

Validity of Neuropsychological Testing in Young African Children Affected by HIV

Miriam C. Chernoff¹ Barbara Laughton² Mmule Ratswana³ Itziar Familiar⁴ Lee Fairlie³
Tichaona Vhembo⁵ Portia Kamthunzi⁶ Enid Kabugho⁷ Celeste Joyce⁸ Bonnie Zimmer⁹
J. L. Ariansen¹⁰ Patrick Jean-Philippe¹¹ Michael J. Boivin^{4,12,13}

¹Center for Biostatistics in AIDS Research, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, United States

²Family Clinical Research Unit, Department of Paediatrics and Child Health, Tygerberg Hospital, Stellenbosch University, Tygerberg, RSA, South Africa

³Wits Reproductive Health & HIV Institute (WRHI), Shandukani Clinic, Johannesburg, RSA, South Africa

⁴Department of Psychiatry, Michigan State University, East Lansing, Michigan, United States

⁵Harare Family Care CRS, University of Zimbabwe, College of Health Sciences Clinical Trials Unit, Harare, Zimbabwe

⁶University of North Carolina Project– Lilongwe, Malawi CRS, Malawi

⁷Makerere University–Johns Hopkins University Research Collaboration (MUJHU CARE LTD) CRS, Kampala, Uganda

⁸University of Witwatersrand, RSA, Perinatal HIV Research Unit, Soweto RSA, South Africa

⁹Frontier Science Foundation, Amherst, New York, United States

¹⁰Clinical Research Management, FHI360, Durham, North Carolina, United States

¹¹Maternal Adolescent Pediatric Research Branch, DAIDS/NIAID/NIH, Rockville, Maryland, United States

¹²Department of Neurology and Ophthalmology, Michigan State University, East Lansing, Michigan, United States

¹³Department of Psychiatry, the University of Michigan, Ann Arbor, Michigan, United States

Address for correspondence Miriam C. Chernoff, PhD, Center for Biostatistics in AIDS Research, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, United States (e-mail: mchernoff@sdac.harvard.edu).

J Pediatr Infect Dis 2018;13:185–201.

Abstract

Introduction Western-constructed neuropsychological tests have been used in low- and middle-income countries to assess the impact of human immunodeficiency virus/acquired immunodeficiency syndrome (HIV/AIDS) and other chronic illnesses. We explore using such instruments cross-culturally in a sub-Saharan African setting.

Methods IMPAACT P1104S was a 2-year observational study performed at six clinical sites (South Africa—three sites, Malawi, Uganda, and Zimbabwe) to assess and compare neuropsychological outcomes in three cohorts of children between the ages of 5 and 11 years: HIV-infected (HIV), HIV-exposed but uninfected (HEU), and HIV unexposed and uninfected (HU). Descriptive statistics compared sociodemographic characteristics among children at sites. Instruments included the Kaufman Assessment Battery for Children, 2nd edition (KABC-II) cognitive ability, Test of Variables of Attention (TOVA) attention/impulsivity, Bruininks–Oseretsky Test of Motor Proficiency, 2nd edition (BOT-2) motor proficiency tests, and Behavior Rating Inventory for Executive Function (BRIEF) executive function problems. Test characteristics were assessed using intraclass and Spearman's nonparametric correlations, linear regression, and principal factor analyses.

Results Of the 611 participants, 50% were males and mean age ranged from 6.6 to 8 years. In Malawi, Uganda, and Zimbabwe, substantial proportions of families lived in rural settings in contrast to the South African sites. Intraclass correlation coefficients between weeks 0 and 48 were highest for the KABC scores, ranging between 0.42 and

Keywords

- ▶ pediatric HIV
- ▶ neuropsychological testing
- ▶ validity
- ▶ reliability
- ▶ Africa

received
October 3, 2017
accepted after revision
February 6, 2018
published online
March 20, 2018

Copyright © 2018 by Georg Thieme
Verlag KG, Stuttgart · New York

DOI <https://doi.org/10.1055/s-0038-1637020>.
ISSN 1305-7707.

0.71. Correlations among similar test domains were low to moderate but significant, with positive correlation between KABC sequential and TOVA scores and negative correlation between BRIEF and KABC scores. TOVA response time scores correlated negatively with the BOT-2 total points score. Strong and significant associations between individual measures of growth, disability, and development with all test scores were observed. Performance-based measures were markedly lower for HIV compared with HEU and HU participants, even after controlling for age, sex, and site. Factor analyses confirmed the underlying theoretical structure of the KABC scaled item scores.

Conclusion The KABC, TOVA, BRIEF, and BOT-2 were valid and reliable tools for assessing the neuropsychological impact of HIV in four sub-Saharan African countries.

Introduction

Access to pediatric prevention and care for human immunodeficiency virus/acquired immunodeficiency syndrome (HIV/AIDS) has expanded throughout the low- and middle-income countries (LMIC), with reduced pediatric infection, morbidity, and mortality, shifting attention to the impact of the HIV virus on neurocognitive development and other longer-term outcomes. As learning and problem-solving approaches may differ across cultures, researchers have questioned whether tests developed in high-income settings are valid elsewhere.^{1–6} Tests need to be reliable, consistent, and robust to challenging testing situations, simple to score and interpret, and sensitive to population variability and changes in health and home environment.^{7–11}

For this study, we chose four Western-developed standard tests, previously used in LMIC HIV/AIDS and other high burden illness settings. These measures have proven sensitive to the neuropsychological effects of pediatric HIV in the sub-Saharan African context, assessing aspects of motor development and proficiency, attention, working memory, learning, visual-spatial analysis and problem solving, and planning and reasoning.^{11–16} This study used the Kaufman Assessment Battery for Children, 2nd edition (KABC-II),¹⁷ the Test of Variables of Attention (TOVA; www.tovatest.com),¹⁸ the Bruininks–Oseretsky Test of Motor Proficiency, 2nd edition (BOT-2),¹⁹ and the Behavior Rating Inventory for Executive Function (BRIEF).²⁰

Some have questioned the validity of adapting “Western-based” tests to the African context, and have recommended instead the development of altogether new neurodevelopmental assessment batteries more appropriate for such settings along with normative data specific to the target culture.⁵ Contrary to this view, Boivin and Giordani have proposed that human neurodevelopment can be characterized within a foundational brain/behavior “omnibus,” universally accessible by neuropsychological assessment batteries (even “Western” ones) that are well designed, carefully adapted, and appropriately applied.^{21,22} To substantiate this view, we evaluate the feasibility, reliability, and validity of cross-cultural neuropsychological testing in a sub-Saharan African setting with children affected by HIV/AIDS.

We hypothesize that these instruments meet adequate test construction criteria and can be used effectively and successfully in cross-cultural neuropsychological research in at-risk African children.

Methods

Study Population and Design

IMPACT P1104S was a 2-year multisite observational study performed at six clinical sites in four sub-Saharan African countries: three in South Africa (Soweto, Tygerberg, and Johannesburg); Kampala (Uganda); Harare (Zimbabwe); and Lilongwe (Malawi). The purpose was to assess and compare neuropsychological outcomes in three cohorts of young children (5–< 11 years): HIV-infected (HIV), HIV-exposed but uninfected (HEU), and HIV-unexposed and uninfected (HU).²³ HIV-infection and exposure was assumed to be perinatal or during breastfeeding. HIV-infected children were eligible if they had previously enrolled and were actively being followed in the Pediatric AIDS Clinical Trials Group (PACTG) 1060, version 5.0, a longitudinal study evaluating nevirapine-based versus lopinavir/ritonavir-based antiretroviral regimens.²⁴ The two P1104S comparison groups (HEU and HU) were recruited from similar neighborhoods as the HIV-infected participants; siblings were encouraged to enroll. Target sample sizes were determined based on the number of eligible P1060 children and a power computation for cross-sectional pairwise group differences. To ensure balance of the two control groups with the HIV-infected participants within sites, site- and age-specific enrollment limits were imposed based on the number and age distribution of the eligible P1060 participants. Each HIV-uninfected cohort was to enroll 182 children (35, < 6 years; 64, 6–< 7 years; and 83, 7–< 11 years). Data for P1104S were collected between September 23, 2013 and December 12, 2016.

Other eligibility criteria included documented evidence of mother’s HIV status during pregnancy (if HIV-infected) and at birth or afterwards (if HIV negative). For the HU control group, exclusion criteria included history of documented brain injury, seizures, or hospitalization from central nervous system (CNS) infection. Local institutional review

boards approved the study and parental/guardian informed consents were obtained. Assent was obtained for children 7 years and older. Study assessments were performed at baseline (week 0), week 48, and week 96; Data from all study weeks are included in the reliability analysis, while only baseline data are included in the validity analysis.

Measures

Sociodemographic characteristics collected included that of the caregiver (education, HIV-status, length of time child was in care of primary caregiver, family income, fuel and water sources, urban or rural living zone, and employment) and child (anthropomorphic measures, illness history, and medications which may influence neuropsychological testing). Children's neurodevelopment was assessed using performance-based measures and subjective assessments. Performance-based measures (e.g., KABC-II, TOVA, and BOT-2) are those in which there is a direct observation of the child performing a specific task using a standardized battery of tests that can include props and toys. Subjective assessments (BRIEF, Multiple Indicators Cluster Survey [MICS]) generally include questionnaires with items relating to the child's behavior that are read out loud to the informant (e.g., parent), who evaluates the extent to which those behaviors describe the child.

Kaufman Assessment Battery for Children, Second Edition (KABC-II)

The KABC-II¹⁷ measures neuropsychological outcomes for ages 3 to 18 years, and has been previously validated in Uganda^{25,26} and South Africa.⁶ In this study, the KABC-II Luria model was used to assesses cognitive abilities in four domains: (1) sequential processing (short-term memory), (2) simultaneous processing (visual-spatial processing and problem solving), (3) learning (immediate and delayed memory), and (4) planning (executive reasoning). All subtests were administered to all participants with instructions in local languages (the child's preferred language). Subtests were administered even if out of age range for the purpose of future analyses by age-band, but only scaled if within appropriate age ranges. Subscale scores and global scores (Mental Processing Index and Nonverbal Index [NVI]) were scaled to a U.S. population by age. The KABC-II takes approximately 1.5 hours to administer. ► **Appendix A** describes the KABC-II subtests. Items of three subtests were administered in English at all sites; Rebus and Rebus delayed and Number recall (designed for one-syllable numbers). As such, the Rebus tests assessed learning in a foreign language. It is common in Anglophone Africa for English numbers to be used and translating would have complicated the subtest by introducing multisyllabic words.

Test of Variables of Attention, Version 8 (TOVA)

The TOVA is a computerized test measuring key components of attention and self-control, variability (consistency), response time (speed), commissions (impulsivity), and omissions (focus and vigilance). Prior studies suggest that the visual TOVA (where a child responds rapidly by pressing a switch when a simple geometric stimulus "signal" appears

on a computer screen) is a sensitive measure of persisting neurocognitive effects among African children surviving cerebral malaria.²⁷⁻²⁹ Following instructions and practice trials, the TOVA takes 22 minutes for children older than 5.5 years. Scoring is based on establishing whether or not response times and attention is at the normal range for the sex and age of the child with attention deficit/hyperactivity disorder (ADHD) scores more negative than -1.8 considered suggestive of ADHD for U.S. norms.

Bruininks-Oseretsky Test of Motor Proficiency, Second Edition (BOT-2), Brief Form

The BOT-2^{19,30} is a standardized test comprehensively assessing gross and fine motor skills in children between the ages of 4 and 21 years through eight subtests: fine motor precision, fine motor integration, manual dexterity, bilateral coordination, balance, running speed and agility, upper limb coordination, and strength. The Brief Form of the BOT-2 has 12 items, each a game-like task, and takes approximately 15 minutes in all to administer. Composite scores include fine manual control, manual coordination, body coordination, strength and agility, and total composite score. Total scores (0-72) are standardized to U.S. norms by age.

Behavior Rating Inventory for Executive Function

The BRIEF Parent Form for school-age children (5-18 years) was especially designed to evaluate executive function behaviors through 86 items rated as Never = 1, Sometimes = 2, and Often = 3, and takes approximately 15 to 30 minutes to administer.³¹ The BRIEF has eight nonoverlapping clinical scales and two statistically derived indexes: (1) Behavioral Regulation Index (BRI; sum of Inhibit, Shift, and Emotional Control), and (2) Metacognition Index (MI; sum of Initiate, Working Memory, Plan/Organize, Organization of Materials, and Monitor). The Global Executive Composite (GEC) score is the sum of the BRI and MI. BRIEF standardized scores are calculated based on the scores of other respondents from a standardized sample, by age and sex. Higher scores suggest a higher level of dysfunction in a specific domain. A negativity and inconsistency score is also obtained which assists with determining the validity of responses.

MICS-4 Early Childhood Development and Childhood Disability Scales

We used the Early Childhood Development portion of the Questionnaire for Children Under 5 (17 items) and Questionnaire Form for Child Disability from the MICS research program, which is sponsored by United Nations International Children's Education Fund (UNICEF).³² The primary caregiver answered the questionnaire and selected questions were scored and summed to develop an overall Disability score (scale, 0-100) and an overall Development score (scale, 0-100), with higher values indicating better child development and worse child disability, respectively.

Procedures

Standardized operating procedures were developed for each battery of tests, including test administration, scoring, and

quality assurance. The KABC-II, TOVA, and BOT-2 tests were designed to be readily adaptable to cross-cultural settings and instructions for the various tasks and items can be spoken by the examiner to the child in the local language or in English (the official educational language at some study sites) during administration, without the need for formal written translation or written explanation for the child. Each subtest of the KABC-II has example “teaching” items to ensure that the child understands the task required. In fact, the KABC-II has a NVI score compiled from the subtests for which language fluency is not required (even administrable by pantomime), and this global performance composite was a principal outcome for our analyses. Likewise, the BOT-2 items can be demonstrated by modeling on the part of the tester, and the TOVA relies on simple geometric figures for the signal and nonsignal stimuli, not on letters or numbers or specific images as is the case in other continuous performance monitoring tests of attention.

The 86 items of the BRIEF questionnaire, read out loud to the parent or principal caregiver of the child, were formally translated by permission of the publisher. The publisher provided translated versions for Luganda, Afrikaans, and Xhosa. For Shona, Zulu, Chichewa, Sesotho, and Setswana, instructions and items were translated (forward and backward) and approved by the authors of the BRIEF, before being administered by trained assessors. The Hopkins Symptom Checklist (HSLC), MICS-4 disability, and MICS-4 child development and household socioeconomic questionnaires were translated into local languages by the sites and administered as interviews. Where possible, tests were administered by testers who were blinded to the child’s HIV status. The TOVA was administered first so that attention was assessed before the child was tired, and also that visual deficits for study children would immediately become evident, allowing for adaptation of subsequent assessments to limited vision capability if necessary. The battery was mostly completed in 1 day unless a child appeared tired, requiring a second session to complete the assessment (within 1 week).

Testing quality assurance procedures included a 1-week training by coauthor M.J.B., weekly practice sessions over a 1-month period by all testers, and a quality assurance plan required each month for every tester (KABC-II video evaluation, team self-video evaluation, standard operating procedures review for all tests). KABC-II testing quality was scored according to a consistent rubric and the scores were tracked monthly for every tester at every site by independent KABC-II testing experts at the Global Health Uganda assessment center in Kampala. The TOVA was automated and the BOT-2 was a straightforward test of motor proficiency so these were not evaluated as carefully as the KABC-II, but they also had standard operating procedures that were to be reviewed by testers monthly. These will be described in depth in a separate manuscript.

Statistical Analysis

Data were retrieved on May 3, 2017. Sociodemographic characteristics were summarized by site and compared using chi-square, analysis of variance, and Wilcoxon rank-sum tests. A

socioeconomic index was based on seven characteristics: fuel, water, refrigerator, caregiver’s work status, social grant main source of income, caregiver education, and sufficiency of household income to meet basic needs. Each item was scored from 0 to 10, and a mean score was computed with 10 reflecting the highest socioeconomic status.

Feasibility was assessed through test completion rates and the ability to complete the tests in 1 day without significant interruptions. Test–retest reliability was assessed through absolute (mean difference, upper and lower limits of agreement, and the coefficient of repeatability) and relative (Pearson’s product–moment correlation, intraclass correlation coefficient [ICC]) measures.^{33–35} Absolute and relative measures assessed the consistency and agreement of test scores from week 0 to week 48 and from week 48 to week 96 (differences were computed as the value at the most recent week minus the value at the earlier week). ICCs were computed using linear mixed models regression with participants nested within cohorts, both unadjusted and adjusted for site, participant age at entry, and sex. Models used restricted maximum likelihood (REML) estimation. Test–retest analyses were repeated on the HIV-infected cohort as a sensitivity analysis to explore potential effects of change in disease status.

An exploratory principal factor analysis was performed on 13 of the 15 KABC-II week 0 items for the full sample and by HIV status (HIV-infected, HIV-uninfected), using the squared multiple correlations as initial communality estimates and varimax orthogonal rotation, with a four-factor model based on the theoretical underlying structure. Factor loadings, percent of common variance, and percent of total variance were summarized. Prior to carrying out the factor analysis, we explored the data distributions graphically and tested for normality using the Shapiro–Wilk test.

Construct validity was assessed using baseline (week 0) data. First, correspondence (convergence) validity among test domains was assessed using linear regression models and generalized estimating equations (GEEs) for robust error estimates. Models, adjusted just for cohort (unadjusted) and also for age, sex, and site (adjusted), were run pooling all study participants, and then separately by group (HIV-infected; comparison, pooling HEU and HU). Parameter estimates and 95% confidence intervals (CIs) were estimated. We also computed Spearman’s rank correlation coefficients for the unadjusted data. Prior cross-cultural work suggests that measures of attention (e.g., TOVA D-prime), memory (e.g., KABC sequential processing), and visual–spatial processing would be associated, that working memory and attention would reflect executive function (assessed in this study through the BRIEF), and that sequential processing and mental processing indices would be inversely correlated with TOVA error scores (impulsivity, inattention).^{5,36,37} We were also interested in assessing comparability of performance based (KABC, TOVA) with parental evaluations (BRIEF).

Second, we assessed whether neuropsychological scores correlated with HIV status, growth, quality of the home environment, and caregiving (child’s development,

disability, and socioeconomic status). For each characteristic and test score, a linear regression model using GEE methods was fit over all cohorts and then separately for the HIV and comparison (pooling HU and HEU) cohorts. Unadjusted regressions and those adjusted for site, sex, and age at study entry were performed. Parameter estimates and their standard errors and significance levels were used to indicate degree of associations. Statistical significance was determined with two-sided tests at $\alpha = 0.05$. All analyses were performed with SAS version 9.4.

Results

Of the original 615 enrolled, 611 participants (246 HIV, 183 HEU, and 182 HU) completed the baseline study visit and were eligible for analysis (►Fig. 1). Of these, 603 (98.7%)

completed the week 48 visit and 588 (96.2%) completed the week 96 visit; 23 (3.8% of 611) discontinued prematurely (prior to week 96; 4 HIV-infected, 11 HEU, 8 HU). Test completion rates were 95% or higher at each study visit with over 91% of children completing test batteries in 1 day. Although only 7% of caregivers completed the entry visit batteries in 1 day, on subsequent visits over 95% did so. Over 97% of child and caregiver tests were considered to be valid by the test administrators.

Sociodemographic Characteristics

►Table 1 shows personal and family characteristics by study site, pooling cohorts. By design, there were approximately 40% HIV and 30% each HEU and HU at each study site. About 50% were male. Mean age ranged from 6.6 to 8 years. Johannesburg and Soweto enrolled older participants while

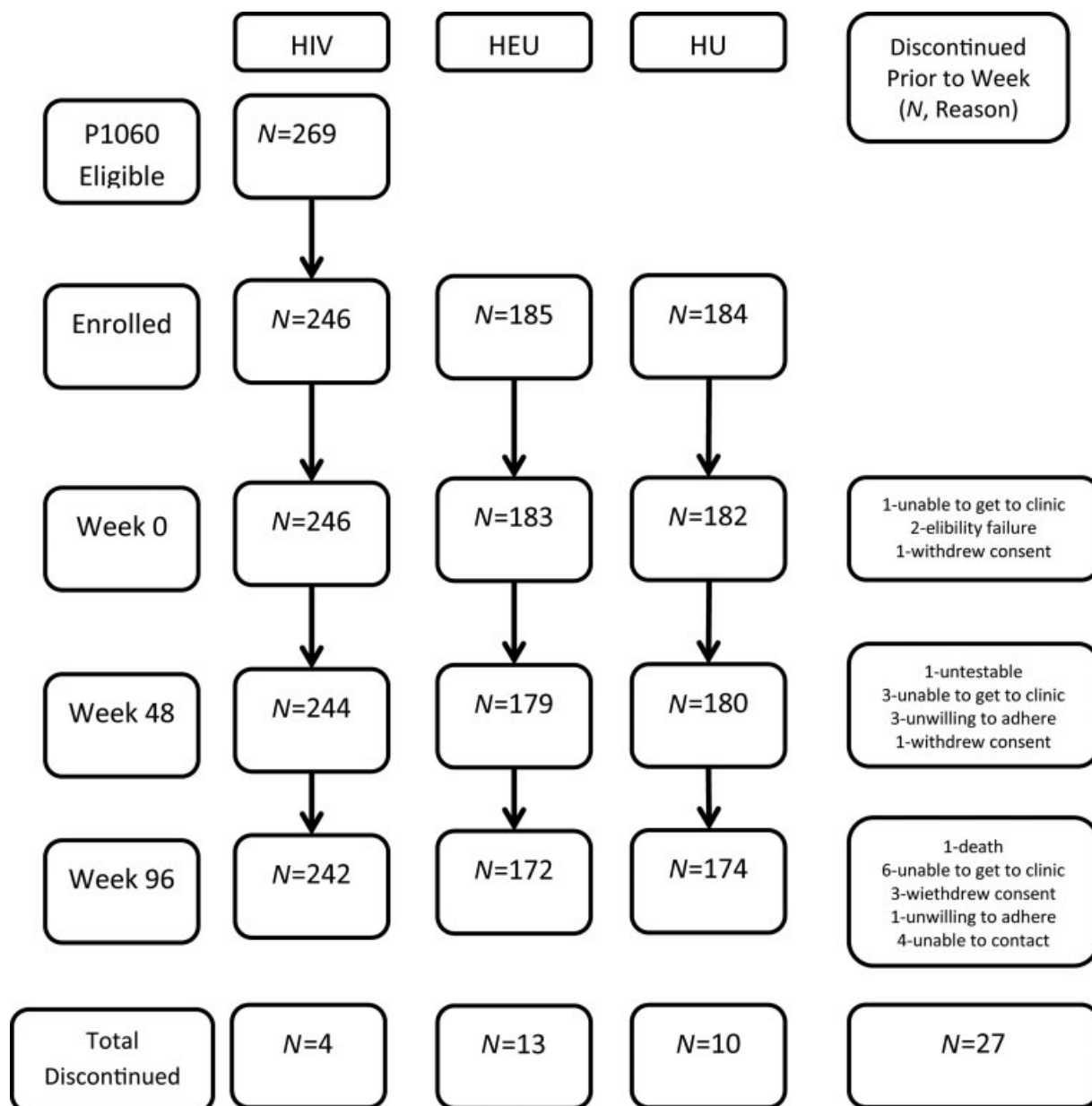


Fig. 1 The number of participants at each P1104S study visit and reasons for discontinuations. HEU, HIV-exposed, uninfected; HIV, HIV-infected cohort; HU, HIV unexposed, uninfected.

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

Table 1 Child, caregiver, and family characteristics at study entry by site

Characteristic	Johannesburg (N = 69)	Soweto (N = 100)	Tygerberg (N = 140)	Malawi (N = 81)	Uganda (N = 89)	Zimbabwe (N = 132)	Total (N = 611)	p-Value ^a
Design characteristics								
Cohort ^b								
HIV	27 (39%)	42 (42%)	56 (40%)	33 (41%)	37 (42%)	51 (39%)	246 (40%)	1.000
HEU	21 (30%)	29 (29%)	42 (30%)	24 (30%)	26 (29%)	41 (31%)	183 (30%)	
HU	21 (30%)	29 (29%)	42 (30%)	24 (30%)	26 (29%)	40 (30%)	182 (30%)	
Child characteristics								
Sex								
Male	32 (46%)	50 (50%)	65 (46%)	43 (53%)	49 (55%)	51 (39%)	290 (47%)	0.182
Race								
Black African	69 (100%)	100 (100%)	97 (69%)	81 (100%)	89 (100%)	132 (100%)	568 (93%)	<0.001
Colored/White/ Other	0 (0%)	0 (0%)	43 (31%)	0 (0%)	0 (0%)	0 (0%)	43 (7%)	
Age (y)								
Mean (SD)	8.0 (1.6)	7.9 (1.3)	7.4 (1.3)	6.6 (0.9)	6.7 (1.2)	6.8 (1.4)	7.2 (1.4)	<0.001
<6 y	8 (12%)	8 (8%)	11 (8%)	18 (22%)	22 (25%)	45 (34%)	112 (18%)	<0.001
6-< 7 y	12 (17%)	21 (21%)	58 (41%)	44 (54%)	37 (42%)	39 (30%)	211 (35%)	
≥ 7 y	49 (71%)	71 (71%)	71 (51%)	19 (23%)	30 (34%)	48 (36%)	288 (47%)	
In school at time of KABC test	58 (84%)	84 (86%)	94 (71%)	55 (77%)	41 (46%)	74 (57%)	406 (69%)	<0.001
Unknown/NA	0	2	8	10	0	3	23	
WHO weight z-score	-0.31 (1.11)	-0.19 (1.09)	-0.18 (1.06)	-1.14 (0.76)	-0.52 (0.83)	-0.48 (0.87)	-0.44 (1.01)	<0.001
WHO height z-score	-0.64 (1.04)	-0.57 (1.05)	-0.38 (0.95)	-1.55 (0.98)	-0.83 (0.96)	-0.38 (1.00)	-0.66 (1.06)	<0.001
WHO BMI z-score	0.08 (0.91)	0.19 (1.04)	0.05 (0.99)	-0.20 (0.73)	0.01 (0.82)	-0.40 (1.02)	-0.06 (0.96)	<0.001
MICS disability score	3.77 (6.21)	4.60 (9.37)	4.29 (8.49)	5.31 (10.01)	10.56 (11.52)	5.76 (8.02)	5.65 (9.25)	<0.001
MICS development score	76.35 (12.23)	84.09 (10.89)	86.28 (12.16)	50.42 (13.92)	69.25 (15.14)	73.17 (13.13)	74.73 (17.14)	<0.001
Caregiver characteristics								
Caregiver is biological mother	66 (96%)	92 (92%)	137 (98%)	75 (93%)	80 (90%)	122 (92%)	572 (94%)	0.18
Caregiver for 5 or more y	61 (88%)	92 (92%)	137 (98%)	74 (91%)	79 (89%)	120 (91%)	563 (92%)	0.09
HIV status of primary caregiver								
HIV-uninfected	21 (31%)	29 (30%)	43 (31%)	28 (35%)	31 (35%)	41 (31%)	193 (32%)	0.96
HIV-infected	47 (69%)	67 (70%)	97 (69%)	53 (65%)	57 (65%)	91 (69%)	412 (68%)	
Not reported	1	4	0	0	1	0	6	
Socioeconomic characteristics								
Current residential zone								
Rural	7 (10%)	3 (3%)	1 (1%)	34 (42%)	25 (28%)	39 (30%)	109 (18%)	<0.001
Periurban	30 (43%)	10 (10%)	138 (99%)	27 (33%)	60 (67%)	3 (2%)	268 (44%)	
Urban	32 (46%)	87 (87%)	1 (1%)	20 (25%)	4 (4%)	90 (68%)	234 (38%)	

Table 1 (Continued)

Characteristic	Johannesburg (N = 69)	Soweto (N = 100)	Tygerberg (N = 140)	Malawi (N = 81)	Uganda (N = 89)	Zimbabwe (N = 132)	Total (N = 611)	p-Value ^a
Major income source								
Mother	24 (35%)	22 (22%)	28 (20%)	30 (37%)	45 (51%)	34 (26%)	183 (30%)	<0.001
Father	16 (23%)	10 (10%)	49 (35%)	43 (53%)	36 (40%)	81 (61%)	235 (39%)	
Social grant	23 (33%)	64 (65%)	43 (31%)	0 (0%)	0 (0%)	0 (0%)	130 (21%)	
Other person/ source	6 (9%)	3 (3%)	20 (14%)	8 (10%)	8 (9%)	17 (13%)	62 (10%)	
Missing	0	1	0	0	0	0	1	
Socioeconomic index ^d	7.1 (5.7, 8.3)	5.7 (5.7, 7.1)	6.4 (5.0, 8.6)	3.0 (2.4, 3.9)	5.8 (4.5, 6.8)	6.2 (4.6, 7.6)	6.0 (4.6, 7.6)	<0.001 ^c

Abbreviations: BMI, body mass index; HIV, human immunodeficiency virus; KABC, Kaufman Assessment Battery for Children; MICS, Multiple Indicators Cluster Survey; SD, standard deviation; WHO, World Health Organization.

^aUnless indicated, differences among sites for categorical measures are tested using the chi-square test and for continuous measures, using the analysis of variance.

^bHIV = HIV-infected; HEU = HIV-exposed, uninfected; HU = HIV-unexposed, uninfected.

^cKruskal-Wallis test.

^dThe socioeconomic index is a composite of caregiver education, fuel and water sources, sufficiency of household income for basic needs, caregiver education and work status, major source of income, and whether household has a refrigerator. Values range from 0 to 10.

Malawi, Uganda, and Zimbabwe enrolled younger participants ($p < 0.001$). Weight and height z-scores were especially low at Malawi (-1.14 and -1.55, respectively, $p < 0.001$). MICS disability scores were highest at Uganda (10.56, $p < 0.001$) compared with the other sites. Development scores were lowest at Malawi (50.4; $p < 0.001$). Over 90% of caregivers across sites were biological mothers at entry and over 88% had been primary caregivers for 5 years or more at that time. Close to 70% of primary caregivers across the three arms were HIV-infected.

Among the HIV cohort, several of the HIV disease characteristics differed across study sites, in particular age at antiretroviral (ARV) initiation (youngest at Soweto and Tygerberg, oldest at Malawi) and length of time on ARVs (longest at the South African sites), World Health Organization (WHO) disease stage (most severe at Johannesburg and Zimbabwe, least severe at Soweto; ► **Supplementary Table S1**, online only).

In Malawi, Uganda, and Zimbabwe, substantial proportions of families lived in rural settings in contrast to the South African sites. In Johannesburg and Malawi, nonrural families were split equally between urban and periurban settings, while in the other sites, the nonrural families lived primarily in one setting ($p < 0.001$). The composite socioeconomic indicator showed that participants from Johannesburg had better socioeconomic circumstances than the others, with Malawi the least resourced ($p < 0.001$; ► **Table 1**).

Test-Retest Reliability

Unadjusted ICC measures between weeks 0 and 48 were highest for the KABC scaled scores. Except for delayed recall, these values ranged from 0.54 to 0.73. The ICC for the BOT-2 total points score was 0.69. BRIEF ICC measures ranged from 0.50 (Monitor), 0.52 (BRI) to 0.56 (MI) and 0.57 (GEC). The ICC values for the TOVA D-prime and D-prime Standard were 0.63 and 0.45, respectively (► **Table 2**). Pearson's correlations were very similar to the unadjusted ICCs (► **Table 2**), while adjusted ICC measures were slightly lower (data not shown).

Both Pearson's correlation and ICCs between measures at week 48 and week 96 were slightly higher than those between weeks 0 and 48 (► **Supplementary Table S2**, online only). In few cases, these differences were substantial. For example, whereas the Pearson correlation between weeks 0 and 48 for KABC learning was 0.56, the same correlation between weeks 48 and 96 was 0.82.

Mean difference scores varied by time and test. For example, mean scores were stable between weeks 0 and 48 (as indicated by the 95% CI for the mean difference including zero) for the KABC NVI and the KABC Mental Processing Index, while there were small but statistically significant increases for the KABC simultaneous scores and decreases for the KABC delayed recall. During this same time frame, BRIEF scores decreased on average, indicating improved function (► **Table 2**). Most of the KABC scores as well as the TOVA ADHD and D-prime scores increased slightly between weeks 48 and 96 while BRIEF scores were stable (► **Supplementary Table S2**, online only). By definition, 95% of individual difference scores would lie between the

Table 2 Absolute and relative measures of agreement for weeks 0 and 48

Outcome	Week 0 N, mean (SD)	Week 48 N, mean (SD)	Pearson ^a	Unadjusted ICC ^b (95% CI)	Mean difference(95% CI)	Limits of agreement ^c	CR ^d
Kaufman Assessment Battery for Children (KABC-II), summary scores							
Sequential	605 83.32 (12.12)	601 82.92 (12.44)	0.72	0.71 (0.67, 0.75)	-0.39 (-1.13, 0.36)	-18.46, 17.69	18.45
Simultaneous	606 77.72 (12.41)	602 79.90 (13.20)	0.66	0.65 (0.61, 0.70)	2.05 (1.20, 2.89)	-18.60, 22.70	21.07
Learning	605 86.69 (16.02)	601 86.34 (14.75)	0.56	0.54 (0.48, 0.59)	-0.35 (-1.51, 0.82)	-28.66, 27.97	28.90
Planning	285 76.08 (11.87)	474 76.59 (13.20)	0.75	0.69 (0.64, 0.74)	1.92 (0.79, 3.06)	-16.99, 20.84	19.31
Delayed recall	608 87.39 (15.75)	601 84.90 (13.88)	0.45	0.42 (0.35, 0.49)	-2.52 (-3.79, -1.26)	-33.29, 28.25	31.40
Nonverbal index	602 76.01 (12.03)	602 76.63 (12.45)	0.75	0.73 (0.69, 0.76)	0.53 (-0.17, 1.23)	-16.53, 17.59	17.41
Mental processing index	607 77.27 (11.63)	601 76.83 (11.98)	0.73	0.71 (0.67, 0.75)	-0.53 (-1.22, 0.16)	-17.40, 16.33	17.21
Bruininks-Oseretsky Test of Motor Proficiency (BOT-2)							
Total score	608 50.85 (8.42)	603 51.59 (8.06)	0.70	0.69 (0.64, 0.73)	0.80 (0.30, 1.31)	-11.63, 13.24	12.69
Behavior Rating Inventory for Executive Function (BRIEF)							
BRI	609 52.21 (12.28)	602 49.81 (11.40)	0.51	0.52 (0.46, 0.58)	-2.31 (-3.25, -1.37)	-25.29, 20.67	23.45
MI	609 51.38 (12.06)	603 48.16 (11.13)	0.56	0.56 (0.51, 0.61)	-3.17 (-4.05, -2.29)	-24.63, 18.30	21.90
GEC	609 51.92 (12.20)	603 48.77 (11.17)	0.56	0.57 (0.52, 0.62)	-3.07 (-3.95, -2.19)	-24.59, 18.44	21.95
Monitor	609 50.32 (12.91)	603 47.26 (12.47)	0.49	0.50 (0.44, 0.56)	-2.98 (-4.00, -1.95)	-28.02, 22.07	25.56
Test of Variables of Attention (TOVA)							
ADHD	582 0.03 (2.86)	595 0.70 (2.96)	0.52	0.53 (0.47, 0.59)	0.76 (0.53, 0.99)	-4.74, 6.26	5.61
Response time variability (ms)	582 247.14 (71.81)	595 208.89 (66.77)	0.66	0.66 (0.61, 0.70)	-41.39 (-46.01, -36.76)	-151.90, 69.12	112.77
Response time (ms)	582 663.05 (153.03)	595 590.65 (129.70)	0.76	0.74 (0.71, 0.78)	-76.71 (-85.00, -68.42)	-274.71, 121.29	202.04
Commission (%)	582 9.52 (6.39)	595 6.94 (5.45)	0.46	0.46 (0.39, 0.52)	-2.71 (-3.21, -2.20)	-14.77, 9.36	12.31
Omission (%)	582 15.28 (16.81)	595 10.25 (13.11)	0.57	0.57 (0.51, 0.62)	-5.65 (-6.78, -4.52)	-32.70, 21.40	27.60
D-prime	582 2.70 (0.99)	595 3.25 (1.10)	0.63	0.63 (0.58, 0.68)	0.58 (0.51, 0.66)	-1.17, 2.33	1.78
D-prime standard score	582 85.88 (13.09)	595 89.22 (13.61)	0.47	0.45 (0.39, 0.52)	3.71 (2.58, 4.83)	-23.16, 30.58	27.42
Selected KABC Raw Scores							
Conceptual thinking	608 12.07 (5.77)	602 14.39 (5.26)	0.68	0.66 (0.62, 0.71)	2.33 (1.97, 2.68)	-6.40, 11.05	8.90
Story completion	608 5.09 (4.09)	602 6.59 (5.03)	0.82	0.80 (0.77, 0.83)	1.48 (1.25, 1.71)	-4.17, 7.14	5.77
Pattern reasoning	608 5.51 (4.76)	602 7.32 (5.60)	0.75	0.74 (0.70, 0.77)	1.78 (1.49, 2.08)	-5.49, 9.06	7.43
Face recognition	608 11.09 (4.13)	602 12.33 (3.67)	0.58	0.57 (0.51, 0.62)	1.25 (0.96, 1.53)	-5.80, 8.29	7.19
Hand movements	608 6.73 (2.77)	602 7.44 (2.88)	0.64	0.64 (0.59, 0.68)	0.73 (0.54, 0.92)	-3.93, 5.39	4.76

Abbreviations: ADHD, attention deficit/hyperactivity disorder; BRI, Behavioral Regulation Index; CI, confidence interval; CR, coefficient of repeatability; GEC, Global Executive Composite; ICC, intraclass correlation coefficient; MI, Metacognition Index.

^aPearson = Pearson's product-moment correlation.

^bICC = Intraclass correlation coefficient; Unadjusted ICC values incorporate cohort and time in regression models.

^c95% of differences should fall within lower and upper limits of agreement.

^dCR = coefficient of repeatability (roughly half the width from lower to upper limit of agreement).

lower and upper levels of agreement and the coefficient of repeatability is about half the width of this interval. Patterns of absolute and relative measures of test–retest reliability were similar for the analysis of data just from the HIV-infected cohort (data not shown).

Correspondence (Convergent) Validity—Association among Neuropsychological Tests

There were patterns of modest though statistically significant positive correlations between KABC sequential scaled scores and TOVA ADHD and D-prime baseline scores. Correlations ranged between 0.18 and 0.31 for the full sample and were slightly higher for the HIV positive group, ranging up to almost 0.4 for the TOVA D-prime (► **Table 3**).

TOVA response time scores correlated negatively with the BOT-2 total points score. The BRIEF global scales and TOVA error measures were either not significantly correlated or correlations were very low. The other pairwise correlations were low but significant and in the expected directions. Subgroup analyses of HIV-infected and -uninfected children followed similar patterns. These relationships were supported by the adjusted regression parameter estimates, most of which were all significantly different from zero in the expected directions (► **Supplementary Table S3**, online only).

Construct Validity—Association with Predictors of Neuropsychological Functioning

TOVA and KABC scores were markedly lower for the HIV cohort in comparison to HEU and HUU (► **Table 4**). There were strong and significant adjusted associations between individual measures of growth, disability, and development with the baseline KABC global scores (NVI, Mental Processing Index; ► **Table 4**). Although these measures were also associated with the BRIEF GEC score, only the child development and disability scores were consistently associated in the expected directions. Strong and significant linear associations also held between individual indicators of socioeconomic status and KABC global domains but largely not with the BRIEF global scores (► **Supplementary Table S4**, online only). For the KABC nonverbal test, associations with growth measures were largely driven by the HIV-uninfected study group (HEU/HU) (► **Supplementary Table S5**, online only) and this held true also for the associations between both KABC measures and socioeconomic indicators (data not shown). The TOVA D-prime and ADHD scores were associated with child development and disability when all participants were analyzed (► **Table 4**), but in subgroup analyses, only child developmental scores were associated with these outcomes (► **Supplementary Table S5**, online only).

Factor Analyses

Distributional analysis indicated item scores were not normally distributed (all Shapiro–Wilk *p*-values < 0.001). However, except for Gestalt closure, graphical analysis indicated approximate normality with slight kurtosis and skewness (► **Supplementary Fig. S1a** and **b**, online only). When we analyzed the KABC-II baseline scores of all participants together, 47% of the total variance was explained by the four-

factor solution. Factor 1 largely loaded on measures reflecting learning (Atlantis, Atlantis delay, Rebus, Rebus delay), factor 2 loaded on simultaneous processing (Block count, Rover) and planning (Pattern reasoning, Story completion), factor 3 loaded on sequential processing (Number recall, Word order), and factor 4 loaded on simultaneous processing (Triangles and Gestalt closure; ► **Fig. 2**). These results also held for the analysis of HIV-infected participants alone and for the analysis of the uninfected group (HEU and HU combined; ► **Table 5**).

Discussion

Despite demographic differences across six research sites in four countries speaking nine languages (including English), we were able to use a Western-based set of neuropsychological tests successfully in an African context with psychometrically valid and reliable outcomes yielding consistent findings across all study sites. Previously, we have shown that scores on these tests were lower in the HIV-infected cohort³⁸ and these results were confirmed in this study. These results demonstrated acceptable testing consistency across 96 weeks for the KABC, TOVA, BRIEF, and BOT-2.

Other researchers have noted moderate to high test–retest correlations at 2 to 3 weeks in three LMIC countries for KABC learning, sequential processing, and planning subtests and tests equivalent to KABC simultaneous processing and TOVA measures of inattention and error rates.⁵ In that study, testing was repeated at 2 to 3 weeks, whereas in this study we repeated testing at 48 and 96 weeks. Although time intervals for the retesting differed, the test–retest reliability measures were comparable.

The lowest KABC ICCs corresponded to subtests that have differing age bands. For example, among the KABC raw scores, Delayed recall is not usually administered to children if their original scores (Rebus and Atlantis) were low (< 4); but we administered these scores to all participants, which could have affected correlations with their scores when they were 1 year older (at week 48). It is also possible that their school status changed between the two study visits, which could also have influenced their learning and delayed recall abilities. On the other hand, conceptual thinking and face recognition were designed for children 7 years and under; these children may have reached a ceiling 1 year later (at the second study visit).

The test–retest analysis of just the HIV-infected cohort demonstrated substantially similar measures of relative agreement (correlations) as on the full set of study participants and absolute measures largely followed the same patterns, mitigating the concern that change in disease status might influence these results.

A recent study of rural, HIV-unexposed isiZulu-speaking South African children from 7 to 11 years of age validated the factor structure of the KABC-II, focusing on eight subtests, two each from the learning, sequential, simultaneous, and planning scales.⁶ Factor analyses of the KABC scaled scores in our study discriminated the theoretical constructs, though explaining less of the total variation than in two previous African studies (over 50% of total variability), one study in

Table 3 Spearman's correlations for the association between selected baseline neuropsychological test scores

Test outcome	Predictor	Full sample			HIV positive			HIV uninfected		
		N	Correlation	p-Value	N	Correlation	p-Value	N	Correlation	p-Value
KABC sequential	TOVA ADHD	580	0.18	< 0.001	230	0.17	0.01	350	0.11	0.05
	TOVA D-prime	580	0.31	< 0.001	230	0.37	< 0.001	350	0.18	< 0.001
	TOVA D-prime standard score	580	0.26	< 0.001	230	0.35	< 0.001	350	0.13	0.01
BRIEF BRI	TOVA commission (%)	580	0.06	0.18	230	0.10	0.13	350	0.02	0.74
	TOVA omission (%)	580	-0.01	0.74	230	-0.05	0.48	350	-0.01	0.92
BRIEF MI	TOVA commission (%)	580	0.14	< 0.001	230	0.20	0.003	350	0.10	0.05
	TOVA omission (%)	580	0.11	0.01	230	0.10	0.15	350	0.11	0.03
TOVA response time variability (ms)	KABC planning	285	-0.09	0.15	120	0.02	0.82	165	-0.11	0.17
	KABC Mental Processing Index	605	-0.06	0.18	242	-0.07	0.25	363	-0.03	0.54
TOVA response time (ms)	BOT-2: Total score	582	-0.29	< 0.001	231	-0.36	< 0.001	351	-0.21	< 0.001
	BOT-2: Total score	582	-0.24	< 0.001	231	-0.26	< 0.001	351	-0.20	< 0.001
BRIEF monitor	TOVA commission (%)	580	0.14	< 0.001	230	0.22	< 0.001	350	0.08	0.15
	TOVA omission (%)	580	0.11	0.01	230	0.15	0.02	350	0.09	0.10

Abbreviations: ADHD, attention deficit/hyperactivity disorder; BOT-2, Bruininks-Oseretsky Test of Motor Proficiency, 2nd edition; BRI, Behavioral Regulation Index; BRIEF, Behavior Rating Inventory for Executive Function; HIV, human immunodeficiency virus; KABC, Kaufman Assessment Battery for Children; MI, Metacognition Index; TOVA, Test of Variables of Attention.
^aIndicates p-value of < 0.05.

healthy Zairian children, and a second in healthy South African children primarily from Soweto.^{3,39} A third study in Uganda, testing children with cerebral malaria, used the same subtests as in this study, but analyzed raw scores, finding a five-factor solution discriminated among learning, simultaneous processing, sequential processing, and planning.²⁶ In our four-factor analysis of all study participants, we found similar interpretations for the first three factors. Whereas Bangirana et al's fourth factor loaded on planning, in our analysis, planning loaded on the second factor. In all of these cited studies, staff administered the KABC in local languages. Modifications to account for cultural differences varied, including omitting certain test items which participants may not have been familiar with, omitting whole subtests, and considering whether to modify time-based administration rules. These studies also reviewed important factors which could influence test-taking: children's prior exposure to individualized testing, culturally distinct learning styles, differences in perceptual discrimination, and whether children are encouraged to complete tasks accurately rather than quickly.

As a measure of convergent validity, test results for similar domains correlated and regression analyses showed there were statistically significant associations in the expected directions. However, Spearman's correlations were low. Since the regression analyses were adjusted for site, sex, and age, we can only posit the unadjusted correlations might have been affected by measured or unmeasured cultural differences such as those noted above.

Test results were correlated with expected health, growth, development, disability, and, for the KABC, socio-economic indicators. BRIEF scores were associated with child development and disability. Subgroup analyses did not show the same strength of association for the HIV-infected group as for the HIV-uninfected. This could have been due to loss of power resulting from smaller samples and/or from actual differences between the groups.

In this study, tests and task and item instructions were spoken to the children as needed in their local language. Despite the manner in which the standardized administration of the tests could have been undermined by speaking the instructions in the local languages, the measures themselves were sensitive enough in their design, and proved highly adaptable to the various cultural and language contexts in this study. This may have been because of the ability to administer the evaluation with the provision of teaching sample items (KABC-II) for all of the subtests, a short practice session prior to testing (TOVA), and the modeling of motor tasks for all of the items (BOT-2). We acknowledge questions of cross-cultural validity in neuropsychological testing remain difficult ones in light of potential differences in social and learning experiences, understanding of test content, and familiarity with test-taking.^{2,3,11,12}

Our study was not without limitations. At the entry visit, the test battery for caregivers proved too lengthy to be completed in 1 day. Several of our analyses were of HIV-status subgroups, for which smaller sample sizes may have reduced the statistical power to distinguish significant effects. There were several conditions that could have

Table 4 The association between selected baseline neuropsychological test scores and participant growth and development characteristics^a

Predictor	Covariate/Level	KABC global domains			MPI			BRIEF global domains			TOVA selected test domains			
		NVI	p-Value	Estimate (SE)	p-Value	Estimate (SE)	p-value	GEC	Estimate (SE)	p-value	ADHD	Estimate (SE)	p-value	
Cohort ^b	HEU		< 0.001	6.56 (1.08)	< 0.001	6.29 (0.99)	< 0.001	*	-1.43 (1.15)	0.21	1.09 (0.26)	< 0.001	5.58 (1.26)	< 0.001
	HU		< 0.001	7.14 (1.05)	< 0.001	7.92 (1.00)	*	*	-2.35 (1.11)	0.03	0.84 (0.26)	< 0.001	5.31 (1.24)	< 0.001
	HIV	REF			REF		REF		REF		REF		REF	
HIV status	HIV-infected		< 0.001	-6.85 (0.90)	< 0.001	-7.10 (0.83)	*	*	1.89 (0.98)	0.05	-0.96 (0.22)	< 0.001	-5.44 (1.09)	< 0.001
	HIV-uninfected	REF		REF		REF		REF		REF		REF		REF
WHO BMI Z-score	Continuous ^c		0.12	0.77 (0.49)		1.22 (0.47)		*	0.98 (0.46)	0.03	-0.03 (0.12)		-0.70 (0.56)	0.21
WHO height Z-score	Continuous		< 0.001	2.60 (0.44)	< 0.001	2.63 (0.42)	*	*	-1.09 (0.44)	0.01	0.15 (0.11)		1.57 (0.54)	0.003
WHO weight Z-score	Continuous		< 0.001	2.39 (0.47)	< 0.001	2.67 (0.46)	*	*	-0.22 (0.46)	0.64	0.13 (0.12)		0.70 (0.55)	0.20
Biological mom	No		0.19	-2.23 (1.70)		-3.12 (1.47)		*	0.35 (2.41)	0.89	-0.95 (0.46)		-2.00 (1.99)	0.31
	Yes	REF		REF		REF		REF		REF		REF		REF
Caregiver less than 5 y	No		0.89	0.22 (1.53)		0.18 (1.47)		0.90	-1.15 (1.67)	0.49	0.18 (0.34)		-2.23 (1.42)	0.12
	Yes	REF		REF		REF		REF		REF		REF		REF
Child, disability score	Continuous		< 0.001	-0.25 (0.05)	< 0.001	-0.31 (0.04)	*	*	0.41 (0.05)	< 0.001	-0.04 (0.01)		-0.14 (0.06)	0.02
	Continuous		< 0.001	0.19 (0.04)	< 0.001	0.18 (0.04)	*	*	-0.24 (0.04)	< 0.001	0.03 (0.01)		0.19 (0.04)	< 0.001

Abbreviations: ADHD, attention deficit/hyperactivity disorder; BMI, body mass index; BRIEF, Behavior Rating Inventory for Executive Function; GEC, Global Executive Composite; HIV, human immunodeficiency virus; KABC, Kaufman Assessment Battery for Children; MPI, Mental Processing Index; NVI, Nonverbal Index; SE, standard error; TOVA, Test of Variables of Attention; WHO, World Health Organization.
^aRegression models adjusted for site, sex, and age at entry; Each cell of the table represents results from a separate analysis.
^bHIV = HIV-exposed and infected; HEU = HIV-exposed and uninfected; HU = HIV-unexposed and uninfected.
^cContinuous measure, parameter estimates change per additional unit.
*Indicates p-value of < 0.05.

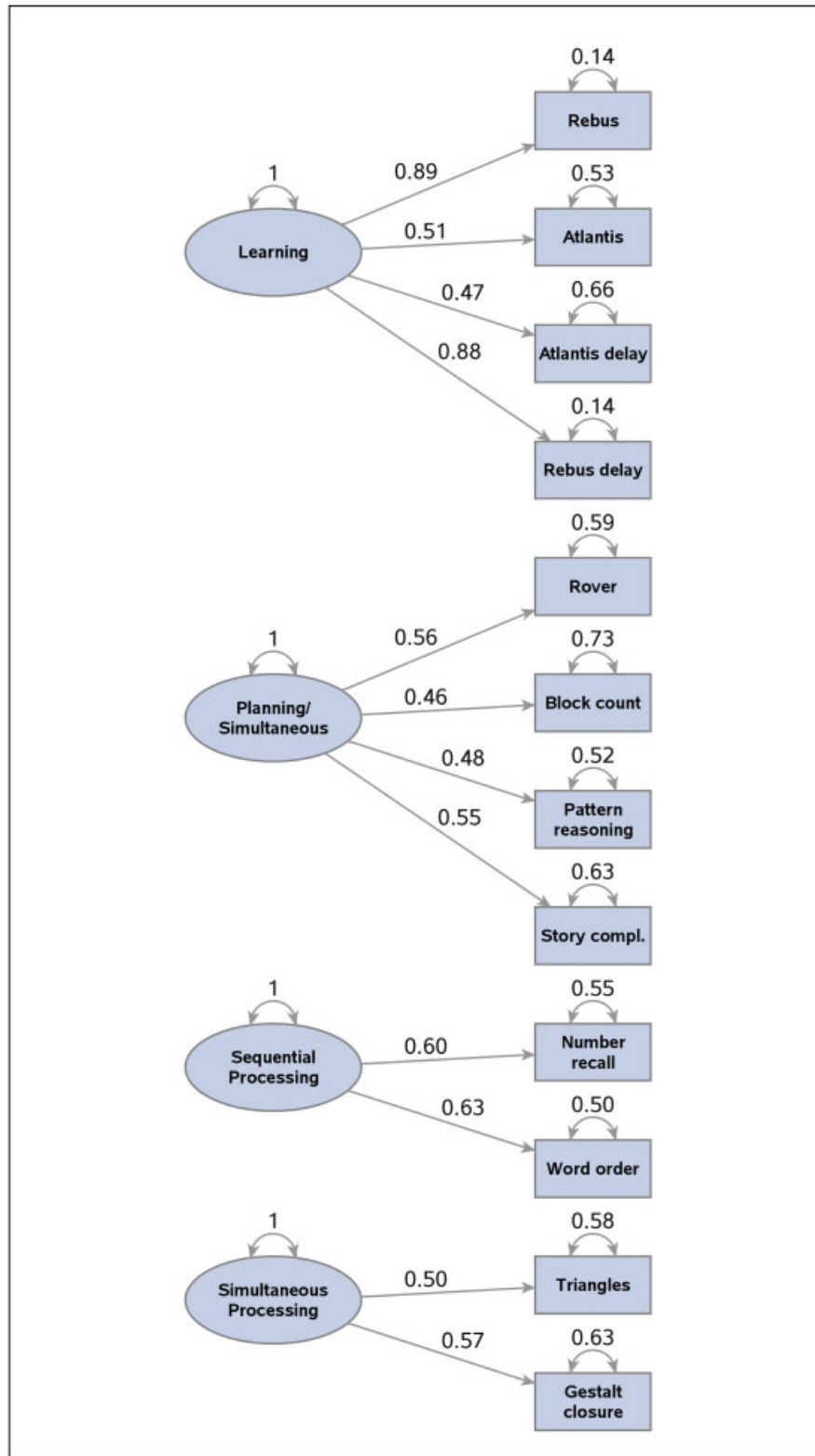


Fig. 2 Factors and factor loadings for the exploratory factor analysis of scaled Kaufman Assessment Battery for Children (KABC) items for all study participants. Note: Factor loadings are only shown if 0.40 or over. The values by the curved arrowed lines indicate factor and error variance estimates. Factor variances are scaled to the value “1.” Factor interpretations are noted in the oval shapes; in order of the amount of total variance explained, they are: Learning, Planning (with Simultaneous), Sequential and Simultaneous processing. Values by the straight arrowed lines indicate factor loadings, which represents the correlation of the specific factor with the indicated KABC item. The square of the factor loading indicates the proportion of variance for that item accounted for by the specific factor. For example, the correlation between the factor representing Learning and Rebus is 0.89. The amount of variance of Rebus accounted for by the Learning factor is $(0.89)^2 = 79\%$. Error variances for the item scores indicate the proportion of total variance that is unique and not explained by the four common factors. For example, for Block count, 73% of the total variance is unique to that variable while 27% is common to the four factors (refer to ► **Table 5** for the communalities).

Table 5 Factor analysis results of baseline KABC-II scores, overall and by cohort (HIV-infected; HIV-uninfected)^a

Data group		Variable	Factors				Communality ^b	
	N		Factor 1	Factor 2	Factor 3	Factor 4		
All cohorts	603	Number recall	0.03	0.03	0.60	0.29	0.45	
		Word order	0.20	0.21	0.63	0.13	0.50	
		Hand movements	0.11	0.26	0.39	0.17	0.26	
		Block count	0.14	0.46	0.16	0.11	0.27	
		Rover	0.29	0.56	0.10	-0.06	0.41	
		Triangles	0.11	0.29	0.28	0.50	0.42	
		Gestalt closure	0.06	0.01	0.20	0.57	0.37	
		Pattern reasoning	0.28	0.48	0.23	0.34	0.48	
		Story completion	0.22	0.55	0.08	0.12	0.37	
		Atlantis	0.51	0.31	0.31	0.14	0.47	
		Rebus	0.89	0.22	0.11	0.13	0.86	
		Atlantis delay	0.47	0.32	0.11	-0.01	0.34	
		Rebus delay	0.88	0.25	0.08	0.13	0.86	
		% common variance		38.78	25.77	20.66	14.79	.
		% total variance		18.06	12.00	9.62	6.89	46.57
		Eigenvalue		2.35	1.56	1.25	0.90	6.05
		HIV-infected	241	Number recall	0.01	0.04	0.61	0.38
Word order	0.22			0.15	0.67	0.12	0.54	
Hand movements	0.14			0.32	0.33	-0.01	0.23	
Block count	0.13			0.53	0.16	0.22	0.37	
Rover	0.35			0.61	0.17	-0.05	0.52	
Triangles	0.09			0.32	0.22	0.49	0.40	
Gestalt closure	-0.00			-0.02	0.09	0.57	0.34	
Pattern reasoning	0.29			0.51	0.13	0.27	0.43	
Story completion	0.19			0.64	-0.00	-0.04	0.45	
Atlantis	0.48			0.24	0.31	0.15	0.41	
Rebus	0.91			0.19	0.11	-0.03	0.88	
Atlantis delay	0.45			0.20	0.11	0.08	0.26	
Rebus delay	0.90			0.25	0.07	-0.01	0.88	
% common variance				38.71	28.03	19.10	14.16	.
% total variance				18.50	13.40	9.13	6.77	47.80
Eigenvalue				2.41	1.74	1.19	0.88	6.21

(Continued)

Table 5 (Continued)

Data group		Variable	Factors				Communality ^b
	N		Factor 1	Factor 2	Factor 3	Factor 4	
HIV-uninfected ^a (combined)	362	Number recall	0.03	-0.04	0.58	0.20	0.38
		Word order	0.17	0.22	0.60	0.12	0.45
		Hand movements	0.06	0.20	0.44	0.20	0.28
		Block count	0.13	0.41	0.10	0.04	0.19
		Rover	0.24	0.53	0.04	-0.04	0.34
		Triangles	0.12	0.25	0.30	0.52	0.44
		Gestalt closure	0.09	-0.04	0.23	0.56	0.38
		Pattern reasoning	0.27	0.44	0.29	0.36	0.48
		Story completion	0.20	0.48	0.10	0.19	0.32
		Atlantis	0.52	0.34	0.24	0.14	0.47
		Rebus	0.87	0.24	0.12	0.18	0.86
		Atlantis delay	0.48	0.40	0.05	-0.05	0.39
		Rebus delay	0.86	0.26	0.08	0.16	0.83
		% common variance	38.41	24.61	21.02	15.96	.
		% total variance	17.12	10.97	9.37	7.11	44.57
		Eigenvalue	2.23	1.43	1.22	0.92	5.79

^aHIV-uninfected combines HIV-exposed and uninfected (HEU) with HIV-unexposed and uninfected (HU).

^bThe last two entries in the Communality column represent the percent of total variance and the total variance accounted for by all four factors. Note: Bold factor loadings meet the cutoff of ≥ 0.40 .

influenced test scores for the same children across time: In some cases (fewer than 5%), the caregiver completing the BRIEF at week 0 differed from that completing it at weeks 48 and 96; there may have been practice effects of repeated tests; HIV-infected children's changes in health or stability could have influenced their test-taking abilities differently at different study visits. In our factor analysis of the KABC scores, we used item scores which were normed to the U.S. population. Since data on the 13 KABC items were not strictly normally distributed, we should treat our results as exploratory findings. The BRIEF instrument may be especially susceptible to bias derived from using nonlocal scoring norms. Further analyses are planned to develop local norms in these four African countries using the P1104S HU cohort data, recognizing these analyses may still yield tentative results due to the small sample sizes of population subgroups categorized by age, sex at birth, and ethnicity/language.

Conclusions and Implications

We found the KABC, TOVA, BRIEF, and BOT-2 to be valid and reliable when using standardized scores from high-income countries in a study of HIV-infected and -uninfected perina-

tally exposed and unexposed children at six research sites in four sub-Saharan African countries. Yet, further research into adapting tests developed in high-income Western settings for LMIC contexts is also warranted.

Boivin and Giordani proposed the concept of a brain/behavior omnibus as a way to conceptually interface foundational neuropsychological functions.²¹ This approach suggests that foundational brain/behavior functions are universal to human neurodevelopment, and therefore, consistent across cultures. However, this approach also recognizes that the neuropsychological assessment of these functions must be sensitive to how such foundational brain/behavior domains are shaped by ecological necessity and cultural experience. The robust validity and reliability of our application of a Western-based neuropsychological assessment battery to evaluate the brain/behavior effects of pediatric HIV across six different study sites representing nine different local languages in four different countries supports this premise.

Neuropsychological measurement can be used to evaluate the extent to which the integrity of the foundational brain/behavior domains is consistently compromised across settings in a given manner for a given disease (for example, HIV).

Likewise, these measures can also be used to evaluate improvement in function, given health interventions.

We need to continue to develop more sensitive and innovative technologies within the cross-cultural neuropsychology of African children. As these technologies become increasingly accessible to resource-constrained settings such as those represented in this study, we will gain a wealth of information that lends interpretive power for a dynamic biocultural co-constructivist paradigm (demonstrating biocultural reciprocity and plasticity) as applied to the neuropsychology of African children. Western-based tests, which have been adapted for use in culturally sensitive, effective, and appropriate manners, will yield valuable results and will, along with dynamic testing, neuroimaging, and genomic analyses, help us characterize the foundational universal brain/behavior domains.

Conflict of Interest

M.J.B. reports grants from National Institutes of Health (NIH), grants from NICHD, and grants from NIMH, during the conduct of the study.

M.C.C reports grants from National Institutes of Health (NIH), grants from NICHD, and grants from NIMH, during the conduct of the study.

C.J. reports grants from IMPAACT Network (NIH), during the conduct of the study.

L.F. reports grants from NIAID, during the conduct of the study.

M.R. reports grants from IMPAACT Network (NIH), during the conduct of the study.

P.K. reports grants from NIH, during the conduct of the study.

T.V. reports grants from National Institutes of Health (NIH), grants from NICHD, and grants from NIMH, during the conduct of the study.

B.Z. reports grants from National Institutes of Health (NIH) during the conduct of the study.

B.L. reports grants from National Institutes of Health (NIH), during the conduct of the study; grants from National Institutes of Health (NIH), outside the submitted work.

Acknowledgments

Overall support for the International Maternal Pediatric Adolescent AIDS Clinical Trials (IMPAACT) Network was provided by the National Institute of Allergy and Infectious Diseases (NIAID) of the National Institutes of Health (NIH) under Award Numbers UM1AI068632 (IMPAACT LOC), UM1AI068616 (IMPAACT SDMC), and UM1AI106716 (IMPAACT LC), with cofunding from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) and the National Institute of Mental Health (NIMH). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. We would like to thank Meredith Warsaw and Stan Morse for carefully reading the manuscript and reviewer comments, Katie McCarthy and J.L. Ariansen for clinical trials support. We would like to thank the

protocol medical officers, Pim Brouwers and Sonia Lee, for their contributions. The following institutions and individuals participated in IMPAACT P1104S: 8051, Wits Reproductive Health and HIV Institute (Wits RHI) Shandukani Research Center CRS, University of the Witwatersrand, Johannesburg, South Africa: Janet Grab, Gurpreet Kindra; 8052, Soweto IMPAACT CRS, Perinatal HIV Research Unit (PHRU), Chris Hani Baragwanath Hospital, Johannesburg, South Africa: Given Leshabane, Hilda Ndiweni; 8950, Family Clinical Research Unit (FAM-CRU) Clinical Research Site (CRS), Tygerberg Hospital, Tygerberg, South Africa: Joan du Plessis, Marisa Groenewald; 30293, Makerere University – Johns Hopkins University (MU-JHU) Research Collaboration (MUHJU CARE LTD) CRS, Kampala, Uganda: Linda Barlow-Mosha, Mary Nyakato; 31890 (30313), Harare Family Care CRS, Parirenyatwa Hospital, Harare, Zimbabwe: Mary N. Tichareva, Petronella Matibe; 12001, University of North Carolina (UNC) Project, Tidziwe Centre, Kamuzu Central Hospital, Lilongwe, Malawi: Mary Chiunda, Tionge Kamvaunamwali.

References

- Boivin MJ, Giordani B. Neuropsychological assessment of African children: evidence for a universal brain/behavior omnibus within a coconstructivist paradigm. *Prog Brain Res* 2009;178:113–135
- Sternberg RJ. Culture and intelligence. *Am Psychol* 2004;59(05):325–338
- Giordani B, Boivin MJ, Opel B, Dia Nseyila DN, Lauer RE. Use of the K-ABC with children in Zaire, Africa: An evaluation of the sequential-simultaneous processing distinction within an intercultural context. *Int J Disabil Dev Educ* 1996;43(01):5–24
- Abubakar A, van de Vijver FJR. How to adapt tests for Sub-Saharan Africa. In: Abubakar A, van de Vijver FJR, eds. *Handbook of Applied Developmental Science in Sub-Saharan Africa*. Springer; 2017:197–212. Available at: <http://www.springer.com/978-1-4939-7326-2>
- Holding P, Anum A, van de Vijver FJR, et al. Can we measure cognitive constructs consistently within and across cultures? Evidence from a test battery in Bangladesh, Ghana, and Tanzania. *Appl Neuropsychol Child* 2018;7(01):1–13
- Mitchell JM, Tomlinson M, Bland RM, Houle B, Stein A, Rochat TJ. Confirmatory factor analysis of the Kaufman assessment battery in a sample of primary school-aged children in rural South Africa. *S Afr J Psychol* 2017;1–19. <https://doi.org/10.1177/0081246317741822>
- Save-the-children. *Cognition Workshop Report*. Paper presented at: Cognition Workshop; July 23–24, 2013; Institute of Education, London; 2013
- Bangirana P, John CC, Idro R, et al. Socioeconomic predictors of cognition in Ugandan children: implications for community interventions. *PLoS One* 2009;4(11):e7898
- Brahmbhatt H, Boivin M, Ssempijja V, et al. Impact of HIV and antiretroviral therapy on neurocognitive outcomes among school-aged children. *J Acquir Immune Defic Syndr* 2017;75(01):1–8
- Ogunlade AO, Kruger HS, Jerling JC, et al. Point-of-use micronutrient fortification: lessons learned in implementing a preschool-based pilot trial in South Africa. *Int J Food Sci Nutr* 2011;62(01):1–16
- Semrud-Clikeman M, Romero RA, Prado EL, Shapiro EG, Bangirana P, John CC. Selecting measures for the neurodevelopmental assessment of children in low- and middle-income countries. *J Child Neuropsychol* 2017;7:761–802
- van Wyhe KS, van de Water T, Boivin MJ, Cotton MF, Thomas KG. Cross-cultural assessment of HIV-associated cognitive impairment using the Kaufman assessment battery for children: a systematic review. *J Int AIDS Soc* 2017;20(01):21412

- 13 Bangirana P, Allebeck P, Boivin MJ, et al. Cognition, behaviour and academic skills after cognitive rehabilitation in Ugandan children surviving severe malaria: a randomised trial. *BMC Neurol* 2011; 11:96
- 14 Bangirana P, Sikorskii A, Giordani B, Nakasujja N, Boivin MJ. Validation of the CogState battery for rapid neurocognitive assessment in Ugandan school age children. *Child Adolesc Psychiatry Ment Health* 2015;9:38
- 15 Boivin MJ, Busman RA, Parikh SM, et al. A pilot study of the neuropsychological benefits of computerized cognitive rehabilitation in Ugandan children with HIV. *Neuropsychology* 2010;24(05):667–673
- 16 Familiar I, Ruisenor-Escudero H, Giordani B, et al. Use of the behavior rating inventory of executive function and child behavior checklist in Ugandan children with HIV or a history of severe malaria. *J Dev Behav Pediatr* 2015;36(04):277–284
- 17 Kaufman AS, Kaufman NL. Kaufman Assessment Battery for Children. 2nd ed. San Antonio, TX: Pearson Inc NCS Pearson; 2004
- 18 The T.O.V.A. (Version 6.X) (Computer Program) [computer program]. Los Alamitos, CA; 1993
- 19 Bruininks RH, Bruininks BD. BOT2: Bruininks-Oseretsky Test of Motor Proficiency. 2nd ed. Minneapolis, MN: Pearson Assessments; 2005
- 20 Gioia GA, Isquith PK, Guy SC, Kenworthy L. Behavior Rating Inventory of Executive Function® (BRIEF®). Lutz, FL: Psychological Assessment Resources (PAR); 2003
- 21 Boivin MJ, Giordani B. Neuropsychological assessment of African children: evidence for a universal basis to cognitive ability. In: Chiao JY, ed. *Cultural Neuroscience: Cultural Influences on Brain Function*. Vol. 178. New York, NY: Elsevier Publications; 2009: 113–135
- 22 Boivin MJ, Giordani B, eds. *Neuropsychology of children in Africa: perspectives on risk and resilience*. In: Fletcher-Janzen E, ed. *Specialty Topics in Pediatric Neuropsychology*. Vol. 1. New York, NY: Springer; 2013
- 23 Boivin MJ, Barlow-Mosha L, Chernoff MC, et al; IMPAACT P1104s Study Team. Neuropsychological performance in African children with HIV enrolled in a multisite antiretroviral clinical trial. *AIDS* 2018;32(02):189–204
- 24 Barlow-Mosha L, Angelidou K, Lindsey J, et al. Nevirapine- versus lopinavir/ritonavir-based antiretroviral therapy in HIV-infected infants and young children: long-term follow-up of the IMPAACT P1060 randomized trial. *Clin Infect Dis* 2016;63(08):1113–1121
- 25 Bangirana P, Giordani B, John CC, Page C, Opoka RO, Boivin MJ. Immediate neuropsychological and behavioral benefits of computerized cognitive rehabilitation in Ugandan pediatric cerebral malaria survivors. *J Dev Behav Pediatr* 2009;30(04):310–318
- 26 Bangirana P, Seggane-Musisi, Allebeck P, et al. A preliminary examination of the construct validity of the KABC-II in Ugandan children with a history of cerebral malaria. *Afr Health Sci* 2009;9(03):186–192
- 27 Bergemann TL, Bangirana P, Boivin MJ, Connett JE, Giordani BJ, John CC. Statistical approaches to assess the effects of disease on neurocognitive function over time. *J Biom Biostat* 2012;S7(016, Suppl 7):1–8
- 28 Boivin MJ, Bangirana P, Byarugaba J, et al. Cognitive impairment after cerebral malaria in children: a prospective study. *Pediatrics* 2007;119(02):e360–e366
- 29 John CC, Bangirana P, Byarugaba J, et al. Cerebral malaria in children is associated with long-term cognitive impairment. *Pediatrics* 2008;122(01):e92–e99
- 30 Deitz JC, Kartin D, Kopp K. Review of the Bruininks-Oseretsky Test of Motor Proficiency, Second Edition (BOT-2). *Phys Occup Ther Pediatr* 2007;27(04):87–102
- 31 Gioia GA, Isquith PK, Retzlaff PD, Espy KA. Confirmatory factor analysis of the Behavior Rating Inventory of Executive Function (BRIEF) in a clinical sample. *Child Neuropsychol* 2002;8(04): 249–257
- 32 UNICEF. Multiple Indicator Cluster Surveys; 2017. Available at: <http://mics.unicef.org/>. Accessed March 14, 2017
- 33 Bonett DG. Sample size requirements for estimating intraclass correlations with desired precision. *Stat Med* 2002;21(09): 1331–1335
- 34 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1(8476):307–310
- 35 Vaz S, Falkmer T, Passmore AE, Parsons R, Andreou P. The case for using the repeatability coefficient when calculating test-retest reliability. *PLoS One* 2013;8(09):e73990
- 36 Boivin MJ, Chounramany C, Giordani B, Xaisida S, Choulamountry L. Validating a cognitive ability testing protocol with Lao children for community development applications. *Neuropsychology* 1996;10(04):588–599
- 37 Holding P, Boivin MJ. The assessment of neuropsychological outcomes in pediatric severe malaria. In: Boivin MJ, Giordani B, eds. *Neuropsychology of Children in Africa: Perspectives on Risk and Resilience*. New York, NY: Springer Science + Business; 2013
- 38 Boivin M, Chernoff M, Laughton B, et al. Neuropsychological outcomes in a two year African-based pediatric observational study. Paper presented at: Conference on Retroviruses and Opportunistic Infections (CROIs) 2017 February 13–16, 2017; Seattle, WA
- 39 Jansen P, Greenop K. Factor analyses of the Kaufman Assessment Battery for Children assessed longitudinally at 5 and 10 years. *S Afr J Psychol* 2008;38(02):335–365

Appendix A: Description of KABC-II global scales and accompanying subtests

Scale	Subtest	Description
Sequential processing	Number recall	Child repeats a series of numbers in the same sequence the examiner said them
	Word order	The child touches a series of silhouettes of common objects in the same order as the examiner said the names of the objects
	Hand movements	The child copies the examiner's precise sequence of taps on the table with the fist, palm, or side of the hand
Simultaneous processing	Face recognition	The child is presented with a photograph of a human face for 5 s. A photograph is then presented with multiple individuals and the child is asked to point to the person with the face previously shown
	Rover	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 quickest path, i.e., the one that takes the fewest moves
	Triangles	For most items, the child assembles several identical foam triangles (blue on one side, yellow on the other) to match a picture of an abstract design. For easier items, the child assembles a set of colorful plastic shapes to match a model constructed by the examiner or shown on the easel
	Gestalt closure	The child mentally "fills the gaps" in partially completed drawings and names or describes the object or action depicted in the drawing. <i>Seven items representing French fries, fire hydrant, windmill, electric cord, map of the United States, egg carton, and ice cream were removed</i>
	Block count	The child counts the exact number of blocks in pictures of stacks of blocks in which one or more blocks is partly or completely hidden from view. As such, Luria considered it a good measure of a child's ability to use covert visual imagery and orientation to correctly count the number of blocks
Planning	Conceptual thinking	The child is shown a series of four stimuli whereby one of them does not belong with the other three. The child identifies the odd item
	Pattern reasoning	The child is shown a series of stimuli that form a logical, linear pattern, but one stimulus is missing. The child completes the pattern by selecting the correct stimulus from an array of 4 to 6 options at the bottom of the page
	Story completion	The child is shown a row of pictures that tell a story, but some of the pictures are missing. The child is given a set of pictures, selects only the ones that are needed to complete the story and places the missing pictures in their correct locations
Learning and delayed recall	Atlantis	The examiner teaches the child nonsense names for fanciful pictures of fish, plants, and shells. The child demonstrates learning by pointing to each picture (out of an array of pictures) when it is named
	Rebus	The examiner teaches the child the word or concept associated with each particular rebus (drawing) and the child then "reads" aloud phrases and sentences composed of these rebuses
	Atlantis delayed	The child demonstrates delayed recall of paired associations learned ~15–25 min earlier during Atlantis by pointing to the picture of the fish, plant, or shell that is named by the examiner
	Rebus delayed	The child demonstrates delayed recall of paired associations learned ~15–25 min earlier during Rebus by "reading" phrases and sentences composed of those same rebuses

Abbreviation: KABC-II, Kaufman Assessment Battery for Children, 2nd edition.