secondary aim of the study was to assess whether performance on the NEmo FER and ToM tasks was dissociable from performance on standardized tasks assessing other domains of cognition.

Below, I discuss the findings regarding the NEmo FER tasks. Then, I discuss the findings regarding the NEmo ToM tasks. Thereafter, I provide a brief overview of findings regarding dissociability of the NEmo tasks from those assessing from other cognitive domains (processing speed, set-shifting, response inhibition, and Performance IQ). Finally, I address the limitations of this study and provide directions for future research based on these findings.

Between-group Differences in NEmo Emotion Recognition

The NEmo Prosodic Facial Emotion Recognition (FER) task requires participants to state whether the emotion depicted in a photograph of a facial expression and an emotionally loaded voice recording were the same or different. The photograph and voice recording are presented simultaneously. Analyses indicated that participants in the English group performed significantly better on this task than those in the Xhosa group (i.e., they were more likely to correctly match the acoustic stimulus [tone of voice] to the visual stimulus [facial expression]).

One way to account for this between-group difference is to note that, although the auditory stimuli were adapted for South African use, the voices used were only those of people who were predominantly English-speakers. Hence, the current finding is consistent with research showing that people are more likely to accurately recognize the emotions expressed within the speech of others when they share the same culture as compared to when they come from different cultural groups (Laukka et al., 2013; Paulmann & Uskul, 2014). However, findings reported by Bryant and Barrett (2008) oppose this notion. In their study, performance on vocal emotion recognition was compared between Shuar hunter-horticulturalists from Amazonian Ecuador and native English speakers from the United States. Participants had to match emotional spoken utterances, produced by the native English speakers, to emotional expressions portrayed in pictured faces. Results showed there were no between-group differences (i.e., the Shuar participants matched the emotions as accurately as the American participants did). These results, in contrast to those from the current study, support the assertion that vocal emotion recognition is expressed similarly across cultures (Sauter et al., 2010).

Analyses also detected significant between-group differences on the NEmo Emotional Stroop Test (EST). There, participants in the English group performed significantly better (i.e., were quicker to complete the task) than those in the Xhosa group. These findings might be interpreted as suggesting that participants in the Xhosa group were more likely to be distracted by the anxious faces that served as stimuli. Given that all of the stimuli depicted the faces of white people, the current findings are consistent with the notion that there is a cross-race bias and in-group advantage in face processing (i.e., people are more likely to process facial information quickly when the face they are decoding belongs to a member of their own race; Laukka et al., 2013; Paulmann & Uskul, 2014; Young, Hugenberg, Bernstein, & Sacco, 2012). Moreover, Matsumoto, Olide, and Willingham (2009) showed that in-group advantages are particularly likely to be observed in the presentation of mimed facial expressions. Future research should ensure that the stimuli used in the NEmo EST task are faces that are representative of different races and ethnicities.

Regarding performance on other measures of emotion recognition (Static FER, Dynamic FER, and Simultaneous FER), analyses detected no significant between-group differences. Because these tasks only assessed emotion recognition via the visual modality, the current findings are consistent with the notion that visual emotion recognition processes are universal and relatively invariable across different cultural and linguistic contexts (Adolphs, 2010).

Between-group Differences in NEmo Theory of Mind

The NEmo Faux Pas task assesses the ability of the test-taker to detect social faux pas committed in a written vignette, and also to infer the mental states of characters in the scenes. Analyses indicated that participants in the English group performed significantly better on this task than those in the Xhosa group (i.e., they were more likely to detect the faux pas and to accurately infer the characters' mental states).

One way to account for this between-group difference is to note that the vignettes reflected Western cultural mores. Hence, the current finding is consistent with the notion that ToM develops and is expressed differently across cultures (Bradford et al., 2018). Any such faux pas task that seeks to be applied fairly across cultures must therefore proceed from the starting point that there are cultural differences as to what constitutes inappropriate behaviour (Molinsky, 2005), and that therefore all stimuli used must reflect, as much as possible, locally relevant social situations and mis-steps.

Analyses also detected significant between-group differences on the NEmo Irony task. There, participants in the English group performed significantly better (i.e., identified ironic statements within spoken expressions more accurately) than those in the Xhosa group. In order to recognise and understand irony, one needs to pick up on subtle cues, such as the speaker's tone of voice, in the context of the utterance (Mauchand, 2019). Although there are few published studies describing cross-cultural comparisons in the detection and comprehension of irony, Krygier-Bartz (2016) reported a result similar to ours. She found that Polish and Canadian adults interpreted the seriousness of irony differently. In this study, participants had to judge the speaker's intent and humour from either the addressee's perspective or the bystander's perspective. The findings revealed that there were differences in the manner in which Polish adults and Canadian adults rated public ironic criticisms: Polish adults rated them as less serious in comparison to private ironic criticisms, whilst Canadian adults rated these conditions similarly. Taken together with the current finding, there are inchoate suggestions that culture influences the detection and comprehension of irony during verbal encounters.

In terms of performance on the MASC test, analyses revealed significant between-group differences. Participants in both the English and Afrikaans groups performed significantly better on this task as compared to those in the Xhosa group. This task involved inferring mental states of the characters in a short film. Although it was adapted for South African use by translating it to English, these differences in performance reflect that the context or setting of this film was one of the factors that affected performance. This film was set in a Eurocentric context (Germany) that the participants in the Xhosa group were not likely accustomed to. Familiarity has been shown to be advantageous in the understanding of a specific social context (Laukka et al., 2013). In this case, the participants in the English and the Afrikaans groups were at an advantage as they were more familiar with Eurocentric social cues as compared to the Xhosa group. However, although between-group differences were observed for this task, all the language groups performed relatively adequate on this task. In this regard, future research should look into adapting this task further by ensuring it is set in a more local context.

For all the ToM tests, there were no significant between-group differences observed between the English group and the Afrikaans group. This suggests that the development and expression of ToM might be similar in these groups. This is consistent with the view that South African cultures (white English- and Afrikaans-speaking) that share several similarities with Western Europe tend to perform similarly on tasks adapted from the West (Claassen et al., 2001).

Dissociability of NEmo Tests Performance from that on Other Cognitive Tests

In terms of correlations between scores on the different NEmo facial emotion recognition tests, as might have been expected, analyses detected significant positive correlations. This is due to these tasks all measuring the same underlying construct, in varying strength, which in this case was emotion recognition. This pattern was also evident with regards to the different NEmo theory of mind tests, as analyses detected significant positive correlations between the Faux Pas, MASC and Irony subtests. Better performance on one ToM test predicted better performance on the others, suggesting these tasks are measuring the same construct. Furthermore, analyses detected that many of the NEmo measures correlated significantly with one another across the boundaries of emotion recognition and ToM. Hence, better performance on one of the measures of emotion recognition tended to predict better performance on tests of theory of mind, suggesting these tasks might have been measuring the same underlying construct (viz., social cognition).

Regarding associations between NEmo test performance and performance on the other cognitive tests, analyses detected a number of significant correlations. There was a moderate positive correlation between the SCWT and the EST. This was expected as the EST was developed analogous to the classic Stroop test. So, similarities in performance on these tasks reflects construct validity, as these tasks were measuring the same construct. Analyses also revealed that there was a moderate negative correlation between TMT Part A and Prosodic FER. High scores on the Prosodic FER test were associated with less times taken to complete the TMT task assessing motor speed. Furthermore, there were a few correlations between tasks assessing performance IQ and those that assessed emotion recognition. Analyses revealed that were moderate positive correlations between the Picture Completion task and the Dynamic FER test. This was also observed in the relationship between performance on the Digit-Symbol Coding task and the Simultaneous FER test.

Regarding the tests assessing ToM, analyses revealed that these were closely related to those assessing performance IQ. Performance on the Picture Completion task was moderately correlated with that on the Faux Pas test, MASC test and Irony test. So, high scores on this task were associated with better performance on the Faux Pas test, MASC test and the Irony test. Moreover, performance on the Block Design task was also moderately correlated with that on the MASC test and Irony test. So, high scores on this task were associated with better performance on the MASC test and the Irony test. These findings suggest that performance IQ is closely related to ToM.

Overall these correlational data suggest that performance on the NEmo social cognition tests is not as dissociable from performance on the other cognitive tests as one might have predicted.

Limitations and Directions for Future Research

Three primary limitations constrain the power of inferences one might make on the basis of the current findings. First, although this was a pilot study, the sample size was fairly small, and the groups were of unequal size. Hence, the study was statistically under-powered and there was an increased likelihood of Type I error (Rusticus & Lovato, 2014). A power analysis for a three-group design using a medium effect size ($f = 0.25$), $\alpha = .05$, and a minimum desired power of .80, indicated that a sample size of 159 would be sufficient (Faul, Erdfelder, Lang, & Buchner, 2007). Thus, future research should ensure that power is improved, and Type I error probability reduced by recruiting larger samples and equal numbers of people in each comparison group.

Second, the sample was non-representative. Participants were all university students between the ages of 18 and 40 years (indeed, only two of them were older than 25), all lived in suburban Cape Town, and all were fluent English speakers. Future research on the adaptation of the NEmo battery for South African use will have to recruit a more diverse sample, particularly with regard to age (the current study cannot answer whether the *development* of ToM differs across groups), language (the current study cannot answer how much differently Xhosa-speakers might have performed had they not been English-fluent), and socioeconomic status (the current study did not estimate SES, and so cannot answer the question of how that variable affects NEmo performance).

Third, all the instructions and tasks on the NEmo battery were delivered in English. South Africa has 11 official languages, and many of its citizens are not fluent in English. Hence, further South African adaptation of the NEmo must ensure not only that tasks are translated but that the voices and faces on the emotion recognition tasks are representative of local/regional context.

Summary and Conclusion

This study is the first to describe and evaluate a South African-adapted version of the NEmo test battery of social cognition. Given the importance of intact social cognition for adaptive functioning in everyday environments, and the presence of cultural influences on the expression of this construct, it is essential to use tools that are linguistically and culturally appropriate. Using a sample of participants from three distinct language groups, I found that English speakers tended to perform better than Xhosa speakers on most emotion recognition and theory of mind tasks. The former between-group difference might be accounted for by in-group advantages for participants in the English group (majority of the auditory stimuli were presented by the voices of native English speakers), and the latter by the fact that the ToM tasks were heavily biased toward Western cultural mores. Future research must ensure that NEmo tasks are adapted more adequately for Xhosa-speaking individuals to ensure that the construct of social cognition is measured reliably and with validity in that language group.

References

Adolphs, R. (2010). What does the amygdala contribute to social cognition? *Annals of the New York Academy of Sciences*. https://doi.org/10.1111/j.1749-6632.2010.05445.x

An, S., Ji, L. J., Marks, M., & Zhang, Z. (2017). Two sides of emotion: exploring positivity and negativity in six basic emotions across cultures. *Frontiers in psychology*, *8*, 610.

App, B., McIntosh, D. N., Reed, C. L., & Hertenstein, M. J. (2011). Nonverbal channel use in communication of emotion: How may depend on why. *Emotion, 11*(3), 603.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"?. *Cognition*, *21*(1), 37-46.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck Depression Inventory manual (2nd ed.)*. San Antonia, TX: Psychological Corporation.

Boiger, M., Mesquita, B., Uchida, Y., & Feldman Barrett, L. (2013). Condoned or condemned: The situational affordance of anger and shame in the United States and Japan. *Personality and Social Psychology Bulletin*, *39*(4), 540-553.

Bradford, E. E. F., Jentzsch, I., Gomez, J. C., Chen, Y., Zhang, D., & Su, Y. (2018). Cross-cultural differences in adult Theory of Mind abilities: A comparison of native-English speakers and native-Chinese speakers on the Self/Other Differentiation task. *Quarterly Journal of Experimental Psychology*. https://doi.org/10.1177/1747021818757170

Bryant, G., & Barrett, H. C. (2008). Vocal emotion recognition across disparate cultures. *Journal of Cognition and Culture*, *8*(1-2), 135-148.

Butters, N., Granholm, E., Salmon, D., Grant, I., & Wolfe, J. (1987). Episodic and semantic memory: A comparison of amnesic and demented patients. *Journal of Clinical and Experimental Neuropsychology, 9*, 479–497

Cacioppo, J. T., & Decety, J. (2011). Social neuroscience: challenges and opportunities in the study of complex behavior. *Annals of the New York Academy of Sciences*, *1224*(1), 162-173.

Choudhury, S., Blakemore, S. J., & Charman, T. (2006). Social cognitive development during adolescence. *Social cognitive and affective neuroscience*, *1*(3), 165-174.

Claassen, N. C. W., Krynauw, A. H., Paterson, H., and Wa ga Mathe, M. (2001). *A Standardisation of the WAIS-III for English-Speaking South Africans*. Pretoria: Human Sciences Research Council.

Cohen, J. (1988). Statistical power analysis for the social sciences.

Cohn, M., St-Laurent, M., Barnett, A., & McAndrews, M. P. (2015). Social inference deficits in temporal lobe epilepsy and lobectomy: Risk factors and neural substrates. *Social Cognitive and Affective Neuroscience*. https://doi.org/10.1093/scan/nsu101

Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). G* power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191.

Gardner, W. L., Gabriel, S., & Lee, A. Y. (1999). "I" value freedom, but "we" value relationships: Self-construal priming mirrors cultural differences in judgment. *Psychological Science*, *10*(4), 321-326.

Giovagnoli, A. R., Parente, A., Villani, F., Franceschetti, S., & Spreafico, R. (2013). Theory of mind and epilepsy: What clinical implications? *Epilepsia*. https://doi.org/10.1111/epi.12255

Gobbini, M. I., & Haxby, J. V. (2007). Neural systems for recognition of familiar faces. *Neuropsychologia*. https://doi.org/10.1016/j.neuropsychologia.2006.04.015

Golden, C. J. (1978). Stroop Colour and Word Test Stoelting. *Chicago, IL*.

Hamilton, K., Hoogenhout, M., & Malcolm-Smith, S. (2016). Neurocognitive considerations when assessing theory of mind in autism spectrum disorder. *Journal of Child & Adolescent Mental Health*, *28*(3), 233-241.

Henry, J. D., Von Hippel, W., Molenberghs, P., Lee, T., & Sachdev, P. S. (2016). Clinical assessment of social cognitive function in neurological disorders. *Nature Reviews Neurology*. https://doi.org/10.1038/nrneurol.2015.229

Hoogenhout, M., & Malcolm-Smith, S. (2017). Theory of mind predicts severity level in autism. *Autism, 21*(2), 242-252.

Jokeit, H. (2010). 12. Social cognition in patients with temporal lobe epilepsies. *Epilepsy & Behavior*. https://doi.org/10.1016/j.yebeh.2010.01.037

Kagee, A., Nel, A., & Saal, W. (2014). Factor structure of the Beck Depression Inventory-II among South Africans receiving antiretroviral therapy. *AIDS care*, *26*(2), 257-262.

Killgore, W. D., Oki, M., & Yurgelun-Todd, D. A. (2001). Sex-specific developmental changes in amygdala responses to affective faces. *Neuroreport*, *12*(2), 427-433.

Kitayama, S., Mesquita, B., & Karasawa, M. (2006). Cultural affordances and emotional experience: Socially engaging and disengaging emotions in Japan and the United States. *Journal of Personality and Social Psychology, 91*, 890-903.

Kobayashi, C., Glover, G. H., & Temple, E. (2006). Cultural and linguistic influence on neural bases of 'Theory of Mind': an fMRI study with Japanese bilinguals. Brain and language, 98(2), 210-220.

Koeda, M., Belin, P., Hama, T., Masuda, T., Matsuura, M., & Okubo, Y. (2013). Cross-cultural differences in the processing of non-verbal affective vocalizations by Japanese and canadian listeners. *Frontiers in psychology*, *4*, 105. doi:10.3389/fpsyg.2013.00105

Krygier-Bartz, M. (2016). Verbal irony comprehension in middle school age children and adults in Polish and English discourse.

Laukka, P., Elfenbein, H. A., Söder, N., Nordström, H., Althoff, J., Iraki, F. K. E., & Thingujam, N. S. (2013). Cross-cultural decoding of positive and negative non-linguistic emotion vocalizations. *Frontiers in Psychology*, *4*, 353.

Lee, J., Altshuler, L., Glahn, D. C., Miklowitz, D. J., Ochsner, K., & Green, M. F. (2013). Social and nonsocial cognition in bipolar disorder and schizophrenia: relative levels of impairment. *American Journal of Psychiatry*, *170*(3), 334-341.

Liebenthal, E., Silbersweig, D. A., & Stern, E. (2016). The language, tone and prosody of emotions: neural substrates and dynamics of spoken-word emotion perception. *Frontiers In Neuroscience*, *10*, 506.

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.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940–967. http://doi.org/1092-4388/07/5004-0940

Matsumoto, D., Olide, A., & Willingham, B. (2009). Is there an ingroup advantage in recognizing spontaneously expressed emotions?. *Journal of Nonverbal Behavior*, *33*(3), 181.

Mayer, A. F., & Träuble, B. (2013). Synchrony in the onset of mental state understanding across cultures? A study among children in Samoa. *International Journal of Behavioral Development*, *37*, 21–28.

Molinsky, A. (2005). Language Fluency and the Evaluation of Cultural Faux Pas: Russians Interviewing for Jobs in the United States. *Social Psychology Quarterly*, *68*(2), 103–120. https://doi.org/10.1177/019027250506800201

Morningstar, M., Nelson, E. E., & Dirks, M. A. (2018). Maturation of vocal emotion recognition: Insights from the developmental and neuroimaging literature. *Neuroscience*

& Biobehavioral Reviews, 90, 221-230.

Myer, L., Stein, D. J., Grimsrud, A., Seedat, S., & Williams, D. R. (2008). Social determinants of psychological distress in a nationally-representative sample of South African adults. *Social science & medicine*, *66*(8), 1828-1840.

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.

Nelson, E. E., Leibenluft, E., McClure, E. B., & Pine, D. S. (2005). The social re-orientation of adolescence: a neuroscience perspective on the process and its relation to psychopathology. *Psychological medicine, 35*(2), 163-174.

Nomura, H., Inoue, S., Kamimura, N., Shimodera, S., Mino, Y., Gregg, L., & Tarrier, N. (2005). A cross-cultural study on expressed emotion in carers of people with dementia and schizophrenia: Japan and England. *Social Psychiatry and Psychiatric Epidemiology*. https://doi.org/10.1007/s00127-005-0924-z

Paulmann, S., & Uskul, A. K. (2014). Cross-cultural emotional prosody recognition: Evidence from Chinese and British listeners. *Cognition and Emotion*. https://doi.org/10.1080/02699931.2013.812033

Perception Research Systems. 2007. Paradigm Stimulus Presentation, Retrieved from http://www.paradigmexperiments.com

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a 'theory of mind'? *Behavioral and Brain Sciences, 4*, 515-526.

Reitan, R. M. (1992). *Trail Making Test: Manual for administration and scoring*. Reitan Neuropsychology Laboratory.

Saunders, J. B., Aasland, O. G., Babor, T. F., De la Fuente, J. R., & Grant, M. (1993). Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption-II. *Addiction, 88*(6), 791-804.

Sauter, D. A., Eisner, F., Ekman, P., & Scott, S. K. (2010). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proceedings of the National Academy of Sciences, 107(*6), 2408-2412. https://doi.org/10.1073/pnas.0908239106

Siebert, J. M. (2018). *Toward Linguistically Fair IQ Screening: Multilingual Vocabulary Test.* (Unpublished Masters dissertation). University of Cape Town, South Africa.

Shulruf, B., Alesi, M., Ciochina, L., Faria, L., Hattie, J., Hong, F., & Watkins, D. (2011).

Measuring collectivism and individualism in the third millennium. *Social Behavior and Personality: An International Journal, 39,* 173–187.

Steele, G.A., Edwards, D.A. (2008). Translation of the BDI-II into Xhosa. *Journal of Psychology in Africa, 18*, 207-216.

Stewart, M. E., Mcadam, C., Ota, M., Peppé, S., & Cleland, J. (2013). Emotional recognition in autism spectrum conditions from voices and faces. *Autism*. https://doi.org/10.1177/1362361311424572

Ward, C. L., Flisher, A. J., Zissis, C., Muller, M., & Lombard, C. (2003). Reliability of the beck depression inventory and the self-rating anxiety scale in a sample of South African adolescents. *Journal of Child and Adolescent Mental Health, 15*(2), 73-75.Wechsler, D. (2008). Wechsler adult intelligence scale–Fourth Edition (WAIS–IV). *San Antonio, TX: NCS Pearson, 22*, 498.

Weightman, M. J., Air, T. M., & Baune, B. T. (2014). A review of the role of social cognition in major depressive disorder. *Frontiers in Psychiatry, 5*, 179.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child development,* 72(3), 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(1), 103-128.

Young, S. G., Hugenberg, K., Bernstein, M. J., & Sacco, D. F. (2012). Perception and motivation in face recognition: A critical review of theories of the cross-race effect. *Personality and Social Psychology Review*, *16*(2), 116-142.

Yurgelun-Todd, D. A., & Killgore, W. D. (2006). Fear-related activity in the prefrontal cortex increases with age during adolescence: a preliminary fMRI study. *Neuroscience letters*, *406*(3), 194-199.

Zimmermann, P., & Fimm, B. (2004). A test battery for attentional performance. In *Applied neuropsychology of attention* (pp. 124-165). Psychology Press.

Appendix A

Beck Depression Inventory-II (BDI-II)

Name: ___________________________________

Age: ____________ Sex: ______

Instructions: This questionnaire consists of 21 groups of statements. Please read each group of statements carefully, and then pick out the one statement in each group that best describes the way you have been feeling during the past two weeks, including today. Circle the number beside the statement that you have picked. If several statements in the group seem to apply equally well, circle the highest number for that group. Be sure that you do not choose more than one statement for any group, including Item 16 (Changes in Sleep Pattern) and Item 18 (Changes in Appetite).

1. Sadness 0 I do not feel sad. 1 I feel sad much of the time. 2 I am sad all of the time. 3 I am so sad or unhappy that I can't stand it. **2. Pessimism** 0 I am not discouraged about my future. 1 I feel more discouraged about my future than I used to be. 2 I do not expect things to work out for me. 3 I feel my future is hopeless and will only get worse. **3. Past Failure** 0 I do not feel like a failure 1 I have failed more than I should have. 2 As I look back, I see a lot of failures.	**6. Punishment Feelings** 0 I don't feel I am being punished. 1 I feel I may be punished. 2 I expect to be punished. 3 I feel I am being punished. **7. Self-Dislike** 0 I feel the same about myself as ever. 1 I have lost confidence in myself. 2 I am disappointed in myself. 3 I dislike myself. **8. Self-Criticalness** 0 I don't criticise or blame myself more than usual. 1 I am more critical of myself than I used to be. 2 I criticise myself for all my faults. 3 I blame myself for everything bad that happens.

3 I feel I am a total failure as a person. **4. Loss of Pleasure** 0 I get as much pleasure as I ever did from the things I enjoy. 1 I don't enjoy things as much as I used to. 2 I get very little pleasure from the things I used to enjoy. 3 I can't get any pleasure from the things I used to enjoy. **5. Guilty Feelings** 0 I don't feel particularly guilty. 1 I feel guilty over many things I have done or should have done 2 I feel quite most of the time. 3 I feel guilty all of the time.	**9. Suicidal Thoughts or Wishes** 0 I don't have any thoughts of killing myself. 1 I have thoughts of killing myself, but I would not carry them out. 2 I would like to kill myself. 3 I would kill myself if I had the chance **10. Crying** 0 I don't cry any more than I used to. 1 I cry more than I used to. 2 I cry over every little thing. 3 I feel like crying, but I can't.

11. Agitation 0 I am no more restless or wound up than usual. 1 I feel more restless or wound up than usual. 2 I am so restless or agitated that it’s hard to stay still. 3 I am so restless or agitated that I have to keep moving or doing something. **12. Loss of Interest** 0 I have not lost interest in other people or activities. 1 I am less interested in other people or things than before. 2 I have lost most of my interest in other people or things. 3 It’s hard to get interested in anything. **13. Indecisiveness** 0 I make decisions as well as ever. 1 I find it more difficult to make decisions than usual. 2 I have much greater difficulty in making decisions than I used to. 3 I have trouble making any decisions. **14. Worthlessness** 0 I do not feel I am worthless. 1 I don’t consider myself as worthwhile and useful as I used to be. 2 I feel more worthless as compared to other people. 3 I feel utterly worthless. **15. Loss of Energy**	**17. Irritability** 0 I am no more irritable than usual. 1 I am more irritable than usual. 2 I am much more irritable than usual. 3 I am irritable all the time. **18. Changes in Appetite** 0 I have not experienced any changes in my appetite 1a My appetite is somewhat less than usual. 1b My appetite is somewhat more than usual. 2a My appetite is much less than usual. 2b My appetite is much more than usual. 3a I have no appetite at all. 3b I crave food all the time. **19. Concentration Difficulty** 0 I can concentrate as well as ever. 1 I can’t concentrate as well as usual. 2 It’s hard to keep my mind on anything for very long. 3 I find I can’t concentrate on anything. **20. Tiredness or Fatigue** 0 I am no more tired or fatigued than usual. 1 I get more tired or fatigued more easily than usual. 2 I am too tired or fatigued to do a lot of the things I used to do. 3 I am too tired or fatigued to do most things I used to do. **21. Loss of Interest in Sex**

0 I have as much energy as ever. 1 I have less energy than I used to have. 2 I don't have enough energy to do very much. 3 I don't have enough energy to do anything. **16. Changes in Sleep Pattern** 0 I have not experienced any change in my sleeping pattern. 1a I sleep somewhat more than usual. 1b I sleep somewhat less than usual. 2a I sleep a lot more than usual. 2b I sleep a lot less than usual. 3a I sleep most of the day. 3b I wake up 1-2 hours early and can't get back to sleep.	0 I have not noticed any recent change in my interest in sex. 1 I am less interested in sex than I used to be. 2 I am much less interested in sex now. 3 I have lost interest in sex completely.

Appendix B

NEmo MASC subtest scoring sheet

	a	b	c	d	
1	0	0	0	0	missing
2	0	0	0	0	missing
3	0	0	0	0	missing
4	0	0	0	0	missing
5	0	0	0	0	missing
6	0	0	0	0	missing
7	0	0	0	0	missing
8	0	0	0	0	missing
9	0	0	0	0	missing
10	0	0	0	0	missing
11	0	0	0	0	missing
12	0	0	0	0	missing
13	0	0	0	0	missing
14	0	0	0	0	missing
15	0	0	0	0	missing
16	0	0	0	0	missing
17	0	0	0	0	missing
18	0	0	0	0	missing
19	0	0	0	0	missing
20	0	0	0	0	missing
21	0	0	0	0	missing
22	0	0	0	0	missing
23	0	0	0	0	missing
24	0	0	0	0	missing
25	0	0	0	0	missing
26	0	0	0	0	missing
27	0	0	0	0	missing
28	0	0	0	0	missing
29	0	0	0	0	missing
30	0	0	0	0	missing
31	0	0	0	0	missing
32	0	0	0	0	missing
33	0	0	0	0	missing
34	0	0	0	0	missing
35	0	0	0	0	missing
36	0	0	0	0	missing
37	0	0	0	0	missing

38 0 0 0 0 missing
39 0 0 0 0 missing
40 0 0 0 0 missing
41 0 0 0 0 missing
42 0 0 0 0 missing
43 0 0 0 0 missing
44 0 0 0 0 missing
45 0 0 0 0 missing

control a b c d
13c 0 0 0 0 mistake
33c 0 0 0 0 mistake
38c 0 0 0 0 mistake
46c 0 0 0 0 mistake
47c 0 0 0 0 mistake
48c 0 0 0 0 mistake

Total "correct"	0		Total "control"	0
Total "exceeding"	0		Total "mistake"	6
Total "less"	0			
Total "no"	0			
Total "missing"	45			

Appendix C

Informed Consent Form

University of Cape Town

Adapting the NEmo Test Battery for Use in South Africa

Research Study

This form provides you with information about this study and seeks your informed consent to participate. Before you agree to take part in this study, please read the information below and ask the researcher (Nwabisa Mlandu) questions about anything that you do not understand. The data collected in this study is to be used towards the completion of an Honours degree in the Department of Psychology at the University of Cape Town.

Purpose

The main aim of this study is to describe and evaluate a South African adaptation of the NEmo tests of emotion recognition and theory of mind. It is the first step in a larger research project that aims to compare the NEmo performance of healthy Swiss and South African adults.

Procedure

If you agree to participate in the second phase of this study, you will complete a set of questionnaires describing yourself, the languages you are fluent in, how you feel and how much alcohol you use. If you are eligible to continue, you will be administered the NEmo tests battery and a set of standardized cognitive tests. This session should last no more than 2 hours.

Participation, Withdrawal, Confidentiality and Risks

Participation in this study is completely voluntary and you may withdraw from the study at any time without any consequences. All information will be treated with confidentiality: at no point will your name or personal details be disclosed to anyone outside the research team.

There are no foreseeable risks involved in this study and there are no costs involved on your behalf. You will be compensated for your time with R100.

Principal Investigators, Ethics Committee, and Contact Details

Nwabisa Mlandu
Department of Psychology
University of Cape Town
nwabisamlandu@rocketmail.com

Kevin G. F. Thomas, Ph.D
Department of Psychology
University of Cape Town
kevin.thomas@uct.ac.za

Research Ethics Committee:

Vicki Ives-Deliperi, Ph.D
Neuroscience Institute
University of Cape Town
vickideliperi@icloud.com

Rosalind Adams
Department of Psychology
University of Cape Town
rosalind.adams@uct.ac.za

I, ____, give my informed consent to participate in this study.

Signature: ______________________Date:_______________________

Researcher's Signature: _______________________________________

Appendix D

Structured interview to assess MASC task

1. What do you think the movie was about?
2. What do you think the task was assessing?
3. Did you find any of the movie scenes particularly interesting?
4. If so, what scene/s were these and why were they interesting?
5. Did you experience any discomfort while watching the movie?
6. If so, at what point of the movie was this discomfort experienced?
7. Did you experience any difficulties answering the MASC task questions?
8. Do you think this task is suitable for South African use? Why/Why not?