

Confirmatory Factor Analysis of the Kaufman Assessment Battery in a sample of Primary School-aged Children in Rural South Africa

Short title: Confirmatory Factor Analysis of KABC-II in Rural South Africa

Joanie M Mitchell^{1,2}, Mark Tomlinson², Ruth M Bland^{1,3,4} Brian Houle^{5,6,7}, Alan Stein^{5,8},
Tamsen J Rochat^{1,8,9,10}

Institutions and affiliations:

1. Africa Health Research Institute, Durban, South Africa
2. Department of Psychology, Stellenbosch University, Stellenbosch, South Africa
3. School of Public Health, Faculty of Health Sciences, University of Witwatersrand, Johannesburg, South Africa
4. Royal Hospital for Sick Children and Institute of Health and Wellbeing, University of Glasgow, Glasgow, United Kingdom
5. MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa
6. School of Demography, The Australian National University, Canberra, Australia
7. CU Population Center, Institute of Behavioral Science, University of Colorado at Boulder, Boulder, Colorado, United States
8. Department of Psychiatry, University of Oxford, United Kingdom
9. Human and Social Development Programme, Human Sciences Research Council, South Africa
10. MRC/Developmental Pathways to Health Research Unit, School of Clinical Medicine, University of Witwatersrand, South Africa

Correspondence:

Tamsen Rochat

trochat@hsrc.ac.za (Tel: 0312425544)

Human and Social Development Programme, Human Sciences Research Council

5th floor, The Atrium, 430 Peter Mokaba Ridge, Berea, Durban, 4001

Abstract

The Kaufman Assessment Battery for Children, Second Edition (KABC-II), measures cognitive processing, includes non-verbal sub-tests, and is increasingly used in low- and middle-income countries. While the KABC-II has been validated in the United States, a psychometric evaluation has not been conducted in Southern Africa. This study aims to establish the reliability and validity of the KABC-II among a sample of 376 primary school-aged children in rural South Africa (7- 11 years). We examined Cronbach alpha and conducted a confirmatory factor analysis. The battery showed good reliability (Mental Processing Index [$\alpha=.78$]); and the originally validated structure of the KABC-II was maintained ($\chi^2= 16.30, p=.432$). Mean scores were low on the Planning sub-scale. On the Simultaneous sub-scale, the mean score was higher for the supplementary sub-test Block Counting versus the core sub-test Triangles. With translation and the inclusion of supplementary subtests, the KABC-II is an appropriate assessment to use in this context.

(150/150)

Keywords: Confirmatory factor analysis, KABC-II, primary school-aged children, South Africa

The number of children at risk of poor development has reduced worldwide in recent years, however Sub-Saharan Africa (SSA) has seen little improvement, and still has the highest number of children failing to fulfil their potential (Lu, Black, & Richter, 2016). Recent evidence has begun to illustrate the cognitive disadvantages SSA children face from primary school age, with children performing well below expected scores for age (Rochat et al., 2016), and often experiencing education failure (measured by repeated grades) early in their schooling careers (Mitchell et al., 2016). Promoting children's cognitive abilities is a key factor in encouraging development. In order to accomplish this, there has been a scale up of early-life interventions (Britto et al., 2016). To understand the impact of these interventions, reliable measures are needed. There is a lack of appropriate and culture-fair cognitive assessments available to use with African children.

In South Africa, the Senior South African Individual Scale (SSAIS-R) have been locally adapted to include children aged 7-16, but the test developers provide only limited normative data for English and Afrikaans speaking children and acknowledge biases based on socio-economic status (Van Eeden, 1997). Furthermore while the SSAIS-R does measure some non-verbal skills, it relies predominantly on the measurement of verbal intelligence (similar to the Wechsler Intelligence Scale for Children (WISC) on which it is based), making it somewhat biased to educational exposure, in particular verbal and linguistic abilities.

The biases introduced by socio-economic status can influence children's performance on cognitive tests through a number of pathways including poor parental education, poor nutrition and health care, low expectations, exposure to violence, and poor quality or absence of school exposure (C. R. Reynolds & Ramsay, 2003). Attempts to address these issues in child cognitive assessments have seen a shift towards assessments measuring underlying

cognitive processes necessary to solve tasks, rather than learned knowledge (Laher & Cockcroft, 2017). These cognitive processing assessments contain testing rules which allow for some practice and learning prior to scored testing. One locally adapted example of this approach is the McCarthy Scales of Children's Abilities, which has been validated in South Africa (Richter, Griesel, & Rose, 1994). However the McCarthy Scales have an upper limit of 8 years, limiting its usability with older samples of primary school aged children.

When considering cognitive assessments which measure underlying processing and have usability with a larger group of children (3-18 years), the Kaufman Assessment Battery, Second Edition (KABC-II), is an increasingly popular option. It has been established to show less bias to school exposure in low-income samples, and to be appropriate with cultural and linguistic minorities (Kaufman & Kaufman, 2004). The KABC-II can be administered from two theoretical approaches, the Cattell-Horn-Carroll Theory (CHC) and Luria's Neuropsychological theory. The choice of approach informs the test format used, and in turn, the expected factor structure. The Luria approach was appropriate to use in this study as it is recommended for low-income children, thus we tested a factor structure reflecting four domains of cognition (Learning, Sequential, Simultaneous and Planning), each of which are made up of their own subtests.

The KABC-II has gone through rigorous psychometric evaluation in the United States where the assessment was developed, it was shown to have construct validity and excellent reliability (mean $\alpha=.90$) with a large normative sample ($N=3025$) (Kaufman & Kaufman, 2004). These findings were largely replicated in subsequent studies in the United States (Morgan, Rothlisberg, McIntosh, & Hunt, 2009; M. R. Reynolds, Keith, Fine, Fisher, & Low, 2007). Consequently, it has been increasingly used in low- and middle-income countries (LMIC), including in Asia and Africa (Bangirana et al., 2009; M. Malda, van de Vijver, Srinivasan, Transler, & Sukumar, 2010; Ogunlade et al., 2011).

Despite being increasingly used in LMIC, validation studies of the KABC-II have been scarce; however two studies have largely replicated findings from the normative sample, showing the instrument to be valid and reliable in non-Western settings. Firstly, in India, the KABC-II underwent adaptation (M Malda et al., 2008), and was found to be reliable and to have maintained its factor structure with a sample of 6-10 year old Kannada-speaking children (N=598) (M. Malda et al., 2010). Secondly, factor analysis demonstrated that the KABC-II maintained its factor structure in a cohort of 7-16 year old children with cerebral malaria in Uganda (N=65) (Bangirana et al., 2009).

While the KABC-II has been used in Malawi (Boivin et al., 2011) and South Africa (Baumgartner et al., 2012; Ogunlade et al., 2011), a psychometric evaluation has not been conducted in Southern Africa. The primary aim of this research was to establish the reliability and test the validated structure of the KABC-II among a sample of primary school-aged children in rural South Africa. Drawing on previous research (Bangirana et al., 2009; M. Malda et al., 2010), we hypothesize that despite the potential for the KABC-II to be influenced by school exposure and cultural variation, the original validated structure of the KABC-II would be maintained in this sample, with each of the subscales reflecting specific subtests and domains of cognition.

Method

Participants

This analysis was conducted retrospectively, with this psychometric evaluation sample being drawn from a large cohort (the ‘Siyakhula’ cohort), which recruited children between 2012 and 2014 (Rochat et al., 2016). This research operated from the Africa Health Research Institute (AHRI), previously known as the Africa Centre for Population Health. AHRI is located in the Umkhanyakude district in northern KwaZulu-Natal, South Africa. This area has one of the highest prevalence of HIV worldwide (Zaidi, Grapsa, Tanser, Newell, & Bärnighausen, 2013), and is predominately isiZulu-Speaking, and rural, however it contains a peri-urban township. AHRI conducts a demographic surveillance in a defined geographical area of 438 square kilometres (the demographic surveillance area [DSA]), routinely collecting data from approximately 85 000 members of 11 000 households (Tanser et al., 2008).

The Siyakhula cohort consisted of HIV-uninfected children born to HIV-infected and uninfected mothers, and included two groups of children: the first was the intervention group, who had been exposed to an early life breastfeeding intervention (Bland, Coovadia, Coutsooudis, Rollins, & Newell, 2010). The second was the comparison group; these children were born in the same years in the DSA and had not been part of the early life intervention.

Children included in this psychometric evaluation sample were drawn from the comparison group of the Siyakhula cohort. Furthermore, these children were HIV unexposed in fetal life (mothers were HIV negative during pregnancy and at birth), and HIV unaffected in childhood (mothers currently HIV uninfected). In conducting this psychometric evaluation,

we felt these criteria aided in selecting a sample which best reflected a normative population sample of children, because:

- i. HIV exposure is known to affect child development and may particularly disadvantage some children;
- ii. Early life interventions are known to influence child development and may particularly advantage some children.

Of the 657 children in the Siyakhula comparison group (see Figure 1), we excluded 151 children exposed to HIV in fetal life and a further 114 whose mothers sero-converted after delivery. 16 women exited the study, leaving a final analytic sample of 376 children.

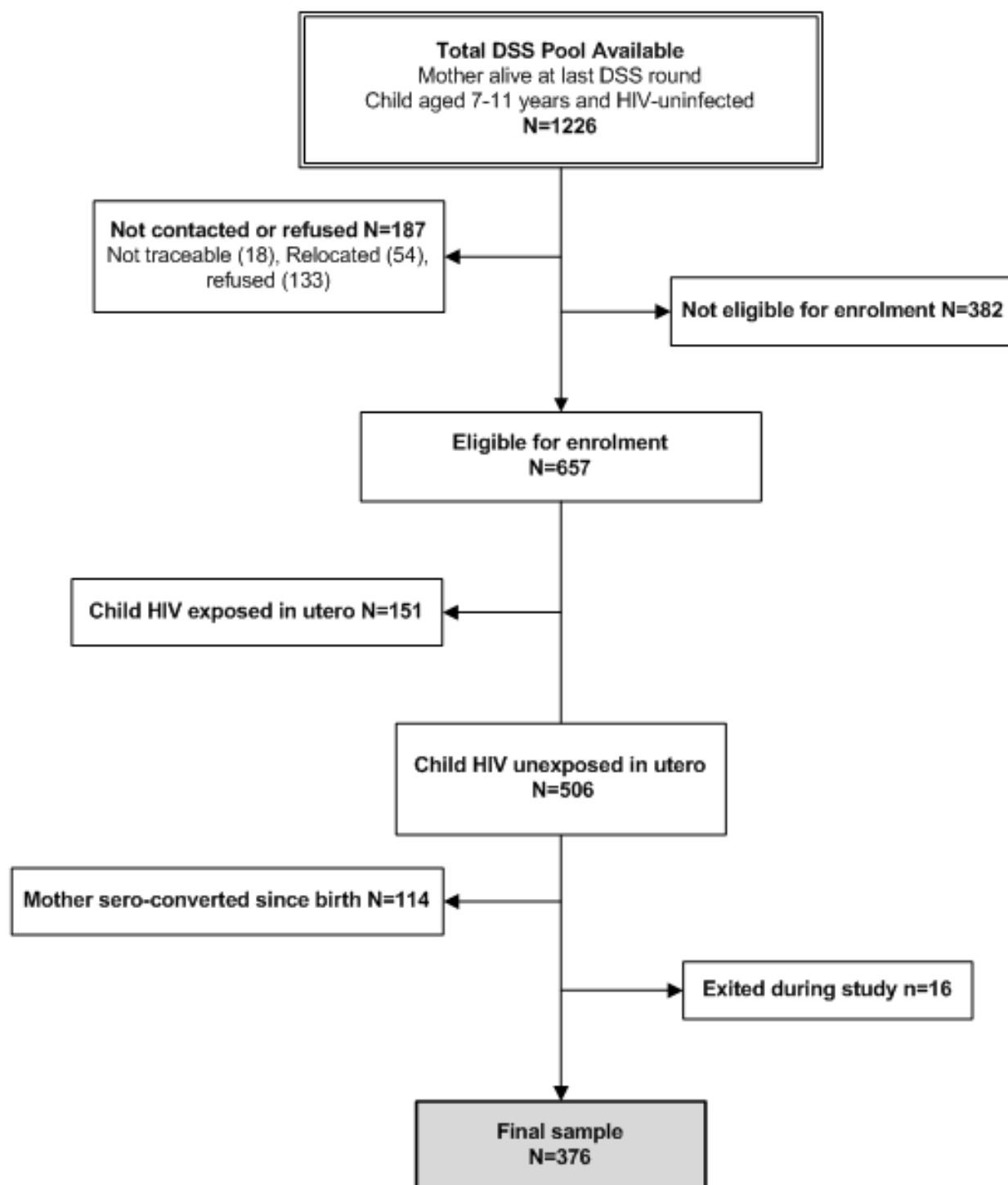


Figure 1. Consort diagram of sample (N=376), Umkhanyakude district of KwaZulu-Natal, South Africa, 2012-2014

Instruments

The KABC-II was the primary outcome measure used in the Siyakhula Cohort study. The KABC-II is an assessment of cognitive processing which is used with children aged 3-18 years (different subtests are used for children aged 3-6 years and 7-18 years), and takes approximately 45 minutes (Kaufman & Kaufman, 2004). As part of a scoping exercise done by Grand Challenges Canada (the funder of this study), the KABC-II was selected as the most widely appropriate battery for a range of different cultural contexts.

A particular strength of the battery is that it can be administered from two theoretical approaches: Luria's Neuropsychological theory reflecting a cognitive processing approach and the Cattell-Horn-Carroll (CHC) approach which reflects an acquired knowledge approach (including a knowledge scale). This allows the battery to be adapted appropriately to each specific sample with which it is used, for example, the CHC approach may be more appropriate for children in a Western context with exposure to good quality schooling. For this study, the Luria approach was chosen as it is less dependent on exposure to formal education and more appropriate for children in this context as they are from a different cultural background to the normative sample.

Using the Luria approach with this sample of children, aged 7 to 11 years, there are 13 sub-tests available to administer, measuring four sub-scales (Learning, Sequential, Simultaneous and Planning sub-scales). Of the 13 sub-tests, eight are core sub-tests and are used to calculate the Mental Processing Index (MPI), a global measure of the KABC-II, and five are supplementary sub-tests designed to provide a more extensive assessment of the sub-scales and allow children multiple opportunities to successfully complete tasks. An expert review team,

including the authors of the KABC-II, selected the sub-tests (both core and supplementary) considered to be most culturally appropriate.

Figure 2 shows an outline of the original battery used for this age group and the battery of sub-tests selected for this study. The expert review team recommended the exclusion of three of the sub-tests from the original battery, namely; Rebus, Rebus Delayed and Gestalt Closure. These would have required extensive adaptation due to the inclusion of stimuli which could show bias to culture, language and school exposure, and this was not feasible in the time frame. As eight core sub-tests are necessary for measuring the MPI, and the core sub-test Rebus was excluded, we included Atlantis Delayed as a core working memory sub-test, which is an acceptable substitution according to the administration guidelines, this substitution was made with advice from the test developers (Kaufman & Kaufman, 2004). Therefore the final battery consisted of 10 sub-tests- eight core and two supplementary (see Table 1 for the descriptions of the sub-tests and their respective sub-scales).

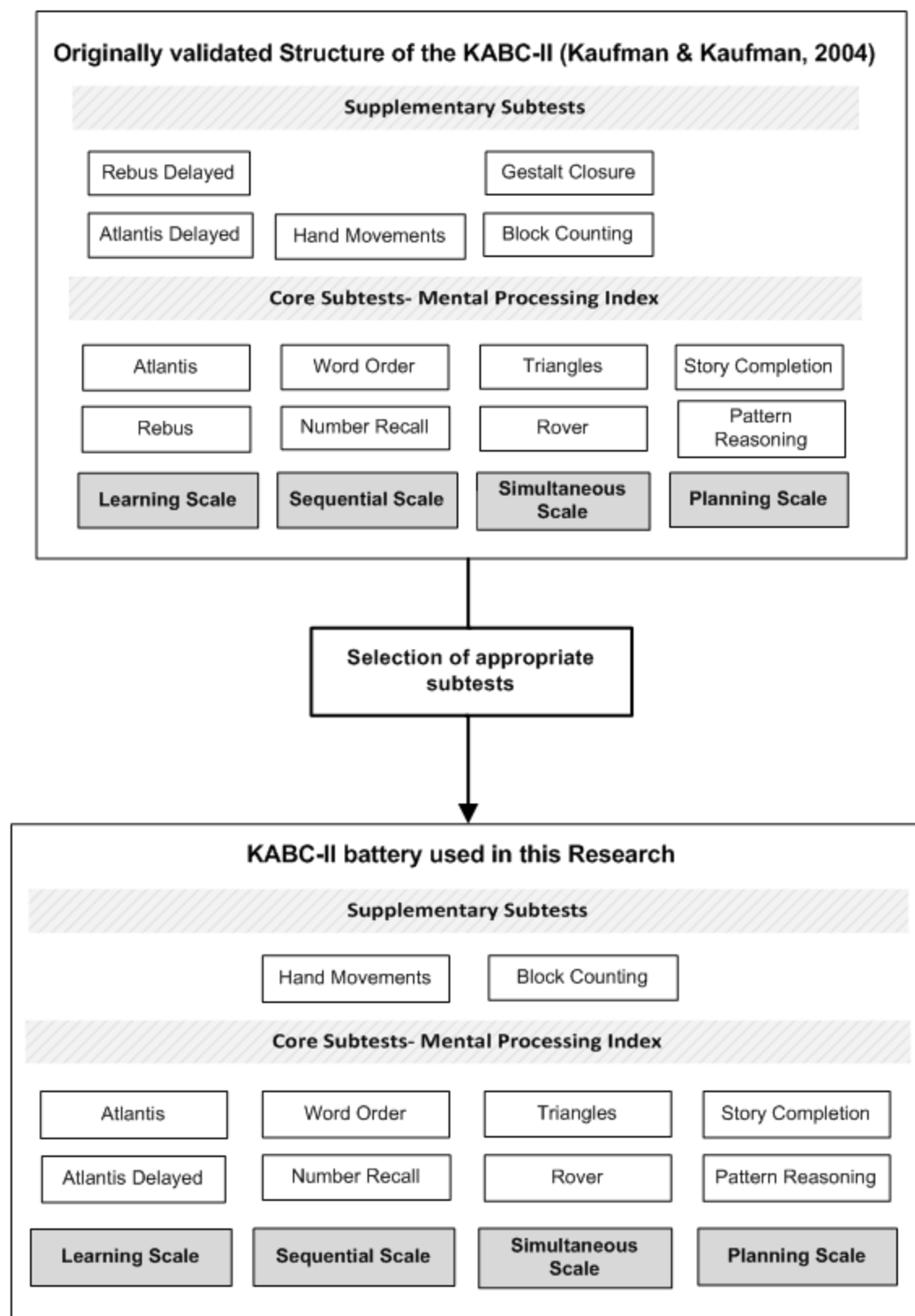


Figure 2. Selected KABC-II battery, Umkhanyakude district of KwaZulu-Natal, South Africa

Table 1

Description of Selected KABC-II Battery, Umkhanyakude district of KwaZulu-Natal, South Africa

Learning Scale	These problems require focused, sustained and selective attention and the ability to code and store information and integrate it with visual stimuli
Atlantis	Tests associative memory and storing newly learned information (The child is taught nonsense names for pictures and demonstrates learning by pointing to each picture, out of an array of pictures, when it is named).
Atlantis Delayed	Tests long-term storage and retrieval, storing and efficiently retrieving previously learned information (The child demonstrates delayed recall of paired associations learned about 15-25 minutes earlier during Atlantis).
Sequential Scale	Completing these activities relies on successive processing, in which the emphasis is on the order of the stimuli
Number recall	Tests memory span, particularly auditory memory, taking in and holding information, and then using it within a few seconds (The child touches a series of silhouettes of common objects in the same order as the examiner said the names of the objects; more difficult items include an interference task (colour naming) between the stimulus and response). The test stimuli were left unaltered as in this context children are taught to count in English at home and in schools.
Word order	Tests memory span and working memory, both visual and auditory, taking in and holding information, and using it within a few seconds (The child touches a series of silhouettes of common objects in the same order as the names them). Test stimuli were translated into isiZulu and during the interference task the child can choose to call out the colours in English or isiZulu.
Hand movements	Tests short-term visual memory span and storage, taking in and holding information, and then using it within a few seconds (The child copies the examiner's precise sequence of taps on the table with the fist, palm, or side of the hand)
Simultaneous Scale	These problems require a process of spatial integration of stimuli, with maximum efficiency. The input must be processed simultaneously, such that separate stimuli are integrated and conceptualised as a whole
Rover	Non-verbal problem solving task using spatial scanning, general sequential reasoning, visualisation, (The child moves a toy dog to a bone on a checkerboard - like grid that contains obstacles (rocks and weeds) and tries to find the shortest path).
Triangles	Spatial relations and visualization, including visual processing, perceiving, storage and

manipulating patterns (The child assembles several identical foam triangles (blue on one side, yellow on the other) to match a picture of an abstract design).

Block Counting Visual orientation in relation to spatial relationships, measures problem solving with visual patterns (The child counts the exact number of blocks in pictures of stacks of blocks. One or more of the blocks are hidden or partially hidden from view).

Planning Scale **These activities involve high-level decision-making and executive processes, the ability to generate hypotheses, to revise one's plan of action and to monitor and evaluate the best solution to the problem**

Pattern Reasoning Tests fluid reasoning, ability to solving novel problems by using reasoning abilities such as induction and deduction (The child is shown a series of stimuli that form a logical, linear pattern and the child completes the pattern by selecting the correct stimulus that is missing from an array of 4 to 6 options at the bottom of the page).

Story Completion Tests verbal mediation, working memory, planning ability and fluid reasoning (The child is shown a row of pictures that tell a story, but some of the pictures are missing, the child must select the pictures needed to complete the story from a set of pictures).
One change was made to this subtest: an item on the measure depicts the story of the preparation and cooking of an egg. The child is given several pictures and is required to place three of these into the three missing places to complete this story. The first picture should be the egg being cracked into a pan, the second the egg in the pan with the egg white still pale and thirdly the egg being cooked through. However one of the picture options is of the pan being washed and the assessors reported that many children were putting this card first; this was further confirmed by the results of the pilot study. In this rural Zulu context it is common for utensils to be washed before cooking food, therefore it was decided that the children would also get maximum points if they placed this picture first.

KABC-II: Kaufman Assessment Battery for Children, Second Edition.

Note. Text adapted from "Kaufman Assessment Battery for Children: Second Edition (p.43-44; 61-69)," by A. Kaufman, and N.L. Kaufman, 2004, Minnesota: AGS Publishing

The test battery is licensed to Pearson Ltd. USA, and test kits and forms were purchased. All subtests in the battery were retained, without significant adaptation, while the administration manual was translated, under license from Pearson Ltd, with fees waived. Initial translation of the instructions from English to isiZulu was undertaken by a panel of local isiZulu-speaking university graduates, who are proficient in English. Once the initial translation was completed, two sets of blind back translations were undertaken by two isiZulu-speaking PhD students. Test stimuli were largely left unaltered, with few exceptions explained in Table 1. Following translation, the battery was sent back to Pearson's who approved the changes.

Each of the included sub-tests contain discontinue and timing rules, with Triangles, Pattern Reasoning and Story Completion using time points (extra points given for rapid responses). The KABC-II does contain no-time-point versions if it is thought that children would be disadvantaged by timing rules. However in the normative sample in the United States (N=3025), it was found that there was little variation in scores when time points were excluded, and we could not find other studies which had excluded time points. Furthermore, the investigators felt that administering the battery in the standardised approach would be beneficial as it had not been evaluated in this context, therefore time points were included.

Procedures

Six non-professional isiZulu-speaking child assessors administered the KABC-II. They had all completed high school and had 5-10 years' experience working in the research field, including administering assessments to children. Before data collection commenced, a one week long training, led by the project investigators (psychologist, psychiatrist, paediatrician), was undertaken. The training was split into several sessions covering: rapport with the child, the content of each sub-test and its specific administration rules (including timing, scoring and

discontinue rules), and calculating and recording overall scores. Prior to starting data collection, each assessor completed a competency test measured against a gold standard assessor, a masters-level psychology student. During data collection, each assessor underwent on average four quality assurance visits by the gold standard assessor to establish equivalence across assessors. In these sessions, the assessments were observed and a set of criteria were evaluated, and responses tape-recorded. These criteria included rapport with the child, adhering to the scoring, timing and discontinue rules.

We adopted a decentralised approach to conducting assessments, whereby children were assessed either at a government primary health care clinic or at their school. This model was appropriate because children were geographically dispersed and bringing them to a single centralised location would have resulted in long travelling times for some children. In each testing setting, the assessor had access to the same equipment, including mobile tables, chairs and cooling fans. The child received refreshments during the course of the assessment session; however care was taken to choose snacks which had few additives and sugar, and a low glycaemic index. These steps were taken to ensure equivalence in the assessments which were approximately 45 to 60 minutes long.

Ethical Considerations

Ethical clearance for the Siyakhula study was received from the Biomedical Research Ethics Committee (BREC) of the University of KwaZulu-Natal (reference number: BF184/12). The activities outlined in this research are covered by this ethical clearance. All mothers and other primary caregivers (if applicable) gave informed consent, including written documentation, and assent was obtained from children prior to data collection.

Data Analysis

Following data collection, raw scores were transformed into scaled scores using the test norms published in the test manual, which allows for comparison across age groups (Kaufman & Kaufman, 2004). All subsequent analyses were conducted using Stata version 13 (StataCorp, 2007).

Psychometric properties and performance

Each sub-test has a specific discontinue rule, for example an incorrect answer for three consecutive items. These rules create missing data, therefore as advised by the test authors, we used the split-half technique adjusted with Spearman-Brown formula to establish reliability for the sub-tests (Kaufman & Kaufman, 2004), and Cronbach alpha for the MPI. We examined the normality of the sub-tests using the Shapiro-Wilk statistic and plotting the distributions of the MPI, the sub-scales and the subtests.

Confirmatory factor analysis

As the MPI is the global score and commonly only the core sub-tests are included, we examined the hierarchical nature of the MPI using three theoretical confirmatory factor analysis (CFA) models, with the eight relevant sub-tests as observed variables. Model one represented a single stratum model, with the sub-tests loading directly onto one overall variable; model two was a two-stratum model with the sub-tests loading onto the sub-scales as the latent factors and correlating together in the absence of an overall factor. The third model represented the originally validated structure of the MPI (Kaufman & Kaufman, 2004), with the sub-tests loading onto the sub-scales which in turn loaded onto an overall factor. Given that multivariate normality was violated, we estimated the overall fit of each model using the asymptotic-distribution free (ADF) method of estimation (Thompson, 2004). We used listwise deletion as

there was minimal missing KABC-II data (<1%). To determine the sufficiency of the sample size to estimate a fit for the models, we used a ratio of 20 participants to one parameter (Jackson, Gillaspy, & Purc-Stephenson, 2009).

We determined goodness of fit using the following fit statistics: model chi-square (χ^2), Root *Mean* Square Error of Approximation (RMSEA) along with the lower and upper bound of its 90% confidence interval, Comparative Fit Index (CFI) and the Standardized Root *Mean* Square Residual (SRMR). A good fit was indicated by: non-significant *p* value of the χ^2 (Hu & Bentler, 1999); a RMSEA value of less than 0.07 (Kline, 2011), with an upper bound 95% CI of <0.08 and a lower bound of close to 0.00; a CFI of .95 and higher; and a SRMR value of less than .08 (Hooper, Coughlan, & Mullen, 2008).

Supplementary model

To further investigate variation in the sub-scales and the battery structure, we examined performance on the supplementary sub-test Block Counting (Simultaneous sub-scale), and tested a fourth model which was similar to the structure of model 3 but included Block Counting in place of Triangles.

Results

Sample Characteristics

The majority of the sample were girls (65%) and the age range was 7-11 years, with a mean of 8.97 years (Table 2). The vast majority (93%) of children were in the primary care of their mothers and were living with their mothers. A quarter of the mothers had completed high school (schooling system in South Africa is described as a footnote in Table 2). All children were enrolled in school at the time of data collection, and most children were in Grade 3. The children had an average birth weight of 3.12kg and the majority had never had TB or been on chronic medication.

Table 2

Sample Characteristics of Children (N=376), Umkhanyakude district of KwaZulu-Natal, South Africa, 2012-2014

	Frequency (%)	<i>M (SD)</i>	Range
Age (<i>n</i> =376)		8.97 (0.83)	7-11
Sex (<i>n</i> =376)			
Female	244 (65%)		
Male	132 (35%)		
Primary Caregiver (<i>n</i> =373)			
Mother	351 (93%)		
Other	22 (6%)		
Caregiver Education Level (<i>n</i> =367)			
None	29 (8%)		
Some Primary	75 (20%)		
Primary Completed	57 (15%)		
Grade 10	104 (28%)		
Matriculation	90 (24%)		
Post matriculation	9 (2%)		
Child Attend School (<i>n</i> =376)			
Yes	376 (100%)		
No	0 (0%)		
Current Grade ¹ (<i>n</i> =376)		3.35 (1.07)	1-7
Birth Weight (kg) (<i>n</i> =320)		3.12 (0.56)	1.5-4.5
Current Weight (kg) (<i>n</i> =376)		29.01 (7.20)	17.2-90.3
Had TB (<i>n</i> =374)			
Yes	16 (4%)		
No	358 (95%)		
Any Chronic Medication (<i>n</i> =372)			
Yes	6 (2%)		
No	366 (97%)		

SD: standard deviation

¹In South Africa school starts at age 5 in Grade R (reception class) and finishes with Grade 12 and includes several phases: Grade R to Grade 3 (foundation phase), Grade 4 to 6 (intermediate phase), Grade 7 to 9 (senior phase), Grade 10 to 12 (further education and training phase)

Psychometric Properties and Performance

The MPI showed good reliability ($\alpha=0.78$), and the sub-scales and sub-tests showed excellent split-half reliability, ranging from 0.81 to 0.96 (Table 3). The majority of sub-tests and sub-scales showed skewed distributions, with the exception of the Sequential ($S-W=0.99$, $p=.243$), Number Recall ($S-W=0.99$, $p=.430$), and Atlantis Delayed sub-scales ($S-W=0.99$, $p=.810$) (see Table 3). The mean score on the MPI was 76.00 ($SD=10.91$), and of the four sub-scales, the mean scores on the Sequential sub-scale were the highest ($M=31.47$, $SD=13.70$), with the Simultaneous ($M=77.02$, $SD=13.73$) and Planning sub-scales ($M=72.46$, $SD=11.74$) having the lowest means (Table 3). Figure 3 demonstrates the distributions of the MPI, the subscales and the subtests, illustrating children's performance. The mean scores were stratified by age group and sex, however few significant differences were found.

Table 3

Performance and Psychometric Properties of the KABC-II Core Battery (N=376), Umkhanyakude district of KwaZulu-Natal, South Africa, 2012-2014

	M¹	SD	Range	Reliability²	S-W (p)
MPI	76.00	10.91	45-115	.78	0.99 ($p=.001$)
Learning	84.93	11.57	52-120	.96	0.99 ($p=.008$)
Atlantis	6.99	2.36	2-14	.96	0.99 ($p=.013$)
Atlantis Delayed	7.52	2.54	1-15	.81	0.99 ($p=.810$)
Sequential	91.47	13.70	51-144	.88	0.99 ($p=.243$)
Number Recall	9.93	3.03	1-19	.82	0.99 ($p=.430$)
Word Order	7.10	2.32	2-17	.83	0.97 ($p<.001$)
Simultaneous	77.02	13.73	43-116	.95	0.99 ($p=.001$)
Rover	7.14	3.03	1-16	.93	0.98 ($p<.001$)
Triangles	5.38	2.49	1-12	.92	0.97 ($p<.001$)
Planning	72.46	11.74	51-117	.94	0.97 ($p<.001$)
Story Completion	4.36	2.44	1-14	.88	0.96 ($p<.001$)
Pattern Reasoning	5.78	2.82	1-19	.94	0.95 ($p<.001$)

KABC-II: Kaufman Assessment Battery for Children, Second Edition; MPI: mental processing index; SD: standard deviation, S-W: Shapiro-Wilk statistic.

¹Standard scores (MPI; sub-scales) possible range of 43-160; scaled scores (sub-tests) possible range of 1-19

² Reliability: Cronbach alpha (MPI); split half reliability (sub-tests and sub-scales)

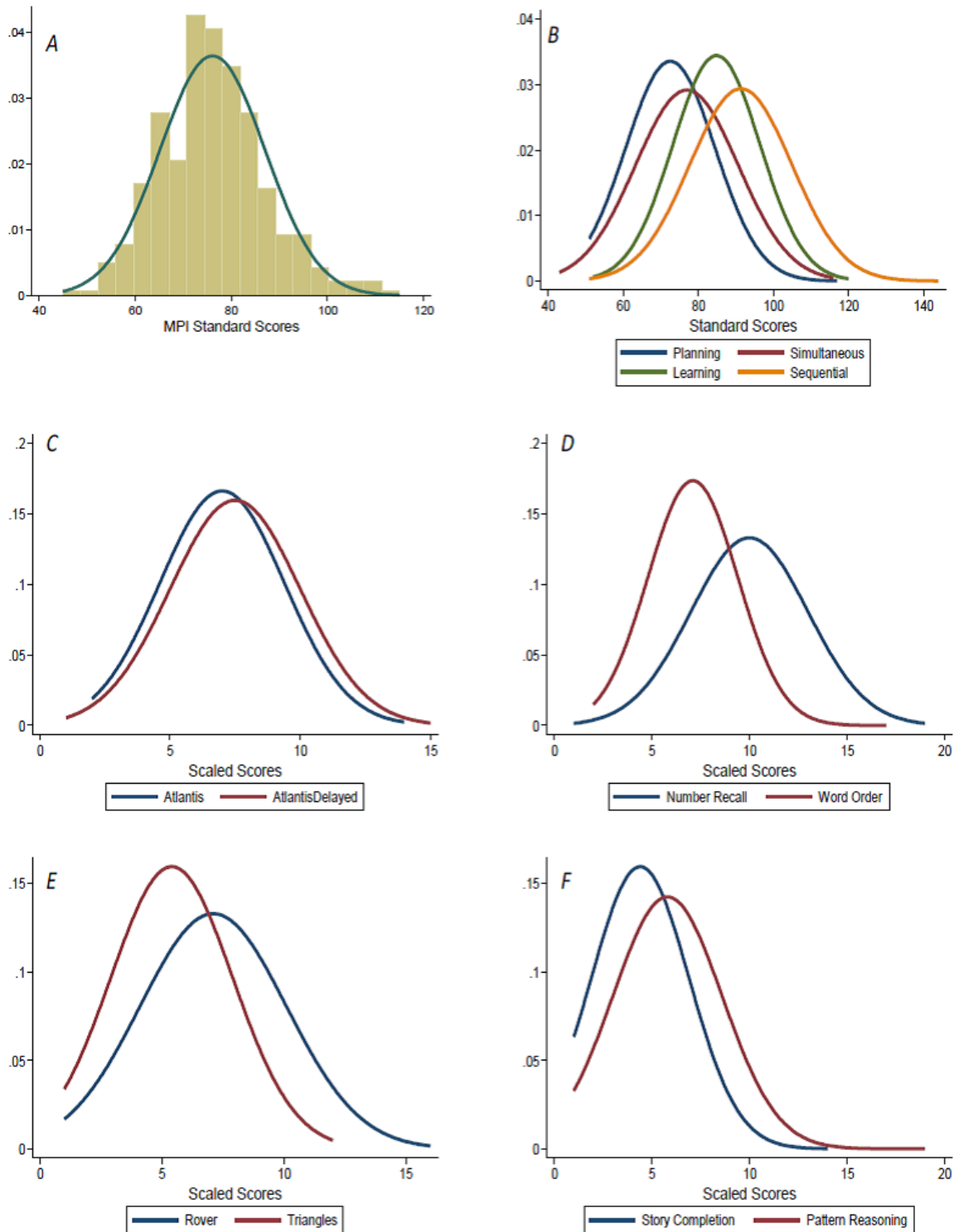


Figure 3. A. Distribution of the MPI scores; B. Distributions of the scale index scores; C. Distribution of scores on the Learning Scale subtests; D. Distribution of scores on the Sequential Scale subtests; E. Distribution of scores on the Simultaneous Scale subtests; F. Distribution of Planning Scale subtests (N=376), Umkhanyakude district of KwaZulu-Natal, South Africa, 2012-2014

Confirmatory Factor Analysis

The resulting standardized regression models for Models 1, 2 and 3 are shown in Figure 4, along with their corresponding fit statistics. Model one (Figure 4A) showed a poor fit to the data ($\chi^2 = 75.29$, $p < .001$) while model two ($\chi^2 = 14.65$, $p = .4022$) and three both showed good fits to the data ($\chi^2 = 16.30$, $p = .432$), with the RMSEA, CFI and SRMR falling within the acceptable fit thresholds. Among the equation level statistics of model two and three (Figures 4B and 4C), the coefficients of the Simultaneous and Planning sub-scales were high (0.99 and 0.91 respectively), indicating a strong association between the MPI and these two sub-scales.

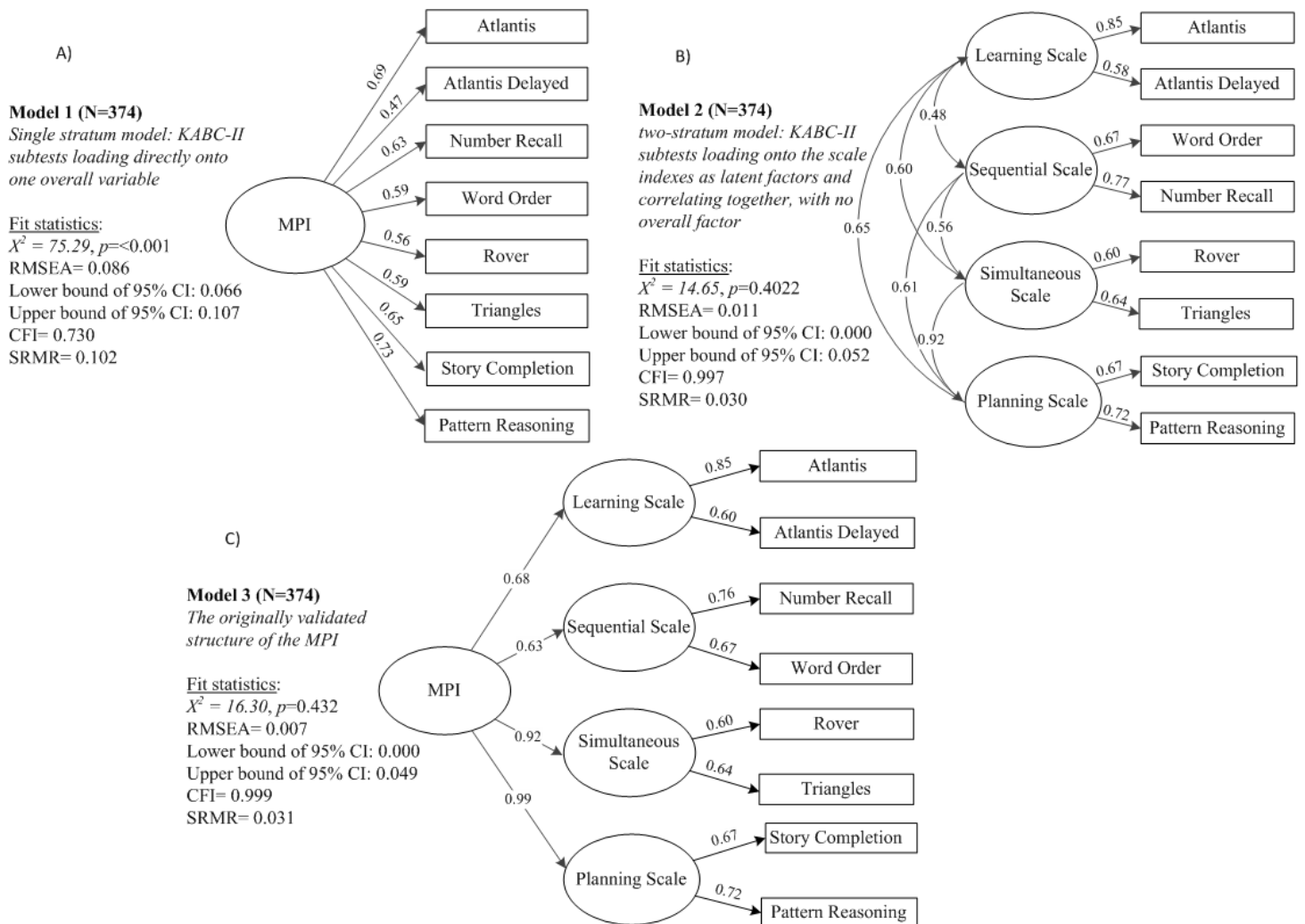


Figure 4. Three theoretical CFA models A) Model 1: single-stratum model, B) Model 2: two-stratum model and C) Model 3: hierarchical

Supplementary Model

The supplementary sub-test Block Counting (Simultaneous sub-scale) was examined to further investigate the strong association and clustering observed in model 3, and because children obtained higher mean scores on Block Counting as compared to the other Simultaneous sub-scale subtests. In comparison to the other Simultaneous sub-tests, children obtained the highest mean on Block Counting ($M=7.93$, $SD=2.7$, range: 1-19) and the lowest mean on Triangles ($M=5.38$, $SD=2.5$).

When used in the place of Triangles, the MPI ($M=78.19$, $SD=11.0$) and the Simultaneous sub-scale score increased ($M=84.79$, $SD=14.30$, range: 43-131). To evaluate whether the increase in the mean of the Simultaneous sub-scale resulted in a change in the model fit, the CFA was re-run, including Block Counting and excluding Triangles. Similar results were yielded, with the CFA model still providing an acceptable fit (see Figure 5).

Model 4 (N=373)

Supplementary model: structure of the MPI, including Block Counting instead of Triangles

Fit statistics:

$\chi^2 = 22.42, p=0.1302$

RMSEA= 0.033

Lower bound of 95% CI: 0.00

Upper bound of 95% CI: 0.062

CFI= 0.966

SRMR= 0.051

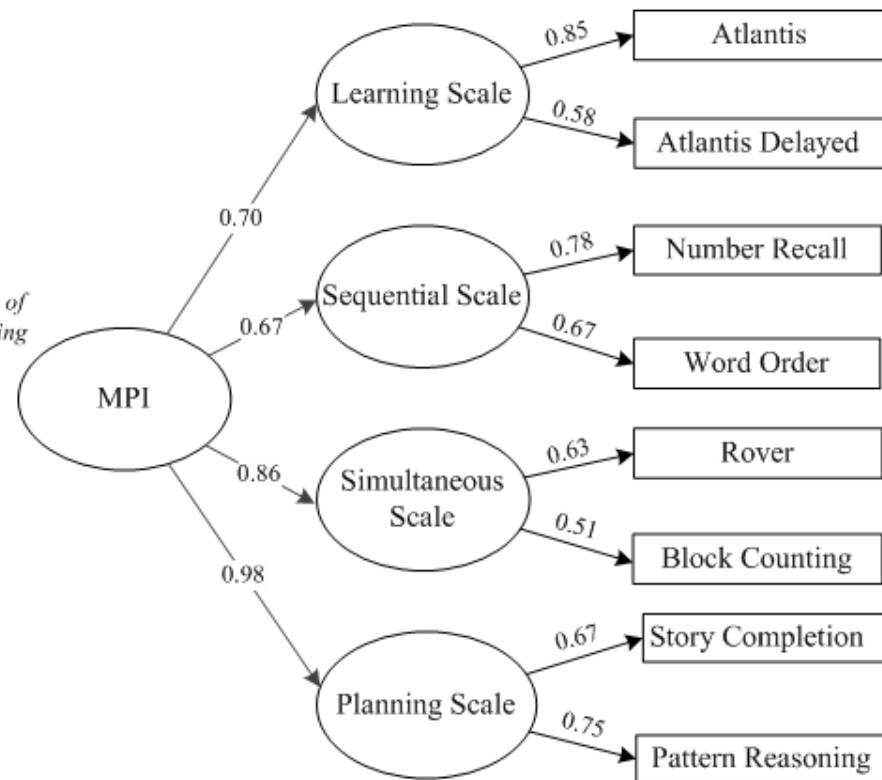


Figure 5. Theoretical hierarchical CFA model including Block Counting replacing Triangles

Discussion

Reliability

Our results showed the KABC-II to have good reliability in a sample of South African children, aged 7 – 11 years, living in a predominately rural area, who were HIV unexposed at birth and HIV unaffected during childhood, and who have only received standardised public health early life care. This is similar to findings in other research using the battery in LMIC settings (M. Malda et al., 2010; Moon, Mclean, & Kaufman, 2003), showing that the battery has good overall internal consistency.

Factor Structure

The originally validated KABC-II structure was maintained in this sample. Factor structure was likewise maintained in samples in Uganda (Bangirana et al., 2009) and India (M. Malda et al., 2010). Our analysis of the underlying structure demonstrated a strong association and clustering between the Simultaneous and Planning sub-scales. An examination of performance revealed majority of children had lower means on the sub-tests of these two sub-scales. The finding of low means and little variation in scores for the Planning and Simultaneous sub-scales were also observed in the Indian cohort, with similar clustering reported (M. Malda et al., 2010). The factor structure of the originally validated KABC-II being maintained is important as it reveals that the battery had construct validity when used with this sample of children in rural South Africa.

Performance trends

In this sample we examined how children performed on the subtests of the four sub-scales (Learning, Sequential, Simultaneous and Planning sub-scales). It was found that overall

children in this sample had the lowest scores on the Planning subtests, and on the Triangles subtest of the Simultaneous sub-scale. This trend in performance has been observed in literature from other LMIC contexts, with samples from Cambodia, Ethiopia, Kenya, India, Tanzania, Democratic Republic of Congo and South Africa (Boivin et al., 2013; O'Donnell et al., 2012; Taljaard et al., 2013; Veena et al., 2010). While these trends have been observed, there is a dearth of literature which has examined potential factors contributing to weaker performance on certain KABC-II subtests. It is not possible to definitively answer this question with the data available in this study; however we were able to explore potential hypotheses, drawing on the results we have and on the broader literature. These are discussed below as test-related and skill-related factors.

Test-related factors in this context

This hypothesis posits that there were test-related factors (including administration rules and the test content) which impacted on children in this context, affecting their performance on the Planning Scale and on the Triangles subtest (of the Simultaneous Scale).

Administration rules

The Story Completion and Pattern Reasoning subtests (Planning sub-scale) and the Triangles subtest (Simultaneous sub-scale) all include time points (extra points given for rapid responses) which are part of the administration rules of the standardised test. There is evidence in the literature which supports the idea that time limits may not be culturally sensitive in non-Western settings (Rosselli & Ardila, 2003). This is based on the suggestion that different cultures tend to put a different emphasis on the importance of completing tasks quickly versus completing a task correctly, irrespective of speed (Oakland, 2009). The authors who adapted the KABC-II for use in India noted that they included a more liberal time limit for some tasks, as they felt that

their local schools did not train children to manage their time and perform quickly (M Malda et al., 2008).

There is literature illustrating this concept in African contexts, suggesting that the constructs of ‘intelligence’ and ‘responsibility’ are linked in these settings, with the responsibility aspects being highly valued in children (Benson, 2003). Thus, this rural South African sample may have felt it more culturally appropriate to take time to consider and carefully complete tasks, in a ‘responsible’ and considered way.

Unfortunately, a limitation of this research is that given that this research only applied standardised testing rules, the testers stopped collecting data on children’s performance after the time limit was reached. Hence we are not able to test this hypothesis comparing the accuracy of childrens responses after the time limit had expired. Future research could explore this hypothesis further.

Test content

Of the subtests of the four scale indexes, the Learning and Sequential Scale subtests present stimuli to which children respond, but do not directly involve the child working with stimuli materials. However in the Simultaneous and Planning Scale subtests the children are required to use materials to solve tasks. The complex figural nature of these tasks which involve different shapes, forms and textures, and a lack of previous exposure to such stimuli given the socio-economic context and impoverished educational contexts, could have resulted in lower scores (Holding & Kitsao-Wekulo, 2004).

To further examine performance on certain sub-scales and to test for potential bias on core subtests, the KABC-II provides supplementary subtests. Supplementary subtests were

included in the battery used in this study, as it was being implemented for the first time in this population. The Planning Scale does not include supplementary subtests, but the Simultaneous Scale does (Block Counting), which was included in the study.

In this research the mean score for Block Counting was the highest of the Simultaneous sub-tests and when used as a substitute for Triangles, the average MPI and Simultaneous sub-scale standard score increased. In examining the content of the test stimuli of the Triangles test we recognise that it may be biased to school exposure or familiarity to test stimuli, as it includes working with specific shapes to form new shapes. Literature shows that children are often taught how to identify shapes but not to recognize and form embedded shapes (using shapes to form a larger picture) (Verdine, Lucca, Golinkoff, Hirsh-Pasek, & Newcombe, 2016). Potentially children in this sample, given their socio-demographic characteristics, poor socio-economic status and limited school exposure, may have been more familiar with games involving wooden blocks in their daily lives. The Indian cohort did not administer Block Counting, but it was similarly noted that children struggled with the Triangles subtests, attributed to potentially having limited exposure to the concept of puzzles (M Malda et al., 2008). While there may be cultural differences between this sample and the Indian cohort, there are similarities in terms of limited school exposure and low socio-economic status.

Block Counting has been shown to be a good measure of the Simultaneous sub-scale and is an acceptable substitution for a core sub-test (M. R. Reynolds et al., 2007). While children in this sample performed substantially better on Block counting than on Triangles, two cautions are important. Firstly Block Counting has not been widely used with children in LMIC; and secondly, the CFA model four in this research (included Block Counting as a replacement for Triangles), was not a better fit than model three. We thus conclude that there is not sufficient

evidence in this research to support the generalised substitution of Block Counting for Triangles, but there is enough evidence that it could be used as an additional sub-test, as Block Counting may allow children an additional opportunity to demonstrate simultaneous processing capacity using more familiar concepts.

While it is not possible to directly compare the performance on the Planning sub-scale sub-tests to alternative sub-tests, we are able to review the literature. Studies which have purposefully substituted Story Completion for an alternative subtest are scarce. However this was done in the Indian cohort, where it was asserted that children were unfamiliar with many of the stories/situations depicted in the subtest (M Malda et al., 2008). In that study, Story Completion was replaced with Picture Arrangement (from the WISC), where each item consists of a series of pictures depicting a story but are in the incorrect order and the child must arrange them in order. The authors (M Malda et al., 2008) acknowledged that while Picture Arrangement seemed to be less culture-specific, items still needed modification.

There has not been extensive enough research on the Story Completion items in African contexts to be able to gauge whether children are disadvantaged by the test content. Furthermore, as the WISC is a more traditional assessment including verbal and acquired knowledge, we would suggest not replacing Story Completion with Picture Arrangement. However administering additional Planning subtests from different batteries in other African settings may allow insight into whether the KABC-II Planning sub-tests show bias to school or test stimuli exposure. An alternative assessment battery which could be useful in administering in addition to the Planning Subtests is the Developmental Neuropsychological Assessment (NEPSY) (Korkman, Kirk, & Kemp, 1998), which predominantly assesses domains of executive functioning.

Skill-related factors

In addition to test-related content influencing performance, children could have performed poorly on these subtests because they have not fully developed the skills necessary to complete these tasks (see Table 1 for the skills required for each subscale). The higher means obtained by children in this sample for Rover and Block Counting may indicate that there is not in a deficit in the visual-spatial abilities required to solve the Simultaneous Scale tasks, but that Triangles was a difficult subtest for most children. In contrast, the performance on the Planning Scale subtests was consistently low, which could reflect underdeveloped executive functioning skills or capacity in this context.

Thus we propose that the Planning subtests are valid in this context, and that the lower scores reflect underdeveloped executive functioning skills. Drawing on literature, it has been well established that the emergence of executive functioning skills are strongly influenced by a child's environment, and that living in the context of poverty and its related risk factors and consequences (on growth, nutrition, other genetic and epigenetic factors) adversely impacts on the acquisition of these executive functioning abilities (Martean & Hall, 2013). Further research would need to include longitudinal studies measuring environment-related factors which influence the emergence of executive functioning skills in children, and how these impact on their performance on KABC-II Planning subtests. This would aid in disentangling whether the lower means on these subtests are reflecting a cultural bias or a skills deficit in this sample.

Limitations of this research included; not piloting the no-time-points version of sub-tests restricted our ability to examine the effects of time points on performance. Furthermore, we were limited in our ability to examine potential cultural influences on the performance on the Planning sub-tests as there was no supplementary sub-test available and no other sub-test measuring

Planning was administered. Additionally, it was beyond the scope of this paper to examine environment-related factors associated with the emergence of executive functioning, and how these potentially influenced performance on the Planning subtests.

Conclusions

Our results showed the KABC-II to have good reliability and to maintain its original structure when used in the sample. Therefore, it can be concluded that, with appropriate translation and inclusion of supplementary subtests, the KABC-II is an appropriate assessment to use in this rural Zulu context. Our results show the importance of administering a whole battery, allowing more insight into performance on different domains of cognition. This includes administering Block Counting in addition to Triangles and in the absence of a supplementary sub-test for the Planning sub-scale, sub-tests from other batteries could be added to strengthen the measurement of the Planning sub-scale.

Acknowledgements

We are grateful to all women and children enrolled in the study, the field and data management team. We thank Ant Snyman, Siyabonga Nxumalo, Dickman Gareta and Kobus Herbst and the research platform management team for support during the analysis. We thank the Community Engagement Unit of the African Health Research Institute and Community Advisory Board for their guidance throughout the study.

Conflict of interest

None.

Funding

The Siyakhula Cohort was funded by Grand Challenges Canada, Saving Brains (Grand Challenges 0063-03). The Africa Centre for Population Health, now called the Africa Health Research Institute (AHRI), where the research took place is funded by the Wellcome Trust (Grant Numbers: Previous Africa Centre 097410/Z/11/Z; Current AHRI 201433/Z/16/Z). The Demographic Surveillance is co-funded by the South African Department of Science and Technology through the DST/MRC South African Population Research Infrastructure Network (SAPRIN). Tamsen RoCHAT is supported by the Newton Advanced Fellowship Scheme (AF160108). Mark Tomlinson is supported by the National Research Foundation, South Africa and is a Lead Investigator of the Centre of Excellence in Human Development (OPP20160035), University Witwatersrand, South Africa. The support of the DST-NRF Centre of Excellence in Human Development towards data analysis is also acknowledged. We are grateful to Pearson Ltd for waiving the translation licence fee for the KABC assessment, and to Alan and Nadeen Kaufman for expert input on subtest selection.

References

- Bangirana, P., Seggane, M., Allebeck, P., Giordani, B., John, C. C., Opoka, O. R., . . . Boivin, M. J. (2009). A preliminary examination of the construct validity of the KABC-II in Ugandan children with a history of cerebral malaria. *African health sciences*, 9(3), 186-192.
- Baumgartner, J., Smuts, C. M., Malan, L., Kvalsvig, J., van Stuijvenberg, M. E., Hurrell, R. F., & Zimmermann, M. B. (2012). Effects of iron and n-3 fatty acid supplementation, alone and in combination, on cognition in school children: A randomized, double-blind, placebo-controlled intervention in South Africa. *The American Journal of Clinical Nutrition*, 96(6), 1327-1338. doi:10.3945/ajcn.112.041004
- Benson, E. (2003). Intelligence across cultures: Research in Africa, Asia and Latin America is showing how culture and intelligence interact. *American Psychological Association Monitor*, 34(2), 56.
- Bland, R. M., Coovadia, H. M., Coutsoodis, A., Rollins, N. C., & Newell, M. L. (2010). Cohort profile: Mamanengane or the Africa centre vertical transmission study. *International Journal of Epidemiology*, 39(2), 351-360. doi:10.1093/ije/dyp165
- Boivin, M., Gladstone, M. J., Vokhiwa, M., Birbeck, G. L., Magen, J. G., Page, C., . . . Taylor, T. E. (2011). Developmental outcomes in Malawian children with retinopathy-positive cerebral malaria. *Tropical Medicine & International Health*, 16(3), 263-271. doi:10.1111/j.1365-3156.2010.02704.x
- Boivin, M., Okitundu, D., Makila-Mabe Bumoko, G., Sombo, M. T., Mumba, D., Tylleskar, T., . . . Tshala-Katumbay, D. (2013). Neuropsychological effects of konzo: A neuromotor

- disease associated with poorly processed cassava. *Pediatrics*, 131(4), 1231-1239.
doi:10.1542/peds.2012-3011d
- Britto, P. R., Lye, S. J., Proulx, K., Yousafzai, A. K., Matthews, S. G., Vaivada, T., . . . Bhutta, Z. A. (2016). Nurturing care: promoting early childhood development. *The Lancet*, 389(10064), 91-102. doi:10.1016/s0140-6736(16)31390-3
- Holding, P., & Kitsao-Wekulo, P. K. (2004). Describing the burden of malaria on child development: What should we be measuring and how should we be measuring it? *The American Journal of Tropical Medicine and Hygiene*, 71(Suppl. 2), 71-79.
- Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods*, 6(1), 53-60.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55. doi:10.1080/10705519909540118
- Jackson, D. L., Gillaspay, J. A., & Purc-Stephenson, R. (2009). Reporting practices in confirmatory factor analysis: An overview and some recommendations. *Psychological Methods*, 14(1), 6-23. doi:10.1037/a0014694
- Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Assessment Battery for Children: Second Edition (KABC-II)*. Minnesota, MN: AGS Publishing.
- Kline, R. B. (2011). *Principles and practice of structural equation modeling, third edition*. New York, NY: Guilford Press.
- Korkman, M., Kirk, U., & Kemp, S. (1998). *NEPSY: A Developmental Neuropsychological Assessment*. San Antonio, TX: Psychological Corporation.

- Laher, S., & Cockcroft, K. (2017). Moving from culturally biased to culturally responsive assessment practices in low-resource, multicultural settings. *Professional Psychology: Research and Practice*, 48(2), 115.
- Lu, C., Black, M. M., & Richter, L. M. (2016). Risk of poor development in young children in low-income and middle-income countries: an estimation and analysis at the global, regional, and country level. *The Lancet Global Health*, 4(12), e916-e922.
doi:10.1016/s2214-109x(16)30266-2
- Malda, M., Van de Vijver, F., Srinivasan, K., Transler, C., Sukumar, P., & Rao, K. (2008). Adapting a cognitive test for a different culture: An illustration of qualitative procedures. *Psychology Science Quarterly*, 50(4), 451-468.
- Malda, M., van de Vijver, F. J. R., Srinivasan, K., Transler, C., & Sukumar, P. (2010). Traveling with cognitive tests: Testing the validity of a KABC-II adaptation in India. *Assessment*, 17(1), 107-115. doi:10.1177/1073191109341445
- Marteau, T. M., & Hall, P. A. (2013). Breadlines, brains, and behaviour. *BMJ : British Medical Journal*, 347, f6750. doi:10.1136/bmj.f6750
- Mitchell, J. M., Rochat, T. J., Houle, B., Stein, A., Newell, M. L., & Bland, R. M. (2016). The effect of maternal and child early life factors on grade repetition among HIV exposed and unexposed children in rural KwaZulu-Natal, South Africa. *Journal of Developmental Origins of Health and Disease*, 7(2), 185-196. doi:10.1017/S2040174415007230
- Moon, S.-B., Mclean, J. E., & Kaufman, A. S. (2003). A Cross-Cultural Validation of the SequentialSimultaneous Theory of Intelligence in Children. *School Psychology International*, 24(4), 449-461. doi:10.1177/01430343030244007

- Morgan, K. E., Rothlisberg, B. A., McIntosh, D. E., & Hunt, M. S. (2009). Confirmatory factor analysis of the KABC-II in preschool children. *Psychology in the Schools*, 46(6), 515-525. doi:10.1002/pits.20394
- O'Donnell, K., Murphy, R., Ostermann, J., Masnick, M., Whetten, R. A., Madden, E., . . . Whetten, K. (2012). A brief assessment of learning for orphaned and abandoned children in low and middle income countries. *AIDS and Behavior*, 16(2), 480-490. doi:10.1007/s10461-011-9940-z
- Oakland, T. (2009). How universal are test development and use? In E. L. Grigorenko (Ed.), *Multicultural psychoeducational assessment*. New York, NY: Springer Publishing Company.
- Ogunlade, A. O., Kruger, H. S., Jerling, J. C., Smuts, C. M., Covic, N., Hanekom, S. M., . . . Kvalsvig, J. (2011). Point-of-use micronutrient fortification: Lessons learned in implementing a preschool-based pilot trial in South Africa. *International Journal of Food Sciences and Nutrition*, 62(1), 1-16. doi:10.3109/09637486.2010.495710
- Reynolds, C. R., & Ramsay, M. C. (2003). Bias in Psychological Assessment: An Empirical Review and Recommendations. In *Handbook of Psychology*: John Wiley & Sons, Inc.
- Reynolds, M. R., Keith, T. Z., Fine, J. G., Fisher, M. E., & Low, J. A. (2007). Confirmatory factor structure of the Kaufman Assessment Battery for Children- Second Edition: Consistency with Cattell-Horn-Carroll theory. *School Psychology Quarterly*, 22(4), 511. doi:10.1037/1045-3830.22.4.511
- Richter, L. M., Griesel, R. D., & Rose, C. (1994). The McCarthy Scales of Children's Abilities: Adaptation and norms for use amongst black South African children. *South African Journal of Occupational Therapy*, 24, 17-30.

- Rochat, T., Houle, B., Stein, A., Pearson, R., Newell, M.-L., & Bland, R. (2016). Cohort profile: Siyakhula or the grand challenges africa centre saving brains cohort in South Africa. *International Journal of Epidemiology*, 1-17. doi.org/10.1093/ije/dyx148
- Rosselli, M., & Ardila, A. (2003). The impact of culture and education on non-verbal neuropsychological measurements: A critical review. *Brain and Cognition*, 52(3), 326-333. doi:https://doi.org/10.1016/S0278-2626(03)00170-2
- StataCorp, L. (2007). *Stata data analysis and statistical Software*, Release, 10 (Special Edition Release). College Station, TX: StataCorp LP
- Taljaard, C., Covic, N. M., Van Graan, A. E., Kruger, H. S., Smuts, C. M., Baumgartner, J., . . . Jerling, J. C. (2013). Effects of a multi-micronutrient-fortified beverage, with and without sugar, on growth and cognition in South African schoolchildren: A randomised, double-blind, controlled intervention. *British Journal of Nutrition*, 110(12), 2271-2284. doi:10.1017/S000711451300189X
- Tanser, F., Hosegood, V., Barnighausen, T., Herbst, K., Nyirenda, M., Muhwava, W., . . . Newell, M.-L. (2008). Cohort profile: Africa Centre Demographic Information System (ACDIS) and population-based HIV survey. *International Journal of Epidemiology*, 37(5), 956-962. doi:10.1093/ije/dym211
- Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington, DC: American Psychological Association.
- Van Eeden, R. (1997). *Senior South African Individual Scale-Revised*. Pretoria: Human Sciences Research Council.
- Veena, S. R., Krishnaveni, G. V., Wills, A. K., Kurpad, A. V., Muthayya, S., Hill, J. C., . . . Srinivasan, K. (2010). Association of birthweight and head circumference at birth to

cognitive performance in 9-10 year old children in South India: Prospective birth cohort study. *Pediatric research*, 67(4), 424-429. doi:10.1203/PDR.0b013e3181d00b45

Verdine, B. N., Lucca, K. R., Golinkoff, R. M., Hirsh-Pasek, K., & Newcombe, N. S. (2016).

The Shape of Things: The Origin of Young Children's Knowledge of the Names and Properties of Geometric Forms. *Journal of cognition and development : official journal of the Cognitive Development Society*, 17(1), 142-161.

doi:10.1080/15248372.2015.1016610

Zaidi, J., Grapsa, E., Tanser, F., Newell, M.-L., & Bärnighausen, T. (2013). Dramatic increases in HIV prevalence after scale-up of antiretroviral treatment: a longitudinal population-based HIV surveillance study in rural kwazulu-natal. *AIDS (London, England)*, 27(14), 2301-2305. doi:10.1097/QAD.0b013e328362e832