

UC Berkeley

UC Berkeley Previously Published Works

Title

Midlife cognitive testing in Africa: validity of the Harmonised Cognitive Assessment Protocol in the Kenya Life Panel Survey

Permalink

<https://escholarship.org/uc/item/1rg1t12n>

Journal

BMJ Open, 15(8)

ISSN

2044-6055

Authors

Gross, Alden L
Duhon, Madeline
Ochieng, Eric
[et al.](#)

Publication Date

2025-08-01

DOI

10.1136/bmjopen-2024-096619

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial License, available at <https://creativecommons.org/licenses/by-nc/4.0/>

Peer reviewed

BMJ Open Midlife cognitive testing in Africa: validity of the Harmonised Cognitive Assessment Protocol in the Kenya Life Panel Survey

Alden L Gross ¹, Madeline Duhon,² Eric Ochieng,³ Jean N Ikanga,⁴ William H Dow,⁵ Jinkook Lee,⁶ Michael W Walker,⁷ Michelle Layvant,⁸ Anthony Ngugi,⁹ Joshua R Ehrlich ¹⁰, Edward A Miguel ⁷

To cite: Gross AL, Duhon M, Ochieng E, *et al.* Midlife cognitive testing in Africa: validity of the Harmonised Cognitive Assessment Protocol in the Kenya Life Panel Survey. *BMJ Open* 2025;**15**:e096619. doi:10.1136/bmjopen-2024-096619

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<https://doi.org/10.1136/bmjopen-2024-096619>).

Received 15 November 2024
Accepted 15 July 2025



© Author(s) (or their employer(s)) 2025. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ Group.

For numbered affiliations see end of article.

Correspondence to

Dr Edward A Miguel;
emiguel@berkeley.edu

ABSTRACT

Objectives Cohort studies of ageing and cognitive decline typically do not begin fielding comprehensive cognitive assessments until older adulthood. However, for identifying preventable dementia risk factors, there is strong value in beginning at earlier ages. The case is especially compelling in sub-Saharan Africa, where the number of older individuals is expected to triple in the next three decades, and where risk factors may operate more intensively at earlier ages. This study reports on the adaptation and validity of the Harmonised Cognitive Assessment Protocol (HCAP) approach in the Kenya Life Panel Survey (KLPS), collected among middle-aged respondents.

Design To evaluate the validity of the HCAP approach in Kenya, this study assesses model fit statistics from confirmatory factor analyses (CFA) and tests measurement invariance by respondent characteristics.

Setting Both rural and urban areas in Kenya.

Participants A sample of n=5878 individuals from the KLPS, who have been surveyed regularly since they were schoolchildren in the 1990s. The HCAP assessment was administered in 2023 at an average age of 37 years (10–90 range 34 to 41).

Primary and secondary outcome measures For each individual, the CFA generates a general cognitive performance score, and cognitive performance scores for five distinct domains, including memory, executive functioning, language, orientation to time and place, and visuospatial functioning.

Results Fit of the models to the data was adequate for general cognitive performance (root mean squared error of approximation (RMSEA)=0.03; comparative fit index (CFI)=0.94; standardised root mean residual (SRMR)=0.05), language (RMSEA=0.02; CFI=0.95; SRMR=0.05) and good for memory (RMSEA=0.05; CFI=0.99; SRMR=0.02) and executive functioning (RMSEA=0.03; CFI=0.98; SRMR=0.03). The CFA indicate that the factor structure is consistent with findings from other countries and that reliability for the general cognitive performance score was high. Statistical models also suggest invariance at the scalar level for leading demographic (gender, age) and socioeconomic (education, occupational complexity) characteristics.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- ⇒ Confirmatory factor analyses indicate that reliability for the general cognitive performance score was high, and statistical models suggest invariance at the scalar level for leading sociodemographic characteristics.
- ⇒ A limitation of the study population is that it is not nationally representative.
- ⇒ A strength of the Kenya Life Panel Survey (KLPS) data is its long-standing longitudinal (panel) dimension, with baseline data collected in 1998, including childhood cognitive test scores, and plans for continued data collection.
- ⇒ The cognitive functioning of mid-age Kenyans appears to be well captured by the adapted KLPS Harmonised Cognitive Assessment Protocol (HCAP) protocol, providing a valuable baseline for studying future cognitive decline.

Conclusions This study demonstrates that the cognitive functioning of mid-age Kenyans appears to be well captured by the adapted protocol. While there is a moderate decline in cognitive performance among older individuals, this relationship appears to be mediated by education, indicating that this KLPS HCAP provides a valuable baseline for studying future cognitive decline.

INTRODUCTION

Cohort studies of ageing and cognitive decline typically do not begin fielding comprehensive cognitive assessments until older adulthood. For example, the Harmonised Cognitive Assessment Protocol (HCAP) global family of studies starts at age 65 in the USA,¹ England (ELSA)² and other European countries (SHARE)³, at age 60 in India (LASI-DAD)⁴ and China (CHARLS)⁵, at age 55 in Mexico (Mex-Cog)⁶ and at age 50 in South Africa (HAALSI-HCAP).⁷ Given the time demands of such assessments, this focus on older adults may be reasonable for measurement of

dementia prevalence and incidence, especially in populations where dementia remains rare until well after 70 years.

For the complementary goal of identifying preventable dementia risk factors, the increasing focus on earlier life course factors suggests strong value in beginning cohort studies at earlier ages.^{8 9} Assessments earlier in life can improve the characterisation of risk exposures across the life course, better measure rates of change in cognitive performance and facilitate earlier testing of relationships with cognition. The case for younger cohort studies is especially compelling in low- and middle-income country populations, where risk factors may operate more intensively starting at younger ages and because of shorter life expectancies. For example, a study in rural Malawi reported evidence of detectable cognitive decline in individuals aged 45–55 years.¹⁰

Sub-Saharan Africa (SSA) is the world's lowest income region¹¹ and the region that has been most neglected in dementia research to date. A decade ago, the World Alzheimer's Report highlighted that although there were few dementia prevalence estimates among African populations, changing risk factors and population ageing could cause the number of African people living with dementia to double within 20 years.¹² In SSA, the population over age 60 is projected to triple from 2020 to 2050, then quadruple by 2100 to reach over 600 million older adults.¹³ There is an urgent need to significantly expand dementia research in SSA to prepare for this challenge.¹⁴

To advance African dementia research, it will be important to validate the adaptation of harmonised cognitive assessment tools to African contexts. This paper reports on the validity of the HCAP battery in the ongoing Kenya Life Panel Survey (KLPS), which has followed a sample of individuals since they were school-children in the late 1990s. The HCAP cognitive assessment was administered to the panel in 2023 at an average participant age of 37 years. This data environment leverages KLPS's extensive history of risk factor exposure data, and early life measures of cognitive ability and academic performance, to be studied in relation to midlife cognitive performance, as well as to future planned waves of repeated HCAP assessments. To assess the validity of these adapted HCAP measures to this Kenya population at this innovatively young survey age, we assess model fit statistics from confirmatory factor analyses (CFA) as previously applied to HCAP data from other countries.¹⁵ We then test measurement invariance by demographic and occupational characteristics and present criterion validation analyses of cognitive performance in relation to demographic characteristics.

METHODS

Participants

KLPS is a cohort of initially over 7000 students who attended primary school in Busia, Kenya, between 1998 and 2002, participated in one of two randomised

interventions and were then recruited into the cohort study.¹⁶ The first intervention (accounting for over 90% of the present sample) was a Primary School Deworming Project, which included children enrolled in participating schools in 1998.¹⁷ A representative subsample of children in these schools was recruited into KLPS. The second was a Girls Scholarship Programme, which included sixth grade girls enrolled in a distinct set of primary schools in 2001–2002.¹⁸

The present study analyses data from a sample of 5878 respondents who participated in the fifth KLPS survey round in 2023, which included a module of HCAP cognitive tests adapted for a Kenyan Swahili-speaking population. The approach builds on both the original Health and Retirement Survey (HRS) HCAP¹ and the India LASI-DAD.⁴ Swahili is one of the two official national languages in Kenya and is widely spoken in all regions of the country, and nearly all KLPS respondents are proficient in the language. KLPS employs a two-stage tracking methodology to maintain high effective tracking rates.¹⁶ The effective survey rate (from baseline) for the KLPS HCAP module stands at 83.1% among the non-deceased (4.8% of the baseline sample is deceased). Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Variables

Participants were administered the HCAP, a neuropsychological assessment battery that comprises tests and items in common with other HRS international partner studies. Tests included items related to orientation to place and time; measures of memory including the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) 10-Word List Learning,¹⁹ East Boston Memory Test (EBMT)²⁰ and Logical Memory test²¹; measures of executive functioning/attention including Raven's Progressive Matrices,²² Similarities and Differences, Token Test, Go-No-Go, Trail Making and Symbol Cancellation; measures of language/fluency including animal fluency and various tests of object naming and measures of visuospatial functioning including CERAD Constructional Praxis, interlocking pentagons and clock drawing.¹⁵

Adaptations of tests were necessary both for proper functioning in Swahili and to ensure appropriate familiarity in the Kenyan context.²³ The process followed standard guidelines for cross-cultural neuropsychological test development, including forward translation by bilingual experts and independent back-translation.²⁴ To improve contextual relevance, certain test items were adapted to replace culturally unfamiliar references (eg, substituting 'roses and daisies' with 'kale and spinach'). Specific phrases were also adapted to preserve the intended level of difficulty, such as replacing the English tongue twister 'No ifs, ands, or buts' with the Swahili phrase 'Tupe tupate tumpatie Taaka'. All measures were pretested and refined iteratively with input from Kenyan enumerators and cognitive specialists. This ensured conceptual and functional equivalence of the measures with those used in

Table 1 Descriptive characteristics of the Kenya Life Panel Survey (KLPS) sample (n=5878)

Variable	Mean (SD) or N (%)	Range	Number (%) of participants with score at ceiling
Panel A: Sample characteristics			
Age in years at baseline, mean (SD)	36.9 (2.7)	31, 47	
Female sex, N (%)	3275 (55.7)		
Urban, N (%)	2511 (43.4)		
Urban, Nairobi, N (%)	1199 (20.7)		
Urban, Mombasa, N (%)	313 (5.4)		
Years of education, mean (SD)	9.4 (2.9)	2, 16	
Secondary education or higher, N (%)	2847 (49.5)		
Occupation: agriculture or fishing, N (%)	1847 (31.4)		
Occupation: retail or commercial occupations, N (%)	1026 (17.5)		
Occupation: unskilled trade, N (%)	1143 (19.4)		
Occupation: skilled and semiskilled trade, N (%)	946 (16.1)		
Occupation: professionals, N (%)	633 (10.8)		
Panel B: Cognitive test performance			
B1. Memory tests			
Three word immediate registration, N correct (%)	5740 (97.7)		
Three word delayed recall, mean (SD)	2.5 (0.8)	0, 3	3781 (64.3)
CERAD immediate sum of 3 trials, mean (SD)	16.5 (3.3)	0, 30	3 (0.1)
CERAD word list delay, mean (SD)	4.4 (2.0)	0, 9	82 (1.4)
CERAD recognition, mean (SD)	18.0 (1.9)	0, 20	1384 (23.5)
Logical Memory immediate, mean (SD)	17.0 (8.2)	0, 49	1 (0.0)
Logical Memory delay, mean (SD)	12.3 (8.0)	0, 43	1 (0.0)
Logical Memory recognition, mean (SD)	11.0 (1.9)	6, 15	79 (1.3)
Brave man immediate (East Boston Memory Test), mean (SD)	10.0 (3.5)	0, 20	2 (0.0)
Brave man delay (East Boston Memory Test), mean (SD)	7.2 (4.4)	0, 20	1 (0.0)
CERAD Constructional Praxis delay, mean (SD)	6.6 (2.7)	0, 11	571 (9.9)
B2. Executive functioning tests			
Raven's Progressive Matrices, mean (SD)	10.2 (2.5)	0, 16	72 (1.3)
Similarities and differences, mean (SD)	3.6 (0.7)	0, 4	3772 (64.2)
Token Test, mean (SD)	6.0 (1.1)	0, 7	1957 (34.1)
Go-No-Go Part 1, mean (SD)	9.6 (1.5)	0, 10	4995 (85.0)
Go-No-Go Part 2, mean (SD)	7.9 (2.8)	0, 10	2448 (42.3)
Digit Span Forward, N correct (%)	2461 (41.9)		
Digit Span Backward, N correct (%)	3047 (51.8)		
Trail Making, alternating shapes, N correct (%)	936 (16.2)		
Reasoning—lost child on the road, N correct (%)	4934 (84.0)		
Backward day naming, mean (SD)	4.5 (1.4)	0, 5	5114 (87.0)
Symbol Cancellation test, mean (SD)	31.5 (10.8)	0, 61	3 (0.1)
Serial 7s, mean (SD)	3.1 (1.6)	0, 5	1450 (24.8)
Mental calculation, N correct (%)	5461 (92.9)		
B3. Language/fluency tests			
Animal fluency, mean (SD)	12.7 (3.8)	0, 49	1 (0.0)
Name scissors, N correct (%)	5474 (93.1)		

Continued

Table 1 Continued

Variable	Mean (SD) or N (%)	Range	Number (%) of participants with score at ceiling
Name watermelon, N correct (%)	3274 (55.7)		
Name elbow, N correct (%)	4023 (69.5)		
Write a sentence, N correct (%)	5227 (89.1)		
Read and follow command, N correct (%)	5448 (94.7)		
Repetition of phrase, N correct (%)	3419 (58.2)		
What do you do with a hammer, N correct (%)	5838 (99.3)		
Following instructions (point to window, door), N correct (%)	5196 (93.0)		
Where is the local market? N correct (%)	4817 (83.2)		
Three step command, mean (SD)	2.7 (0.6)	0, 3	4242 (73.7)
Object naming, N correct (%)	5044 (187.1)		
B4. Orientation tests, N correct (%)			
Day of month	4718 (80.9)		
Month	5720 (97.3)		
Day of the week	5520 (94.7)		
Year	5509 (93.7)		
What county are we in	5499 (95.0)		
What city are we in	5585 (96.5)		
Season of year	5149 (88.3)		
Floor of building	5320 (91.9)		
How did you get here from the nearest town?	5648 (97.5)		
Name the president	5776 (98.4)		
B5. Visuospatial functioning tests			
CERAD Constructional Praxis, mean (SD)	9.3 (2.0)	0, 11	2044 (34.8)
Copy interlocking pentagons, N correct (%)	4244 (73.3)		
Clock drawing, mean (SD)	2.5 (0.7)	0, 3	3737 (65.0)
Means and SD are presented (unweighted). CERAD, Consortium to Establish a Registry for Alzheimer's Disease.			

other HCAP implementations. Additional details of these tests and their adaptations are in supplemental materials (see online supplemental information).

Analysis plan

We described the sample using means and SD for continuous characteristics and counts and proportions for binary variables. For continuous cognitive tests, we calculated the proportion of responses at the ceiling, indicating perfect test performance. We then performed CFA to organise cognitive test items into summary domains of general cognition, memory, executive functioning, language, orientation to time and place and visuospatial functioning. This analysis included evaluation of marginal reliability and measurement invariance. Next, we explored criterion-related validity of the resulting factors by correlating them with demographic characteristics including age, gender and education.

Confirmatory factor analyses

We estimated a series of CFA models to represent general and domain-specific summary factors in the KLPS sample. We evaluated the fit of models using the root mean squared error of approximation (RMSEA), the comparative fit index (CFI) and the standardised root mean residual (SRMR). These global fit statistics are preferred over χ^2 fit statistics because they are not affected by model complexity or sample size. RMSEA values of ≤ 0.05 , CFI values of ≥ 0.95 and SRMR values ≤ 0.08 suggest excellent model fit.²⁵ To improve model fit as needed, we incorporated bifactor structures to account for correlations between test items not explained by a general factor.²⁶ To assess the marginal reliability of resulting factor analysis models for each domain, we calculated the SE of measurement from the CFA models.²⁷

Measurement invariance

After constructing factors to represent general and domain-specific cognitive performance, we tested

Table 2 Model fit statistics from confirmatory factor analysis models of general and domain-specific cognitive performance: Results from Kenya Life Panel Survey (KLPS) (n=5878)

Cognitive domain	RMSEA	CFI	SRMR	Interpretation
General cognitive performance	0.030	0.941	0.053	Adequate
Memory	0.047	0.992	0.022	Good
Executive function/attention	0.028	0.977	0.025	Good
Language/fluency	0.018	0.951	0.051	Adequate
Orientation	0.025	0.811	0.083	Poor
Visuospatial	0.000	1.000	0.000	Perfect

This table shows global absolute model fit statistics for confirmatory factor analyses of general and domain-specific cognitive performance in the KLPS sample.

Preferred fit is indicated by $RMSEA < 0.05$, $CFI > 0.95$ and $SRMR \leq 0.08$. To summarise overall model fit based on these statistics, we interpreted model fit to be perfect if $CFI = 1$ and $RMSEA = 0$ and $SRMR = 0$; good if $CFI \geq 0.95$ and $RMSEA \leq 0.05$ and $SRMR \leq 0.05$; adequate if $CFI \geq 0.90$ and $RMSEA \leq 0.08$ and $SRMR \leq 0.08$ and poor if either $CFI < 0.9$ or $RMSEA > 0.08$ or $SRMR > 0.08$.¹⁵

CFI, comparative fit index; RMSEA, root mean squared error of approximation; SRMR, standardised root mean residual.

whether cognitive factors have similar measurement properties across gender, education and occupational characteristics. Specifically, we evaluated measurement invariance across these factors. We tested configural, metric and scalar levels of invariance using a multiple-groups structural equations modelling framework of the general cognitive performance factor.²⁸ Configural invariance refers to a multiple-groups model in which all parameters are free to vary between groups. Metric invariance refers to a more constrained model in which factor loadings are fixed to be equal across group; lack of metric invariance implies group differences in random error.²⁹ Scalar invariance refers to a more constrained model in which item thresholds are fixed equal across

group; lack of scalar invariance implies group differences in systematic error.²⁹ The groups were defined by gender (female/male), education (at least some secondary education vs lower), occupational analytical complexity index (median split), occupational interpersonal index (median split) and occupational routine/manual index (median split).^{30 31} To evaluate whether a null hypothesis of invariance should not be rejected, we used a change in $CFI \leq 0.01$ as the general criterion.³²

Criterion validation

We investigated the correlation between the general cognitive performance score and demographic characteristics. We calculated Pearson correlations and plotted

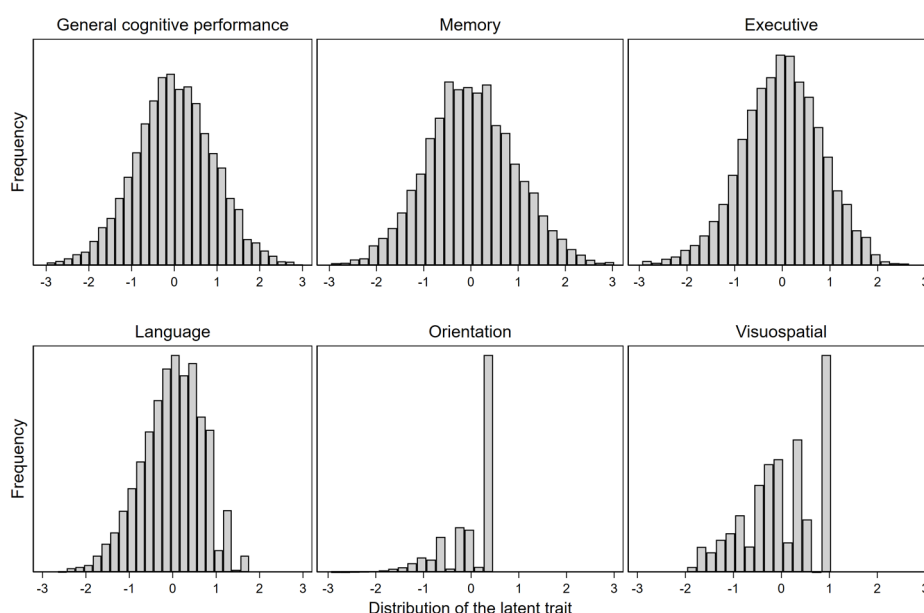


Figure 1 Distribution of general and domain-specific factor scores: Results from Kenya Life Panel Survey (KLPS) (n=5878). Histograms of factor scores, derived from confirmatory factor analyses, for general cognitive performance (top left), memory (top middle), executive functioning/attention (top right), language (bottom left), orientation (bottom middle) and visuospatial ability (bottom right). The scale of the y-axis differs across panels to show the general shape and distribution of each score.

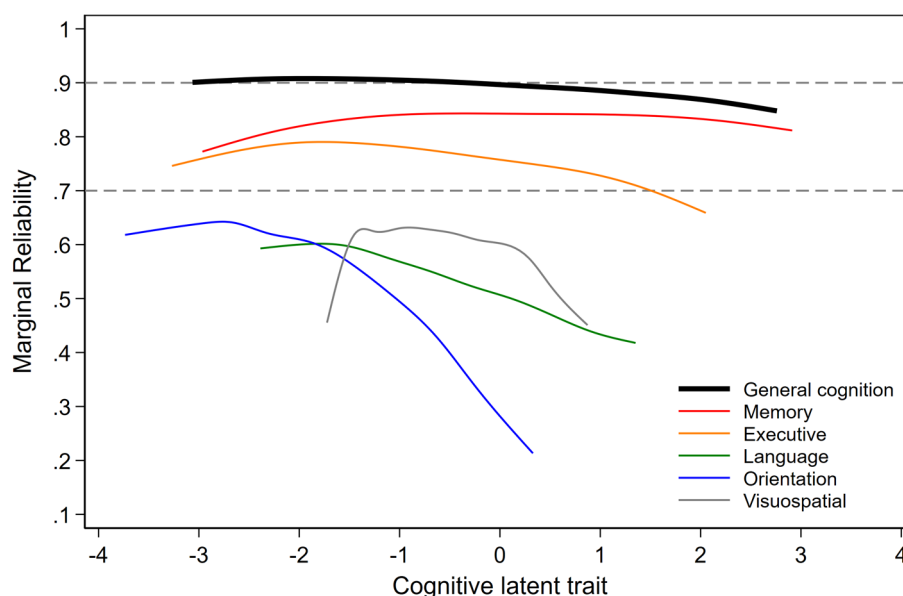


Figure 2 Model-estimated marginal reliability for general and domain-specific cognitive performance: Results from Kenya Life Panel Survey (KLPS) (n=5878). These plots show the marginal reliability of general and domain-specific measurement models. This figure illustrates differences in the reliability of estimated factor scores over levels of the latent trait.

means of general and domain-specific factor scores for cognitive performance against age and years of education. These analyses were then further stratified by gender and education to examine potential differences.

RESULTS

The KLPS Round 5 respondents consisted of n=5878 adults aged 31 to 47 years (mean 36.9 years). Of these participants, 56% are female and 43% live in urban areas, primarily in Nairobi (21%) or Mombasa (5%) (table 1). The sample has 2 to 16 years of schooling (mean 9.4 years), and just under half completed some secondary education, which is similar to the national Kenyan average for these birth cohorts.³³ Regarding occupational information, 31% of the sample reports agriculture or fishing as their primary occupation, with the remainder employed in retail or commercial occupations (18%), unskilled trade (19%), skilled and semiskilled trade (16%) or as professionals (11%). Means of raw cognitive test scores are shown in table 1, alongside counts and percentages of participants performing at the ceiling for each test item.

Confirmatory factor analyses

Model fit of cognitive data to unidimensional CFA models was perfect for visuospatial function (as expected given the domain has only three items and is thus a saturated model), good for memory and executive functioning/attention, adequate for language/fluency and general cognitive performance and poor for orientation (table 2). Consistent with previous HCAP studies from other countries,^{34 35} the memory domain required bifactor structures between immediate, delayed and recognition subtests of the CERAD word list recall, EBMT story memory

and Logical Memory tests. The fit of a unidimensional measurement model to the orientation domain was poor according to CFI and SRMR, largely due to pronounced ceiling effects and thus little variance on every orientation item, making it difficult to fit a parametric model (table 1). The suboptimal fit reflects a known limitation in these types of cognitive test items and has been observed elsewhere.¹⁵

Standardised factor loadings of each item on each domain, as well as the general cognitive factor, are in online supplemental table S1. For memory, loadings were in an acceptable range of 0.3 to 0.9 for all items except for those with pronounced ceiling effects, namely, three word recall (immediate and delayed). For the language/fluency domain, several binary items showed low standardised factor loadings, reflecting low variability due to ceiling effects, including 'Name scissors' (0.17), 'What do you do with a hammer' (0.20), 'Where is the local market' (0.20) and 'Object Naming' (0.24). Similarly, for orientation, standardised factor loadings were almost all uniformly poor, consistent with the poor fit statistics for this domain. Loadings for the visuospatial functioning domain were within an acceptable range.

Distributions of factor scores for general cognitive performance, memory, executive functioning/attention and language are relatively normally distributed, while distributions for orientation and visuospatial function show strong ceiling effects (figure 1). These distributions can be explained by examining the relative thresholds of items, estimated from the factor analysis models. Online supplemental figures S1–S5 plot item threshold parameters along the latent trait for each cognitive test indicator in each domain. The broad coverage of item thresholds

Table 3 Tests of measurement invariance by demographic and work characteristics: Results from Kenya Life Panel Survey (KLPS) (n=5878)

Characteristic	Level	RMSEA	CFI	SRMR
Gender				
	Configural invariance	0.031	0.929	0.062
	Metric invariance	0.028	0.944	0.063
	Scalar invariance	0.030	0.928	0.063
Secondary education or higher versus lower				
	Configural invariance	0.032	0.910	0.064
	Metric invariance	0.029	0.925	0.065
	Scalar invariance	0.030	0.914	0.065
Analytical index, median split				
	Configural invariance	0.029	0.937	0.062
	Metric invariance	0.026	0.947	0.065
	Scalar invariance	0.026	0.944	0.065
Interpersonal index, median split				
	Configural invariance	0.029	0.938	0.061
	Metric invariance	0.026	0.948	0.065
	Scalar invariance	0.026	0.943	0.065
Routine/manual index, median split				
	Configural invariance	0.029	0.939	0.061
	Metric invariance	0.025	0.952	0.062
	Scalar invariance	0.025	0.951	0.062

This table shows results of measurement invariance testing to evaluate whether the factor structure of the general cognitive performance domain is similar across gender, education and occupational characteristics, as described in the Methods section. CFI, comparative fit index; RMSEA, root mean squared error of approximation; SRMR, standardised root mean residual.

in the memory and executive functioning domains over a broad range of the latent trait results in comparatively normal distributions in [figure 1](#), while sparseness of item thresholds for orientation and visuospatial functioning at higher levels of ability is a driver of ceiling effects due to poor marginal reliabilities for these domains in [figure 1](#).

[Figure 2](#) shows the marginal reliability of estimated measurement models for cognitive domains over the range of each latent trait. Consistent with the distribution of item thresholds in online supplemental figure S1, reliability of the measurement model for memory is very high ($r > 0.8$) across nearly the entire range of observed scores, while mean reliability of the executive functioning/attention score falls below $r = 0.7$ above scores of 1.5 SD, beyond which encompasses 3.2% ($n = 187$) of the sample. Similarly, the reliability of the measurement model for general cognitive performance is high across the range

of observed scores, given that items for the general factor are from all domains. Reliabilities for language/fluency, orientation and visuospatial function are uniformly low ($r < 0.6$), corresponding to the fact that these domains were mostly comprised of easier binary items (in the case of language and orientation) or had few items (in the case of visuospatial function).

Measurement invariance

Multiple-group CFA models, with groups for gender, education and occupational characteristics (analytical, interpersonal and routine/manual), were estimated to evaluate configural, metric and scalar levels of measurement invariance ([table 3](#)). The models suggested invariance at the scalar level for each of these characteristics.

Criterion validation

General cognitive performance is negatively correlated with older age ($r = -0.20$) even in this young midlife sample with a relatively restricted age range of 31 to 47 years ([figure 3A](#)). Additionally, mean general cognitive performance is higher on average with more years of schooling ($r = 0.54$; [figure 3B](#)). The relationship between general cognitive performance and age was observed in both females and males ([figure 3C](#)), and similar findings hold for specific cognitive domains (online supplemental figure S6).

To further explore the negative correlation of general cognitive performance with age, we stratified this relationship by education level, across individuals with some secondary education versus those without. Educational attainment rises substantially across these cohorts of Kenyan individuals for both females and males ([figure 3D](#)), leading to a potential confounding of age and education when interpreting the raw correlation of age and cognition in [figure 3A](#). Conditional on secondary education ([figure 3E](#)), the age-cognition pattern flattens ($r = -0.089$ for those with secondary education, $r = -0.088$ for those without); excluding those aged ≥ 41 years, the correlations decline to $r = -0.047$ and $r = -0.043$, respectively. Results for specific cognitive domains are similar (online supplemental figure S7). This result implies that in KLPS Round 5, we have assessed cognitive performance early enough in midlife (with 90% of respondents between ages 34 and 41) that adult aging-related cognitive decline is still minimal and thus establishing a useful baseline sample for prospectively studying adult cognitive decline.

DISCUSSION

Research on cognitive ageing in SSA populations remains relatively sparse compared with other world regions but is of great scholarly and policy interest given the rapid growth of the older population. In this study, we evaluated the factor structure of a relatively comprehensive neuropsychological test battery in a sample of 5878 middle-aged adults in Kenya. The HCAP cognitive battery was

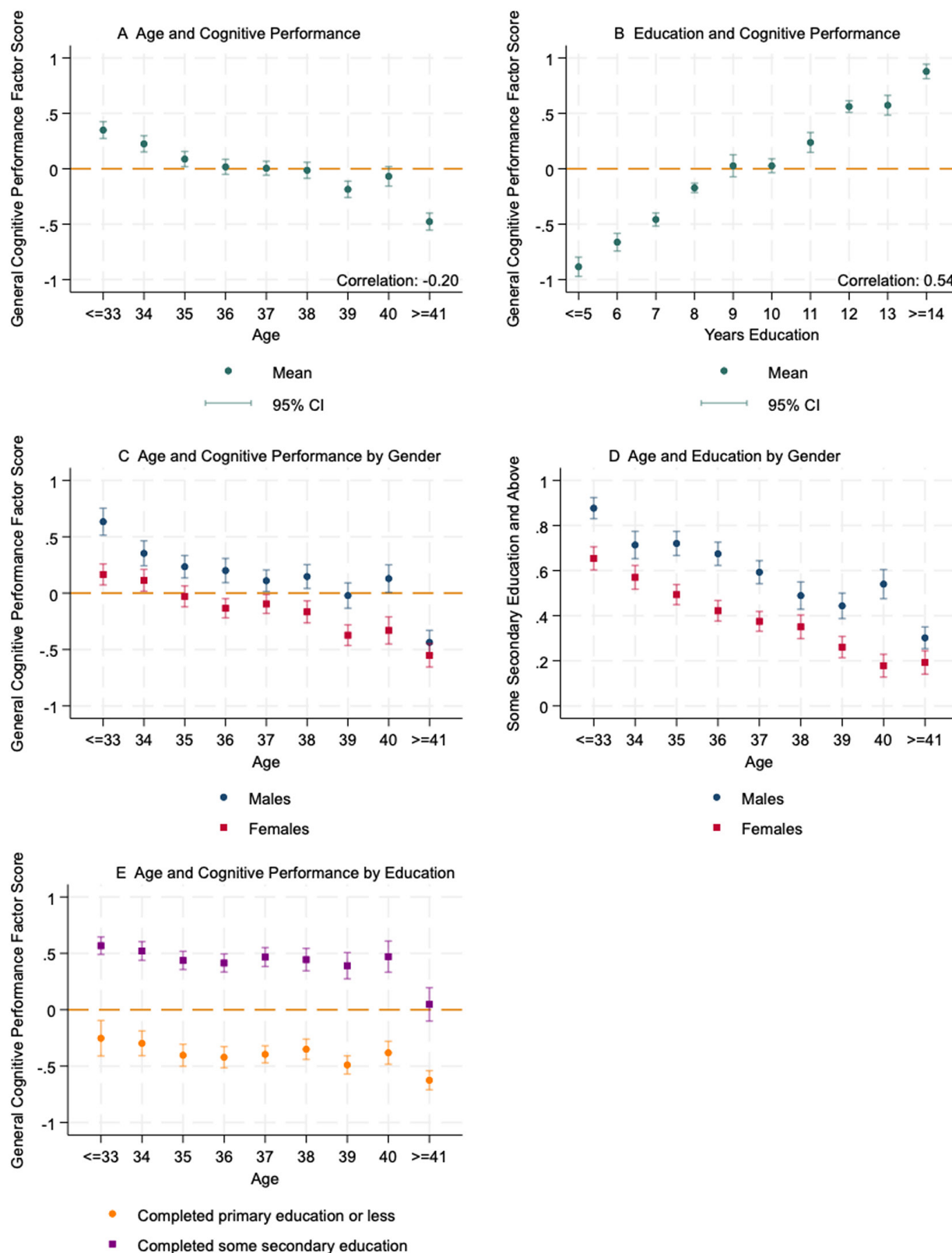


Figure 3 Mean general cognitive performance by age, education and gender: Results from Kenya Life Panel Survey (KLPS) (n=5878). The figure presents mean values of general cognitive performance factor scores (A, B, C, E) or completion of some secondary education (D) within age (A, C, D, E) or years of education completed (B), computed using weights to account for the sampling and survey design. (C, D) Means separately by gender. (E) Means separately by completion of any secondary education. Reported ages of 33 or less are grouped, and reported ages of 41 or greater are grouped (each containing approximately 10% of the sample). Levels of reported educational attainment up to and including 5 years are grouped (containing approximately 7.5% of the sample), and those of 14 years or greater are grouped (containing approximately 10% of the sample).

carefully adapted to the Kenyan context and the widely used Swahili language. Although factors for language, orientation and visuospatial function showed poorer reliability relative to other domains, the factor structure of the KLPS HCAP battery is consistent with findings from other countries outside Africa. Given the large sample size and young age of participants, we expect the scores to be useful in future longitudinal work as this cohort ages.

Reliability for the general cognitive score and for memory was high, while reliability was somewhat lower for some domains, including orientation, due to ceiling effects in cognitive test items. The high levels of performance on simple orientation questions are unsurprising given the relatively young age of study participants who are community-living and not demented.^{7 23} Test items with such ceiling effects may still be useful for dementia screening, as well as longitudinal tracking as the sample ages. Statistical models suggested invariance at the scalar level for leading demographic and socioeconomic dimensions, indicating that individual differences in cognitive performance can be measured consistently across groups defined by these characteristics.

There is a meaningful positive correlation between education and general cognitive performance and modest differences by gender. While cognition is somewhat lower with older age in KLPS, this relationship appears to be mediated by education: conditional on secondary schooling attainment (which is higher for younger cohorts), the relationship between age and general cognitive performance among sample individuals appears to be minimal. Future research regarding education's effect on cognition is needed to disentangle cohort effects from ageing effects. While the full HCAP battery was administered only once in KLPS, tests within the HCAP, namely, animal fluency and Raven's Progressive Matrices, have been administered repeatedly since the cohort was school-aged and efforts are currently underway to model longitudinal within-person changes in cognition.

We note several limitations of this study. First, observed ceiling effects in visuospatial functioning and orientation limit the sensitivity of scores for these domains to detect subtle variations in cognition. This is a well-documented issue in cognitive testing, particularly when applied to younger or midlife populations, and is not unique to the Swahili translation or Kenyan context.^{15 36} Second, the study population is not nationally representative. While the sample was initially drawn from rural western Kenyan schools, these individuals now reside in both urban and rural areas throughout Kenya (and some in neighbouring Uganda); for instance, 43% of the KLPS Round 5 sample lived in urban areas, which is somewhat higher than the national average.³⁷ A further limitation is the cross-sectional nature of the current analysis, which may lead to potential confounding between age and the cultural adaptation of the HCAP protocol. The adapted instrument was applied to a younger Kenyan population, making it difficult to clearly distinguish between

age-related differences and the impact the question adaptations had on cognitive performance. KLPS investigators are also working closely with the team developing the Longitudinal Study of Health and Ageing in Kenya (LOSHAK), which is proposing to similarly field HCAP measures harmonised with the KLPS HCAP in a new nationally representative cohort of older Kenya adults.²³ Results of the LOSHAK study alongside future HCAP survey rounds of KLPS will help address these limitations and clarify the relative contributions of age, cohort and cultural adaptation on cognition.

A strength of the KLPS data is its long-standing longitudinal (panel) dimension, with baseline data collected in 1998, including childhood cognitive test scores, and plans for continued data collection. Another is the ability to estimate long-run causal impacts on cognition of the childhood school health (deworming) intervention from which the sample is derived. This round of data collection serves as a midlife baseline of cognition that can be contrasted with performance in future rounds as the sample enters old age. It is well documented that higher education is associated with better cognitive outcomes among older adults.^{38–41} Yet existing research has found that controlling for adolescent cognitive ability greatly attenuates the association between education and dementia.⁴² The planned KLPS data collection, combined with the midlife baseline presented here together with earlier life cognitive and academic measures in the data, may facilitate research to distinguish between midlife differences in cognition across education levels—as documented in this study—versus a protective effect of education in slowing cognitive decline as individuals age. Ongoing efforts to calibrate cognitive performance scores across survey rounds will enable analysis of within-individual trajectories, advancing understanding of cognitive ageing across the life course in low-resource settings.

A meaningful future direction of the HCAP network, beyond the scope of the present study, will be to conduct comparative analyses of psychometric properties of the HCAP battery across international populations. To date, in addition to KLPS in Kenya, the HCAP battery has been administered in the USA, Mexico, Chile, England, SHARE countries, South Africa, China and India.^{15 43} Efforts have been made to statistically co-calibrate cognitive function cross-nationally, but as the network of studies expands, so do efforts to monitor performance of items and cross-national differences.¹⁵

Author affiliations

¹Department of Epidemiology, Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland, USA

²Department of Economics, Pepperdine University, Malibu, California, USA

³Department of Economics, Maseno University, Maseno, Kenya

⁴Emory University, Atlanta, Georgia, USA

⁵School of Public Health, University of California, Berkeley, California, USA

⁶University of Southern California Dana and David Dornsife College of Letters Arts and Sciences, Los Angeles, California, USA

⁷Department of Economics, University of California, Berkeley, California, USA

⁸Agricultural and Resource Economics, University of California, Berkeley, California, USA

⁹Aga Khan University, Nairobi, Kenya

¹⁰Department of Ophthalmology and Visual Sciences, University of Michigan, Ann Arbor, Michigan, USA

X Alden L Gross @alden.gross

Contributors ALG: conceptualisation, methodology, analysis, interpretation of data and writing. MD, JNl and ML: analysis, interpretation of data and writing. EO: led fieldwork, interpretation of data and writing. WHD, JL, AN and JRE: methodology, interpretation of data and writing. MWW: interpretation of data and writing. EAM: conceptualisation, methodology, interpretation of data and writing. EAM is the guarantor.

Funding This work was supported by the U.S. National Institutes of Health National Institute on Aging (R01AG077001, R21AG077042).

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants. Ethics approval was obtained from the Institutional Review Board (IRB) at the University of California, Berkeley (CPHS Protocol #: 2022-08-15578), the Kenya Medical Research Institute Scientific and Ethics Review Unit (Protocol #: 4619), Maseno University (Approval #: MUERC/00069/14) and the Uganda National Council for Science and Technology. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open-access repository. Data from the KLPS study is posted on the Harvard Dataverse at <https://dataverse.harvard.edu/dataset/KLPS>.⁴⁴ Custom code was used for the statistical analyses in this manuscript.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Alden L Gross <http://orcid.org/0000-0002-4678-0562>

Joshua R Ehrlich <http://orcid.org/0000-0002-0607-3564>

Edward A Miguel <http://orcid.org/0000-0001-6927-709X>

REFERENCES

- Langa KM, Ryan LH, McCammon RJ, *et al*. The Health and Retirement Study Harmonized Cognitive Assessment Protocol Project: Study Design and Methods. *Neuroepidemiology* 2020;54:64–74.
- Cadar D, Abell J, Matthews FE, *et al*. Cohort Profile Update: The Harmonised Cognitive Assessment Protocol Sub-study of the English Longitudinal Study of Ageing (ELSA-HCAP). *Int J Epidemiol* 2021;50:725–726.
- Börsch-Supan A, Douhou S, Fernández I, *et al*. Release note to preliminary SHARE-HCAP classification of mild and severe cognitive impairment in five European countries, Available: <https://www.meashare.eu/wp-content/uploads/2024/05/04-2024.pdf>
- Lee J, Dey AB. Introduction to LASI-DAD: The Longitudinal Aging Study in India-Diagnostic Assessment of Dementia. *J Am Geriatr Soc* 2020;68:S3–4.
- Chen H, Huang Y, Lv X, *et al*. Prevalence of dementia and the attributable contributions of modifiable risk factors in China. *Gen Psych* 2023;36:e101044.
- Mejia-Arango S, Nevarez R, Michaels-Obregon A, *et al*. The Mexican Cognitive Aging Ancillary Study (Mex-Cog): Study Design and Methods. *Arch Gerontol Geriatr* 2020;91:104210.
- Bassil DT, Farrell MT, Wagner RG, *et al*. Cohort Profile Update: Cognition and dementia in the Health and Aging in Africa Longitudinal Study of an INDEPTH community in South Africa (HAALSI dementia). *Int J Epidemiol* 2022;51:e217–26.
- Ritchie K, Ritchie CW, Yaffe K, *et al*. Is late-onset Alzheimer's disease really a disease of midlife? *A D Transl Res Clin Interv* 2015;1:122–30.
- Livingston G, Huntley J, Sommerlad A, *et al*. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet* 2020;396:413–46.
- Payne CF, Kohler IV, Bandawe C, *et al*. Cognition, Health, and Well-Being in a Rural Sub-Saharan African Population. *Eur J Population* 2018;34:637–62.
- World Bank Group. Databank. World Development Indicators, 2024. Available: <https://databank.worldbank.org/Sub-Saharan-Africa-Countries/-id/3cf46253>
- Prince M, Wimo A, Guerchet M, *et al*. World Alzheimer Report 2015: The global impact of dementia: An analysis of prevalence, incidence, cost and trends, 2015. Available: <https://www.alzint.org/resource/world-alzheimer-report-2015/>
- United Nations. Department of Economics and Social Affairs. Population Division. World Population Prospects, 2024. Available: <https://population.un.org/wpp/>
- Duhon ME, Miguel E, Njuguna A, *et al*. Preparing for an Aging Africa: Data-Driven Priorities for Economic Research and Policy. *Nat Bureau Eco Res* 2023. Available: <https://www.nber.org/papers/w31750>
- Gross AL, Li C, Briceño EM, *et al*. Harmonisation of later-life cognitive function across national contexts: results from the Harmonized Cognitive Assessment Protocols. *Lancet Healthy Longev* 2023;4:e573–83.
- Hamory J, Miguel E, Walker M, *et al*. Twenty-year economic impacts of deworming. *Proc Natl Acad Sci USA* 2021;118.
- Miguel E, Kremer M. Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities. *Econometrica* 2004;72:159–217.
- Kremer M, Miguel E, Thornton R. Incentives to Learn. *Review of Economics and Statistics* 2009;91:437–56.
- Ofstedal MB, Fisher G, Herzog AR. Documentation of cognitive functioning measures in the health and retirement study. Institute for Social Research, University of Michigan, 2005. Available: <http://hrsonline.isr.umich.edu/sitedocs/userg/dr-006.pdf>
- Gfeller JD, Horn GJ. The East Boston Memory Test: A clinical screening measure for memory impairment in the elderly. *J Clin Psychol* 1996;52:191–6.
- Wechsler D. Wechsler memory scale - WMS-IV technical and interpretive manual. 4th edn. Pearson Clinical Assessment, 2009.
- Raven J. The Raven's Progressive Matrices: Change and Stability over Culture and Time. *Cogn Psychol* 2000;41:1–48.
- Nagarajan N, Burns SD, Riang'a RM, *et al*. Development of the Longitudinal Study of Health and Ageing in Kenya (LOSHAK). *Innov Aging* 2024;8:igad111.
- International Test Commission. The ITC Guidelines for Translating and Adapting Tests, 2017. Available: https://www.intestcom.org/files/guideline_test_adaptation_2ed.pdf
- Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equ Modeling* 1999;6:1–55.
- Reise SP. The Rediscovery of Bifactor Measurement Models. *Multivariate Behav Res* 2012;47:667–96.
- Harvill LM. Standard Error of Measurement. *Educational Measurement* 1991;10:33–41.
- Bontempo DE, Hofer SM. Assessing factorial invariance in cross-sectional and longitudinal studies. In: *Oxford handbook of methods in positive psychology*. New York, NY: Oxford University Press, 2007: 153–75.
- Cernat A, Couper MP, Ofstedal MB. Estimation of Mode Effects in the Health and Retirement Study Using Measurement Models. *J Surv Stat Methodol* 2016;4:501–24.
- Lo Bello S, Sanchez Puerta ML, Winkler H. *From Ghana to America: the skill content of jobs and economic development*. Washington, DC: World Bank, 2019. Available: <https://openknowledge.worldbank.org/handle/10986/31332>
- DiCarlo E, Bello SL, Monroy-Taborda S, *et al*. The Skill Content of Occupations across Low and Middle Income Countries: Evidence from Harmonized Data IZA Discuss Pap Ser (Discussion Paper no. 10224), 2016. Available: <https://docs.iza.org/dp10224.pdf>
- Cheung GW, Rensvold RB. Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. *Struct Equ Modeling* 2002;9:233–55.

- 33 Kenya National Bureau of Statistics The National Treasury and Planning. Kenya Population and Housing Census: Analytical Report on Education and Training Volume XVII, 2019. Available: <https://www.knbs.or.ke/wp-content/uploads/2023/09/2019-Kenya-population-and-Housing-Census-Analytical-Report-on-Education-and-Training.pdf>
- 34 Gross AL, Khobragade PY, Meijer E, *et al.* Measurement and Structure of Cognition in the Longitudinal Aging Study in India-Diagnostic Assessment of Dementia. *J Am Geriatr Soc* 2020;68:S11–9.
- 35 Jones RN, Manly JJ, Langa KM, *et al.* Factor structure of the Harmonized Cognitive Assessment Protocol neuropsychological battery in the Health and Retirement Study. *J Int Neuropsychol Soc* 2024;30:47–55.
- 36 Mani S, Rianga R, Nagarajan N, *et al.* Validation of cognitive and psychosocial tools in Kenya: Findings from the LOSHAK feasibility pilot. 2025.
- 37 World Bank. World Bank Open Data, 2023. Available: <https://data.worldbank.org/indicator/SP.RUR.TOTL.ZS?locations=KE>
- 38 Stephan BCM, Cochrane L, Kafadar AH, *et al.* Population attributable fractions of modifiable risk factors for dementia: a systematic review and meta-analysis. *Lancet Healthy Longev* 2024;5:e406–21.
- 39 Zhang YS, O'Shea B, Yu X, *et al.* Educational Attainment and Later-Life Cognitive Function in High- and Middle-Income Countries: Evidence From the Harmonized Cognitive Assessment Protocol. *J Gerontol B Psychol Sci Soc Sci* 2024;79:gbae005.
- 40 Lövdén M, Fratiglioni L, Glymour MM, *et al.* Education and Cognitive Functioning Across the Life Span. *Psychol Sci Public Interest* 2020;21:6–41.
- 41 Sharp ES, Gatz M. Relationship between education and dementia: an updated systematic review. *Alzheimer Dis Assoc Disord* 2011;25:289–304.
- 42 Foverskov E, Glymour MM, Mortensen EL, *et al.* Education and adolescent cognitive ability as predictors of dementia in a cohort of Danish men. *PLoS ONE* 2020;15:e0235781.
- 43 Kobayashi LC, Jones RN, Briceño EM, *et al.* Cross-national comparisons of later-life cognitive function using data from the Harmonized Cognitive Assessment Protocol (HCAP): Considerations and recommended best practices. *Alzheimers Dement* 2024;20:2273–81.
- 44 Kenya Life Panel Survey. Kenya Life Panel Survey (KLPS) Dataverse, Available: <https://dataverse.harvard.edu/dataverse/KLPS>