Title

Measurement invariance of a neuropsychological battery across urban and rural older adults in Costa Rica.

Authors
Valdivieso-Mora, Esmeralda
Salazar-Villanea, Monica
Johnson, David K

Publication Date
2022-01-25

DOI
10.1080/23279095.2021.2023153

Peer reviewed

HHS Public Access
Author manuscript
Appl Neuropsychol Adult. Author manuscript; available in PMC 2023 July 25.

Published in final edited form as:
Appl Neuropsychol Adult. ; : 1–12. doi:10.1080/23279095.2021.2023153.

Measurement Invariance of a Neuropsychological Battery across Urban and Rural Older Adults in Costa Rica

Valdivieso-Mora Esmeralda[1], **Salazar-Villanea Monica**[2], **Johnson David K.**[3]

[1]Department of Psychology and Public Health, Universidad Centroamericana José Simeón Cañas– El Salvador

[2]Institute for Psychological Research, Universidad de Costa Rica

[3]Alzheimer's Disease Research Center, University of California at Davis – Sacramento

Abstract

This study evaluated the measurement invariance of a neuropsychological battery across rural and urban older adults from Costa Rica. Rural and urban older adults ($N = 295$) from the Epidemiology and Development of Alzheimer′s Disease (EDAD) study in Costa Rica were assessed. The baseline factor model for the EDAD neuropsychological measures was identified with nine neuropsychological measures and three cognitive constructs: Verbal Memory, Spatial Reasoning and Cognitive Flexibility. Measurement and structural invariance were established, and then, group comparisons of the latent cognitive factors were conducted to explore regional disparities. The findings showed that most of the neuropsychological tests in EDAD can be directly compared across the groups, allowing for cognitive constructs comparisons. The rural sample showed a disadvantage in the Spatial Reasoning and Cognitive Flexibility abilities. When age and education were included in the models, differences between the regions disappeared. Having more years of education was associated with higher cognitive abilities, with a larger effect for the rural group. Norms for Costa Rican older adults should consider age and education adjustments. This study contributes to the growing area of measurement invariance in neuropsychological assessment as it highlights the importance of examining the comparability of assessment measures across different cultural groups.

Keywords

Neuropsychological tests; measurement invariance; rural and urban disparities; aging; Costa Rica

---

Studies of pathological and healthy cognitive aging use neuropsychological measures or screening tests for comparing the cognitive functioning between two or more groups (e.g., Masel & Peek, 2009; see also Fields, Ferman, Boeve & Smith, 2011). These studies often find significant disparities by gender (Singh, Jasilionis & Oksuzyan, 2018), socioeconomic status (Goveas et al., 2016; Lye & Burr, 2016), early life health (Case & Paxson, 2010), race and ethnicity (Masel & Peek, 2009; Sloan & Wang, 2005; Zahodne et al.,2016) and income-

Correspondence concerning this article should be addressed to David K. Johnson, Alzheimer's Disease Research Center, UC Davis. School of Medicine, 4860 Y Street, Suite 3900, Sacramento, CA 95817. dkj@ucdavis.edu.

level of countries (Alzheimer's Disease International [ADI], 2020). As disparities in the cognitive health of older adults have become more apparent, examination of the equivalency of neuropsychological test scores across population groups is necessary. Of interest in this study is the examination of the measurement comparability (i.e., measurement invariance) of a neuropsychological battery across rural and urban older adults in Costa Rica.

Urban and rural disparities in cognition of older adults

Where older adults grow and live influences their physical and cognitive health. Level of income and national economic development status have received increasing attention in cognitive studies because two thirds of the worldwide dementia incidence occur in low- and middle-income countries (LMICs) (ADI, 2020; World Health Organization [WHO], 2017a). Economic development in most LMICs is uneven, with urbanized areas have disproportionately higher economic investment, resulting in large differences on most health indices based solely on region of domicile (Friel et al., 2011). Thus, region of domicile (urban versus rural) effects need to be better understood before meaningful comparative testing is used to render cognitive-based diagnoses.

Most of prior research on dementia and cognition of older adults in rural and urban regions has been conducted in high-income countries from Europe, North America and LMICs from Asia. The direction of the disparities found in these countries is consistent across studies, with rural dwellers having higher prevalence of dementia and poorer global cognition (Cassarino, Sullivan, Kenny & Setti, 2015; Russ, Batty, Hearnshaw, Fenton & Starr, 2012). In the Latin American (LA) region, only the 10/66 Dementia Research Group has looked at regional disparities in cognition of older adults. Their results suggest a higher prevalence of dementia was found in urban LA; however, the authors discussed important discrepancies in their results based on the diagnostic criteria used (see Rodriguez, Ferri, Acosta, Guerra, Huang, Jacob, et al., 2008). The studies across the globe that have examined disparities in older adults' cognitive health using neuropsychological assessment tools or screening tools have assumed measurement equivalency across regions (i.e., construct validity) and, consequently, proceeded with hypothesis testing about group differences (e.g., *t* tests, analysis of variance or logistic regressions). This practice can be problematic as most neuropsychological measures have been designed and normed in higher-income countries and for urban-dweller older adults. Nevertheless, they are assumed to be comparable across nations and regions. If measurement equivalence, also known as measurement invariance, is not examined for these tests, there is no certainty of whether the rural and urban disparities found are due to a true difference between groups, or if it is the result of a biased measurement.

Measurement invariance of neuropsychological batteries

Most of the studies on regional disparities of dementia and cognitive functioning of older adults have not addressed the measurement equivalence of the neuropsychological instruments used across regions. The equivalency of assessment tools can be studied with measurement invariance, which is a methodology with a mathematical and theoretical basis that aims at examining if constructs and indicators are similar in different socio-cultural settings, allowing to test meaningfully group mean differences (Little, 2013). Measurement invariance is relevant in the context of clinical and cultural neuropsychology, as it examines if neuropsychological tests measure equivalent cognitive constructs, the tests are related to the construct in the same way across groups, and whether group differences in test scores reflect actual differences in a cognitive ability. If the equivalence or invariance of a battery of neuropsychological tests is established, then data comparisons between groups are valid. Failure to meet measurement invariance of a battery of neuropsychological tests indicates that comparisons across groups should not be conducted. Measurement invariance is a valuable methodology for testing and understanding group differences.

There has been a limited, but growing, amount of research examining the comparability or equivalence of neuropsychological measures in aging across groups. Measurement invariance has been used to examine neuropsychological tests and batteries across different groups of older adults: levels of education (low vs. high education), neurological status (healthy vs. AD), language (English vs. Spanish), and ethnicity and gender (Ávila, Rentería, Witkiewitz et al., 2020; Bertola, Benseñor, Barreto et al., 2020; Blankson & McArdle, 2015; Bowden et al., 2004; Brewster, Tuokko & MacDonald, 2014; Mungas, Widaman, Reed & Farias, 2011; Siedlecki, Honig & Stern, 2008; Tuokko et al., 2009). However, aside from a study conducted in Brazil (see Bertola et al., 2020), no study has examined the measurement invariance of a neuropsychological battery in the LA region. Further, no study of measurement invariance across rural and urban regions has been found in other parts of the world. Research focusing on the measurement equivalence of neuropsychological batteries across rural and urban regions are warranted to learn if the neuropsychological tests and cognitive constructs are comparable across regions. The regional disparities in older adults' pathological and healthy cognition and the limited number of studies in the LA region argue for the importance of evaluating the comparability of neuropsychological measures across urban and rural regions in LA countries.

Costa Rica

Costa Rica offers a unique test case for examining measurement invariance and urban-rural cognitive disparities in older adults. Costa Rica is a Central American country, with an advanced healthcare system and socioeconomic context that differentiates it from its near and distant neighbors. According to the Pan-American Health Organization ([PAHO], 2017), Costa Rica is more urbanized and educated than other countries in Central America. Costa Ricans spend a higher percentage of their gross domestic product in public health and have more healthcare professionals than the average of Central American and LA countries. Their social security system provides universal coverage of medical insurance, in both urban and rural regions with social programs like 'Salud Rural y Comunitaria' (Salas Chavez, 2010). The overall health and socioeconomic advantages are more uniform across the country, reflected in their lower inequality index (GINI) compared to other LA countries. In fact, the meaning of rurality in Costa Rica is evolving due to the decrease in demographic density, displacement of agriculture for factory-related jobs, increased internet use and accelerated growth of rural tourism (Samper & González, 2020). The social and economic development of Costa Rica challenges the traditional views of what rurality is and provides a context to study urban and rural disparities based on geographical location and the natural environment.

Costa Rica has a growing older adult population, with a median age larger than the one for the average LA region. In fact, this country has the highest percentage of adults 65 and older of the Central American and LA region (PAHO, 2017), probably because of their healthcare system and adequate health behaviors (Rosero-Bixby & Dow, 2009). Costa Rica's social, health and population characteristics provide the appropriate context to examine urban-rural cognitive disparities that are independent of differences in literacy and minimize SES disparity.

The present study

In LA, the study of measurement invariance and regional disparities using a neuropsychological battery is warranted. Most of the professional neuropsychologists in the region conduct assessments using instruments that have not been developed in this region of the world and for its own population. Some practitioners use tests that are not culturally adapted or use scoring norms from other countries (Arango-Lasprilla, 2015). These practices could result in biased diagnoses for using measures with unexamined construct validity (Arango-Lasprilla, Stevens, Morlett-Paredes, Ardila & Rivera, 2017). As recent and increasing efforts in designing culturally adapted assessment instruments in the LA region emerge, the study of measurement invariance of internationally and commonly used neuropsychological tests across demographic groups offers practitioners and researchers a way to examine the validity of these instruments.

The purpose of this study was to test the construct validity of several neuropsychological measures across rural and urban regions in a sample of older adults from Costa Rica (are these measures equivalent across regions?), and then, examine the group mean differences in the cognitive domains the measures assess (if equivalent, are there differences in the cognitive domains between the regions?). Importantly, we sought to estimate the contribution of education in the regional disparities in cognition between the regions (are the differences between regions explained by education?). For these purposes we used baseline neuropsychological measures from the Epidemiology and Development of Alzheimer's Disease study in Costa Rica (EDAD). We hypothesized that a parsimonious and conceptually meaningful neuropsychological battery would be invariant (i.e., equivalent or comparable) across urban and rural groups, and that mean differences in the neuropsychological constructs would exist across regions. Further, we anticipated that the urban and rural disparities would be explained by an effect of education.

Methods

The EDAD was a two-year (2014 – 2016) exploratory/developmental research project that aimed to characterize cognitive aging differences between rural/developing and urban/developed regions of Costa Rica. The EDAD study applied a comprehensive neuropsychological battery used at a U.S. federally funded Alzheimer's Disease Research Center to characterize Costa Rican older adults' cognitive functioning. The cognitive data presented here are one part of a larger fitness, psychosocial and health-related assessment battery. The present study is framed within this larger project.

Participants

Older adult residents of San Jose, the capital city of Costa Rica (urban), were compared to residents of Liberia (rural). San Jose is characterized by a highly urbanized area. Liberia is a rural region in the state of Guanacaste that is slowly becoming suburbanized. The main economic activity in all the Guanacaste region is tourism and agriculture. The EDAD project focused on the Guanacaste region because of its low population density and preserved rurality. According to the 2011 Costa Rican Census, Guanacaste is the state with lower changes in their rural population; that is, this region is not becoming urbanized as fast as other states in Costa Rica and has the lowest internal and external migration rates of the country (Instituto Nacional de Estadistica y Censos [INEC], 2011). The previously described characteristics suggest that Guanacaste is a rural region with a steady economy and population, with potentially more homogeneity than in other regions of Costa Rica. The stability of this region is ideal for the purposes of this study given the homogeneity of its population and context.

Sample.—Participants of EDAD were 295 Costa Rican older adults between the ages of 60 and 85, who lived in the Greater Metropolitan Area of San Jose (n = 181) and Liberia (n = 114). For all participants, Spanish was their primary or only language. Relevant demographic information about the sample is presented in Table 1.

Recruitment.—The EDAD advertised to recruit two different convenience samples of Costa Rican older adults who were living either in rural (Guanacaste) or urban (San Jose) regions of the country. Participants were included if they were cognitively healthy older adults. Eligibility requirements for EDAD included: adults be between 65 to 85 years of age, community dwellers, be free of cognitive impairment (MMSE > 23, Blesa et al., 2001; Folstein, Folstein, & McHugh, 1975), able to read and write, have adequate visual and auditory abilities to complete study procedures, have a stable dose of medication for a minimum of 30 days prior to screening, sign an informed consent, and verbally assent to participate in all scheduled evaluations. Participants were excluded based on the following criteria: moderate cognitive impairment (determined by a MMSE score less than 24), current clinically significant major psychiatric disorder or significant psychiatric symptoms, history of clinically-evident stroke, brain trauma and neurocognitive disorder, clinically-significant infection within the last 30-days, history of drug or alcohol abuse or dependence within the past 2 years, and significant pain or musculoskeletal disorder.

Neuropsychological Assessment

A comprehensive neuropsychological battery was compiled to assess different cognitive domains in EDAD. The 16 cognitive measures included were: *Logical Memory I* and *II* from the WMS-III (Wechsler, 1997a), *Verbal Fluency Animals* and *Vegetables* (Goodglass & Kaplan, 1983), *Trail Making Test A* and *B* (Armitage, 1946), *Digit Symbols* and *Block Design* from the WAIS-R (Wechsler, 1997b), *Stroop Color Naming* (Golden, 1978), *Boston Naming Test* (see Goodglass & Kaplan, 1983; Fernandez & Fulbright, 2015; Jahn et al., 2013), *Selective Reminding Test* (Grober, Ocepek-Welikson & Teresi, 2009), *Crossing off* (Botwinick & Storandt, 1973), *Spatial Relations* from the DAT (Bennett, Seashore, Wesman, 1947), *Paper Folding* test (Workman & Lee, 2004), *Hidden Patterns* (Vandenberg & Kuse, 1978), and *Identical Pictures* (Ekstrom, French, Harman & Dermen, 1976). Spanish versions of the Wechsler's subtests were used. For the remaining measures, a committee of bilingual U.S. and Costa Rican researchers on aging reviewed all the measures administered in Spanish. Preliminary analysis of these measures found that some of these tests violated univariate normality assumptions. These tests were excluded from the baseline model. Descriptive statistics and results from this analysis are provided in the Supplementary Material.

Procedures

Participants were recruited through the Programa Institucional para la Persona Adulta Mayor (Institutional Program for the Older Adults of the University of Costa Rica), the Asociación Gerontológica Costarricense (Costa Rican Gerontology Association), the Programa de Ciudadano de Oro de la Caja Costarricense del Seguro Social (Golden Citizen Program of the Costa Rican Social Security Bureau), and other groups such as retired teachers and community groups. Flyers with information about the study and eligibility criteria were posted on different state and community centers of San Jose and Guanacaste, where older adults are users. Eligible participants were read and had the informed consent form verbally explained the informed consent form to them. All participants signed the informed consent form and verbally agreed to participate in each assessment session. One-on-one interviews and testing were conducted in private offices of the UCR campus and the community centers. Each participant attended two-to-three sessions of data collection. During the first session, participants completed socio-demographic, psychosocial and health questionnaires. During the second assessment session, participants were administered the neuropsychological test battery described previously.

Once all the data was collected, the testing protocols were revised by a clinical neuropsychologist in Costa Rica and scores were transferred to a summary score sheet. All administered materials and score sheets were digitalized to populate a predesigned database. The finalized and cleaned database was used for the analyses.

Statistical Analyses

The baseline model was identified using an Exploratory Factorial Analysis (EFA) / Confirmatory Factorial Analysis (CFA) approach (see Siedlecki, Honing & Stern, 2008). The identified model consisted of three factors: Verbal Memory, Spatial Reasoning and Cognitive Flexibility, and nine indicators: Verbal Fluency – animals, Logical Memory I, Selective Reminding Test, Block Design, Hidden Patterns, Spatial Relations, Stroop, Digit-Symbols, Trails Making Test – B (see Figure 1). The model identification steps and results can be found in the Supplementary Material.

Measurement invariance testing.—Multigroup CFA (MGCFA) was used to examine three levels of measurement invariance: configural, metric and scalar. All levels of measurement invariance testing were conducted in Mplus version 8.2 (Muthén & Muthén, 1998-2017). MGCFA allowed simultaneous analysis of CFA in the urban and rural sample (see Brown, 2015) to evaluate the equivalence of the factor model at different levels (i.e., factor loadings, intercepts, latent means). Urban was set as the reference group, with differences estimated for rural participants. Covariate for age and education were included in the analyses. Measurement invariance testing followed Vandenberg and Lance (2000) and Little (2013) suggestions of progressively restricting parameters in the model by testing nested models.

The configural invariance model (Model 1) examined if the conceptual framework in the factor model was the same across the urban and rural group, that is, if there was an identical factor structure in each group. This model allowed factor loadings, intercepts and residuals to be estimated freely. Once configural invariance was established, the next nested models in measurement invariance assured that latent factors had the same meaning in different groups. The metric invariance model (Model 2), also referred as the weak invariance model (see Brown, 2015), tested if the urban and rural groups responded to the indicators in the same way. This model was tested by constraining all factor loadings to equality and allowing intercepts and residual variances to differ across groups. The scalar invariance model (Model 3), also called strong invariance (Brown, 2015), tested if latent factor means and variances can be compared across the urban and rural groups, and the relationship of latent factors with external variables (e.g., covariates). In other words, scalar invariance provides evidence of construct comparability (see Little, 2013). The metric and scalar invariance models were tested by examining model fit indicators: χ^2, *df*, RMSEA, TLI and CFI. Similarly, the change in value of those indicators was examined (e.g., Model 2 vs. Model 1, Model 3 vs. Model 2).

Model fit test and indices were reviewed using the same criteria used in the EFA and CFA steps. The metric and scalar invariance models were tested by examining different criteria. Model fit test and indices were reviewed using the same criteria used in the EFA and CFA steps. Metric and scalar measurement invariances were also assessed by evaluating the change of χ^2 and *df* ($\Delta\chi^2$, Δdf), RMSEA value (ΔRMSEA), TLI value (ΔTLI) and CFI value (ΔCFI) from the previous model (e.g., Model 2 vs. Model 1, Model 3 vs. Model 2), following criteria conventionally used (see Chen 2007, cited in Putnick & Bornstein, 2016; Little, 2013; Cheung & Rensvold, 2002) in other neuropsychological studies (see Blankson & McArdle, 2013; Siedlecki, Honing & Stern, 2008; Tuokko et al., 2009). Meeting all the previously listed criteria indicated that the more restricted model (e.g., Model 3) fit the data better than the less restricted ones (e.g., Models 2 and 1).

When the full metric and/or scalar invariance tests violated the model-fit criteria, modification indices were used to identify and freely estimate non-invariant parameters. The resulting model with some freed parameters and many invariant parameters was called *partial invariance model* (Byrne, Shavelson & Muthén, 1989). Testing partial invariance models is important because full or partial metric invariance must be established to evaluate scalar invariance; similarly, full or partial scalar invariance is required to compare the factor variances and means across groups (Vandenberg & Lance, 2000). The identification of the partial invariance model was based on a backward method for testing partial factorial invariance (Jung & Yoon, 2016).

Group-mean differences with structural invariance testing.—Once the final partial scalar invariance model was established, structural invariance was tested. Structural invariance consists of two models: (a) factor variance invariance model, and (b) factor mean invariance model. The factor variance invariance model compared the variances of the latent constructs across the groups. Model fit test and indices were compared to those of the final measurement invariance model (i.e., partial scalar model). Once factor variance invariance was established, then the factor means were compared across the groups. Again, model fit test and indices were compared to those of the factor variance invariance model. If $\Delta\chi^2$ was not significant, the model indicated that the latent means of the cognitive constructs were equivalent across regions. However, if $\Delta\chi^2$ is significant, new models that sequentially released constrained factor means should be tested, until the non-invariant and invariant cognitive constructs are identified. This stage of model testing was conducted to determine if regional disparities existed between rural and urban Costa Rican older adults at a factor level.

Effect of age and education in the measurement and structural invariance models.—New models with age and education as covariates were tested following the steps previously described for measurement and structural invariance testing. Model fit test and indices were examined using the same criteria mentioned above. In addition, regression coefficients of each cognitive construct on the covariates were examined.

Results

Demographic Characteristics

Sample demographic characteristics are presented by region in Table 1. The percentage of female participants was similar between the urban and rural regions. The mean age and education markedly differed across regions. Urban CR older adults had a lower mean age and more years of education, compared to their counterparts in the rural region. The significant differences in age and education found between the groups were considered in the testing of the measurement invariance models. Table 2 presents the means, standard deviations and univariate statistics of the neuropsychological measures identified in the baseline model. Results of the CFA baseline model (see Figure 1) can be found in the Supplementary Material.

Measurement invariance across regions

Using MGCFA, different nested models progressively evaluated the configural, metric and scalar invariance of the three-factor model with nine indicators across the urban and rural groups. The configural model (Model 1) tested whether the relationship among the Verbal Memory, Spatial Reasoning and Cognitive Flexibility latent constructs and the neuropsychological tests was invariant across the regional groups. As shown in Table 3, results indicated that the baseline model had an excellent fit across groups. Configural invariance of the model was established.

The metric model (Model 2) tested if the magnitude of the factor loadings was equivalent across groups. The resulting fit test and indices of Model 2 were compared to those of

Model 1 (see Table 3). The $\Delta\chi^2$ suggested that the metric invariance model resulted in a significant decrease in fit relative to Model 1. In follow-up analyses using the backward method, the partial metric invariance model was established (Model 2.1), with *Spatial Relations* freely estimated. The established partial metric model indicates that cross-region comparisons are acceptable if the cognitive constructs tested (i.e., Verbal Memory, Spatial Reasoning and Cognitive Flexibility) were measured with the neuropsychological tests that have invariant factor loadings. That is, all observed variables included in the model (with the exception of *Spatial Relations*, which is noninvariant) relate to the latent factors in the same way across regions. The test of model fit and fit indices for Model 2.1 are presented in Table 3.

Based on the partial metric model, the full scalar measurement invariance model (Model 3) was tested by constraining the model intercepts to equality across groups. A partial scalar model (Model 3.1, in Table 3) was established when the *Logical Memory I* and *Spatial Relations* intercepts were released. This finding suggested that it could be appropriate to compare group-mean differences in the cognitive constructs (Verbal Memory, Spatial Reasoning and Cognitive Flexibility) as they capture the mean differences in the scores of the neuropsychological measures across regions, except for *Logical Memory I* and *Spatial Relations* which cannot be directly compared across groups (noninvariant intercepts).

Structural invariance across rural and urban regions

Structural invariance was tested with two additional models with the purpose of comparing the cognitive constructs at the group-mean level. The first model tested the factor variance invariance, which constrained all factor variances to 1 (i.e., to be equal across regions). The resulting $\Delta\chi^2$ indicated that Costa Rican older adults in the urban and rural region had equivalent amounts of individual differences in each cognitive factor (i.e., range of scores on each latent factor does not vary across groups). The second model tested to factor mean invariance, which constrained to 0 (zero) the factor means to be equal across regions. As shown in Table 3, the resulting $\Delta\chi^2$ indicated that Verbal Memory was the only invariant (i.e., equivalent or comparable) factor, while the other two latent factors had factor means significantly different from zero (i.e., not comparable across groups). The model suggested that the sample of rural and urban Costa Rican older adults had a comparable functioning of their Verbal Memory on average ($p = 0.25$); with standardized factor means for Spatial Reasoning and Cognitive Flexibility lower in the rural group ($\mu = -0.764$, SE = 0.140, $p < 0.001$ and $\mu = -1.081$, SE = 0.145, $p < 0.001$, respectively) than in the urban group. The statistics of the final structural invariance model tested across regions can be found in the Supplemental Material.

Effect of education and age in the urban-rural disparities in the cognitive factors

All levels of measurement invariance (i.e., configural, metric and scalar) and structural invariance (variance and mean invariance) testing were conducted with age and education. All model fit test and indices are presented in Table 4. With the established configural, metric and partial scalar invariance models, the structural invariance model was across regions. The changes in model fit and indices indicated that there was an equal amount of interindividual variation in the three cognitive constructs across groups. As shown in

Table 5, the mean invariance model was not significant either. Unlike the previous set of models tested (regional comparison without age and education as covariates), the latent mean invariance model with the covariates suggested that factor means for all three cognitive constructs were invariant (i.e., comparable) across the rural and urban groups of older adults. No follow-up comparisons were warranted.

Regression coefficients for the latent cognitive constructs regressed on age and education are presented in Table 6. In both urban and rural groups, Verbal Memory, Spatial Reasoning and Cognitive Flexibility had a negative relationship to age and had a positive relationship with education. These results indicated that with older age, latent scores were lower, while with more years of education the latent scores were higher. An examination of the effect sizes (R^2) for age and education suggested that these covariates had a larger effect in the rural group than the one found in the urban group.

Summary of results.—A broad neuropsychological battery made up of nine subtests used in the EDAD study resulted in a three-factor model. All subtest except *Logical Memory I and Spatial Relations* show equivalence between rural and urban dwelling older adults. These two subtests required relaxation of factor loadings and intercepts. Group-mean differences for the test constructs (latent means) of Spatial Reasoning and Cognitive Flexibility were found. These group-mean differences were eliminated when age and education were included as covariates. Age had a negative small effect on the mean latent scores of the cognitive constructs of both regional groups, while education had a positive and larger effect on the constructs for the rural group in comparison to the urban sample.

Discussion

The present study examined and established measurement invariance across a sample of urban and rural older adults in Costa Rica who were part of the EDAD study. We found a three-factor model consisting of Verbal Memory, Spatial Reasoning and Cognitive Flexibility, with nine neuropsychological indicators. The results showed that direct comparisons across regions can be made for the cognitive factors and most of the neuropsychological tests. When regional differences in the cognitive factors were tested, a rural disadvantage was found for the Spatial Reasoning and Cognitive Flexibility constructs. Further measurement invariance model testing suggested that the rural disadvantage can be explained by the influence age and education had in the cognitive functioning of the Costa Rican older adults in the EDAD study.

In this study, the process of testing for invariance followed the logical model specified by Little (2013) (see also Vandenberg & Lance, 2000), the first test was designed to test configural invariance, followed by metric (weak) and scalar (strong) invariance. The results of this study established that the reduced set of nine measures from the EDAD neuropsychological battery demonstrated configural and partial metric and scalar invariance. The partial metric invariance model indicated that all factor loadings associated to their specific latent constructs were comparable across regions (with Spatial Relations being the exception). The size of the factor loadings and explained variance can be considered estimates of the reliability (i.e., consistent results across the two samples) of the neuropsychological measures (see Brown, 2015). This suggests that in the EDAD sample, these measures were consistently and meaningfully related to their cognitive constructs across groups. Further, establishing partial metric invariance evidenced discriminant (i.e., how distinguishable the tests are by the cognitive ability they measure) and convergent validity (i.e., how related the tests that measure one specific cognitive ability are) measures that of the neuropsychological measures. Discriminant and convergent validity are types of construct validity, which informs that when the neuropsychological tests were administered to the urban and rural sample of older adults, the tests measured the same cognitive abilities (see Little, 2013).

The finding of partial scalar invariance established that urban and rural Costa Rican older adults with the same cognitive ability (e.g., Cognitive Flexibility) were measured with the neuropsychological tools used in the EDAD study, they would produce comparable scores on the measures related to that cognitive ability (e.g., Digit Symbol). Therefore, we would conclude that the measures are unbiased across urban and rural populations. The invariance found in the loadings and intercepts indicates that any differences across groups on the cognitive constructs can be attributed to a true latent variable group difference. Establishing partial scalar invariance is critical because only in the context of scalar invariance can group comparisons of the latent means be considered valid (see Little, 2013).

Identification of the factor model and assessment of the measurement invariance model has important implications for the clinical and scientific understanding of pathological and normal cognition of older adults in the LA region as well as in the health disparities research. Most of the neuropsychological measures used in the LA region have been developed in other outside countries and translated and adapted into Spanish. These measures are commonly used in this region for clinical and research purposes without assessing or questioning their equivalency (see Rivera, Mascialino, Brooks, et al., 2020). The use of univariate and multivariate statistics (e.g., t tests, correlations, ANOVAs, etc.) is frequent in the health disparities literature, with an implicit assumption that the scales and measures used are comparable (invariant) across groups. Studies examining the factor structure and measurement invariance are important so that when disparities are found, they are not the result of biased measures, as opposed to true mean scores. The same rationale applies to the administration of neuropsychological tests for clinical and diagnostic purposes. In this sense, this study represents an important step toward the ongoing process of examining construct validity and reliability of neuropsychological measures. And, therefore, this study contributes to the recent, but growing, area of cultural neuropsychology (see Cagigas & Manly, 2014) and the call to action that clinical neuropsychologists have expressed to increase cultural awareness in research and clinical services (Arango-Lasprilla, 2015; Rivera-Mindt, Byrd, Saez & Manly, 2010).

The findings of invariance across groups excluded two measures. *Logical Memory I* and *Spatial Relations* were noninvariant measures that significantly contributed to model misfit in the metric and scalar invariance testing. There are two major implications of this finding. Firstly, the observed scores of these measures were not directly comparable across regions. In the rural sample, this neuropsychological measure did not load in Spatial Reasoning, suggesting that *Spatial Relations* was a biased measure of Spatial Reasoning in the rural older adult group. Secondly, the non-invariance found in *Logical Memory I* and *Spatial Relations* suggested that the differences by region in their mean scores were not due to true cognitive differences. In other words, the lower mean scores observed in these tests among rural Costa Rican older adults were likely not due to poorer abilities in their Verbal Memory and Spatial Reasoning, as these observed measures appeared to have a different meaning across groups. Caution is advised in the administration and interpretation of *Logical Memory I* and *Spatial Relations* in Costa Rican older adults. In the present study, the sources of model misfit in each nested model of measurement invariance were examined and partial invariance was established. This approach proposed by Byrne and colleagues (1989) has been applied in other studies examining neuropsychological tests across different groups of older adults (see Mungas et al., 2011; Tuokko et al., 2009). Allowing the model to be partially invariant is considered a strength of the present study.

The test of the structural invariance models showed that there were mean group differences between urban and rural Costa Rican older adults on Spatial Reasoning and Cognitive Flexibility factors. However, the disadvantage for rural Costa Ricans in these cognitive factors disappeared when age and education were included in the analysis. This finding has two main implications. First, scores and norms should be adjusted by age and education when interpreting neuropsychological measures and cognitive factors for both urban and rural samples of Costa Rica. And, secondly, the findings indicate that the regional disparities in cognition can be explained by the older age and lower number of years in education in the rural sample. The contribution of age and education on cognition has been documented in previous studies (Cassarino et al., 2015; Guerchet et al., 2013; Hall, Gao, Unverzagt & Hendrie, 2000; Jia et al., 2014; Sharma, Salig Ram & Anupam, 2013; Weden, Shih, Kabeto & Langa, 2018). As of now, there are no previous studies examining measurement invariance of a neuropsychological battery across regions (i.e., urban and rural) in a sample of Latin American older adults. Yet, despite the well-known effect that region/geography has on health disparities (Healthy People 2020, 2008), the study of measurement invariance in cognitive tests across rural and urban samples of older adults was inexistent until the present study.

Although age and education had a significant influence on the cognitive constructs in the urban and rural sample, the effect of these two demographic variables behaved differently across groups. The negative regression coefficients associated with age were similar in their sizes across the regions. As it has been well established (WHO, 2017b), we found a negative association between age and cognitive functioning of older adults. In contrast, the positive regression coefficients associated to education differed across regions. Larger coefficients were found in the rural group than in the urban group, which indicated that the influence of education on the latent constructs was stronger in the rural sample compared to the urban region.

Why would education have a larger positive influence on cognition in the rural than in the urban region? One likely explanation is that within the context of a low average in years of education in the rural sample, any increase in years of education would have a larger impact in the cognitive functioning of rural older adults, than in the urban region. In simpler words, having more years of education brings larger health advantages in underprivileged regions. In this way, if the regions and environmental contexts in which individuals grow and age are disadvantaged, even the smallest increase in years of education for rural older adults can result in a positive impact in their cognitive health and overall development. These education-related findings are consistent with the socioecological model (Kaplan, Everson & Lynch, 2000) and the bioecological model (Bronfenbrenner, 2005), which coincide in the inter-dependence that exists between an individual's development and their environmental context (e.g., resources, social network, access to health services, living conditions and culture).

There are some limitations of this study that should be noted. First, results of this study cannot be generalized to all older adults living in rural and urban regions of LA as the present study was limited to the cognitively healthy older adults of the urban and rural regions sampled in Costa Rica for the EDAD study. Because this model was not intended for applied use, and values of its latent factors have not been cross validated in other independent samples (i.e., urban and rural older adults from other provinces in Costa Rica), further testing and replication is warranted. Comparison of the identified model with conceptually different older adult groups would help evaluate the construct validity of the latent factors (see Delis et al., 2003).

Further analyses are needed to understand the use of *Logical Memory I and II, Verbal Fluency Vegetables, Boston Naming Test, Trails Making Test* and *Paper Folding* in the neuropsychological assessment of Costa Rican older adults. From the findings of the present study, caution is advised in the administration and interpretation of these measures for the population of interest. It was out of the scope of this study to analyze the psychometric properties of each individual measure and, thus, more analysis is warranted.

Despite these limitations, the present study was unique in that a comprehensive evaluation of measurement invariance was conducted across rural and urban groups of older adults in Costa Rica. Before this project, such studies had not been conducted for regional groups in Central America, LA and other regions of the world. This study highlights the importance of testing for measurement equivalence. An important strength of our analysis was its ability to test partial measurement invariance models. For the EDAD study, this approach revealed that the EDAD neuropsychological tests could be used to measure the verbal memory, spatial reasoning and cognitive flexibility abilities in Costa Rican older adults, and results could be similarly interpreted across the urban and rural regions. Only Logical Memory I and Spatial Relations were the exceptions of this finding, as their absolute levels of performance may not be comparable across groups and may give rise to misleading interpretations of their test scores. The use of country-specific norms that adjust scores by age and years of education when assessing older adults in rural Costa Rica are warranted. Additionally, this study becomes a step closer to understanding regional health disparities on older adults living in LMICs. This study can serve as a push to advance the research of cognition in older adulthood in the Central American region.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Analyses were conducted as part of Esmeralda Valdivieso Mora´s dissertation in fulfillment of a dual doctorate in Clinical Psychology and Gerontology. This study was supported by Fogarty International Center (FIC) Epidemiology and Development of Alzheimer´s Disease in Costa Rica (R21 TW009665) and National Institute on Aging (NIA) Alzheimer´s Disease Core Center (P30 AG010129).

References

Alzheimer's Disease International (2020). World Alzheimer Report 2020. Design, dignity, dementia: Dementia-related design and the built environment. Alzheimer's Disease International (ADI), London.

Arango-Lasprilla JC (2015). Commonly used neuropsychological tests for Spanish speakers: normative data from Latin America. Neurorehabilitation, 37(4), 489–491. doi: 10.3233/NRE-151276 [PubMed: 26577888]

Arango-Lasprilla JC, Stevens L, Morlett Paredes A, Ardila A, & Rivera D (2017). Profession of neuropsychology in Latin America. Applied Neuropsychology: Adult, 24(4), 318–330. doi: 10.1080/23279095.2016.1185423 [PubMed: 27282450]

Armitage SG (1946). An analysis of certain psychological tests used for the evaluation of brain injury. Psychology Monographs, 60, 1–47.

Avila JF, Rentería MA, Witkiewitz K, Verney SP, Vonk JMJ, & Manly J (2020). Measurement invariance of neuropsychological measures of cognitive aging across race/ethnicity by sex/gender groups. Neuropsychology, 34, 3–14. 10.1037/neu0000584 [PubMed: 31464473]

Bennett GK, Seashore HG, & Wesman AG (1947). Differential aptitude tests. San Antonio, TX, US: Psychological Corporation.

Bertola L, Benseñor IM, Barreto SM, Moreno A, Griep RH, Viana MC, Lotufo PA, & Suemoto CK (2020). Measurement invariance of neuropsychological tests across different sociodemographic backgrounds in the Brazilian Longitudinal Study of Adult Health (ELSA-Brasil). Neuropsychology, 34(2), 227–234. 10.1037/neu0000597 [PubMed: 31829661]

Blankson AN, & McArdle JJ (2015). Measurement Invariance of Cognitive Abilities Across Ethnicity, Gender, and Time Among Older Americans. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 70(3), 386–397. 10.1093/geronb/gbt106 [PubMed: 24170715]

Blesa R, Pujol M, Aguilar M, Santacruz P, Bertran-Serra I, Hernández G, Peña-Casanova J (2001). Clinical validity of the 'mini-mental state' for Spanish speaking communities. Neuropsychologia, 39(11), 1150–1157. doi: 10.1016/S0028-3932(01)00055-0 [PubMed: 11527552]

Botwinick J, & Storandt M (1973). Speed functions, vocabulary ability, and age. Perceptual and Motor Skills, 36(3), 1123–1128. [PubMed: 4711963]

Bowden SC, Cook MJ, Bardenhagen FJ, Shores EA & Carstairs JR (2004). Measurement invariance of core cognitive abilities in heterogeneous neurological and community samples. Intelligence, 32(4), 363–389.

Brewster P, Tuokko H, & MacDonald SWS (2014) Measurement equivalence of neuropsychological tests across education levels in older adults. Journal of Clinical and Experimental Neuropsychology, 36(10), 1042–1054. doi: 10.1080/13803395.2014.967661 [PubMed: 25396690]

Bronfenbrenner U (2005). Making human beings human: Bioecological perspectives on human development. Sage Publications.

Brown TA (2015). Confirmatory factor analysis for applied research. New York, NY: Guilford Press.

Byrne BM, Shavelson RJ & Muthén B (1989). Testing the equivalence of factor covariance and mean structures: the issue of partial measurement invariance. Psychological Bulletin, 105(3), 456–466.

Cagigas XE, & Manly JJ (2014). Cultural neuropsychology: The new norm. In Parsons MW, Hammeke TA, & Snyder PJ (Eds.), Clinical neuropsychology: A pocket handbook for assessment (pp. 132–156). Washington, DC, US: American Psychological Association.

Case A & Paxson C (2010). Causes and consequences of early-life health. Demography, 47(1), s65–s85. [PubMed: 21302429]

Cassarino M, O'Sullivan V, Kenny RA, & Setti A (2015). Environment and cognitive aging: A cross-sectional study of place of residence and cognitive performance in the Irish Longitudinal Study on Ageing. Neuropsychology, 1–15.

Cheung GW & Rensvold RB (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. Structural Equation Modeling, 9(2), 233–255.

Delis D, Jacobson M, Bondi M, Hamilton J, & Salmon D (2003). The myth of testing construct validity using factor analysis or correlations with normal or mixed clinical populations: Lessons from memory assessment. Journal of the International Neuropsychological Society, 9(6), 936–946. doi:10.1017/S1355617703960139 [PubMed: 14632252]

Ekstrom RB, French JW, Harman HH, & Dermen D (1976). Manual for kit of factor-referenced cognitive tests. Princeton, NJ: Educational Testing Service.

Fernandez AL, & Fulbright RL (2015). Construct and Concurrent Validity of the Spanish Adaptation of the Boston Naming Test. Applied Neuropsychology: Adult, 22(5), 355–362. doi: 10.1080/23279095.2014.939178 [PubMed: 25668293]

Fields JA, Ferman TJ, Boeve BF, & Smith GE (2011). Neuropsychological assessment of patients with dementing illness. Nature Reviews Neurology, 7, 677–687. [PubMed: 22045270]

Folstein MF, Folstein SE, & McHugh PR (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatric Research, 12(3), 189–198. doi: 10.1016/0022-3956(75)90026-6 [PubMed: 1202204]

Friel SM, Akerman T, Hancock J, Kumaresan M, Marmot T, & Melin D (2011). Addressing the social and environmental determinants of urban health equity: evidence for action and a research agenda. Journal of Urban Health, 88(5), 860–874. doi:10.1007/s11524-011-9606-1. [PubMed: 21877255]

Golden CJ (1978). Stroop Color and Word Test: A Manual for Clinical and Experimental Uses. Chicago: Stoelting Co.

Goveas JS, Rapp SR, Hogan PE, Driscoll I, Tindle HA, Smith JC, … Espeland MA (2016). Predictors of Optimal Cognitive Aging in 80+ Women: The Women's Health Initiative Memory Study. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences, 71(Suppl 1), S62–S71. 10.1093/gerona/glv055 [PubMed: 26858326]

Goodglass H, & Kaplan E (1983). Boston Diagnostic Aphasia Examination. Philadelphia, PA: Lea & Febiger.

Grober E, Ocepek-Welikson K, & Teresi JA (2009). The free and cued selective reminding test: evidence of psychometric adequacy. Psychology Science Quarterly, 51(3), 266–282.

Guerchet M, Mbelesso P, Bandzouzi B, Pilleron S, Clément J-P, Dartigues J-F, & Preux P-M (2013). Comparison of rural and urban dementia prevalences in two countries of central Africa: The EPIDEMCA study. Journal of the Neurological Sciences, 333, e328–e328. 10.1016/j.jns.2013.07.1219

Hall KS, Gao S, Unverzagt FW, Hendrie HC (2000). Low education and childhood rural residence: Risk for Alzheimer's disease in African Americans. Neurology, 24. doi: 10.101212/WNL.54.1.95

Healthy People 2020. (2008). Phase I Report: Recommendations for the framework and format of Healthy People 2020. Secretary's Advisory Committee of Health Promotion and Disease Prevention Objectives for 2020.

Jahn DR, Mauer CB, Menon CV, Edwards ML, Dressel JA, & O'Bryant SE (2013). A brief Spanish-English equivalent version of the Boston Naming Test: A project FRONTIER study. Journal of Clinical and Experimental Neuropsychology, 35(8), 835–845. [PubMed: 23998641]

Jia J, Wang F, Wei C, Zhou A, Jia X, Li F, … Dong X (2014). The prevalence of dementia in urban and rural areas of China. Alzheimer's & Dementia: The Journal of the Alzheimer's Association, 10(1), 1–9. 10.1016/j.jalz.2013.01.012

Jung E & Yoon M (2016). Comparison of three empirical methods for partial factor invariance: forward, backward, and factor-ratio tests. Structural Equation Modeling: A Multidisciplinary Journal, 23(4), 567–584. doi: 10.1080/10705511.2015.1138092

Kaplan GA, Everson SA, & Lynch JW (2000) The contribution of social and behavioral research to an understanding of the distribution of disease: a multilevel approach. In Smedley BD and Syme SL (Eds.) Promoting health: intervention strategies from social and behavioral research. Washington, D.C.: National Academies Press.

Little TD (2013). Longitudinal Structural Equation Modeling. New York, NY: Guilford Press.

Lye J, & Burr JA (2016). Socioeconomic status across the life course and cognitive function among older adults. Journal of Aging and Health, 28(1). doi: 10.1177/0898264315585504.

Masel MC, & Peek MK (2009). Ethnic differences in cognitive function over time. Annals of Epidemiology, 19(11), 778–783. 10.1016/j.annepidem.2009.06.008 [PubMed: 19656690]

Mungas D, Widaman KF, Reed BR, & Farias ST (2011). Measurement Invariance of Neuropsychological Tests in Diverse Older Persons. Neuropsychology, 25(2), 260–269. 10.1037/a0021090 [PubMed: 21381830]

Muthén LK, & Muthén BO (1998-2017). Mplus user's guide (8th ed.). Los Angeles, CA: Muthén & Muthén.

Putnick DL & Bornstein MH (2016). Measurement invariance conventions and reporting: the state of the art and future directions for psychological research. Developmental Review, 41, 71–90. doi: 10.1016/j.dr.2016.06.004. [PubMed: 27942093]

Rivera D, Mascialino G, Brooks BL, Olabarrieta-Landa L, Longoni M, Galarza-Del-Angel J & Arango-Lasprilla JC (2020) Multivariate base rates of low scores on tests of executive functions in a multi-country Latin American sample. Developmental Neuropsychology, 1–15. Advance online publication. 10.1080/87565641.2020.1863407

Rivera-Mindt M, Byrd D, Saez P & Manly J (2010) Increasing culturally competent neuropsychological services for ethnic minority populations: A call to action. The Clinical Neuropsychologist, 24(3), 429–453, doi:10.1080/13854040903058960 [PubMed: 20373222]

Rodriguez JJL, Ferri CP, Acosta D, Guerra M, Huang Y, Jacob K, … for the 10/66 Dementia Research Group. (2008). Prevalence of dementia in Latin America, India, and China: a population-based cross-sectional survey. Lancet, 372(9637), 464–474. 10.1016/S0140-6736(08)61002-8 [PubMed: 18657855]

Russ TC, Batty GD, Hearnshaw GF, Fenton C, & Starr JM (2012). Geographical variation in dementia: systematic review with meta-analysis. International Journal of Epidemiology, 41, 1012–1032. [PubMed: 22798662]

Samper M & González H (2020). Characterization of rural spaces in Costa Rica and alternative methods for its measurement [Caracterizacion de los espacios rurales en Costa Rica y propuesta de alternativas metodológicas para su medición]. United Nations - Economic Comission for Latin America and the Caribbean [Comision Economica para America Latina y el Caribe, CEPAL],

Siedlecki KL, Honig LS, & Stern Y (2008). Exploring the structure of a neuropsychological battery across healthy elders and those with questionable dementia and Alzheimer's disease. Neuropsychology, 22(3), 400–411. doi:10.1037/0894-4105.22.3.400 [PubMed: 18444718]

Singh PK, Jasilionis D, & Oksuzyan A (2018). Gender difference in cognitive health among older Indian adults: A cross-sectional multilevel analysis. SSM - Population Health, 5, 180–187. 10.1016/j.ssmph.2018.06.008 [PubMed: 30073185]

Sharma D, Salig Ram M, & Anupam P (2013). Prevalence of cognitive impairment and related factors among elderly: A population-based study. Journal of the Dr. NTR University of Health Sciences, 2(3), 171–176.

Sloan FA, & Wang J (2005). Disparities among older adults in measures of cognitive function by race or ethnicity. Journals of Gerontology, Psychological Sciences and Social Sciences, 60(5), p242–p250.

Tuokko H, Chou P, Bowden S, Simard M, Ska B & Crossley M (2009). Partial measurement equivalence of French and English version of the Canadian Study of Health and Aging neuropsychological battery. Journal of the International Neuropsychological Society, 15(3), 416–425. [PubMed: 19402928]

Vandenberg RJ & Lance CE (2000). A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. Organizational Research Methods 3(1), 4–70.

Vandenberg SG & Kuse AR (1978). Mental rotations, a group test of three-dimensional spatial visualization. Perceptual and Motor Skills, 47(2), 599–604. [PubMed: 724398]

Wechsler D (1997a). Administration and scoring manual: Wechsler Memory Scale. San Antonio, TX: Harcourt Brace.

Wechsler D (1997b). Administration and scoring manual: Wechsler Adult Intelligence Scale. San Antonio, TX: Harcourt Brace.World Health Organization (2012). Dementia: a public health priority. World Health Organization Press, United Kingdom

Weden MM, Shih RA, Kabeto MU, & Langa KM (2018). Secular Trends in Dementia and Cognitive Impairment of U.S. Rural and Urban Older Adults. American Journal of Preventive Medicine, 54(2), 164–172. 10.1016/j.amepre.2017.10.021 [PubMed: 29246677]

World Health Organization (2017a). Dementia - Fact Sheets. Available at: http://www.who.int/news-room/fact-sheets/detail/dementia

World Health Organization (2017b). Global Action Plan on the Public Health Response to Dementia 2017-2025. Geneva: WHO.

Workman JE, & Lee S-H. (2004). A cross-cultural comparison of the apparel spatial visualization test and paper folding test. Clothing and Textiles Research, 22(1-2), 22–30.

Zahodne LB, Manly JJ, Azar M, Brickman AM, & Glymour MM (2016). Racial Disparities in Cognitive Performance across Mid and Late Adulthood: Analyses in Two Cohort Studies. Journal of the American Geriatrics Society, 64(5), 959–964. 10.1111/jgs.14113 [PubMed: 27225354]

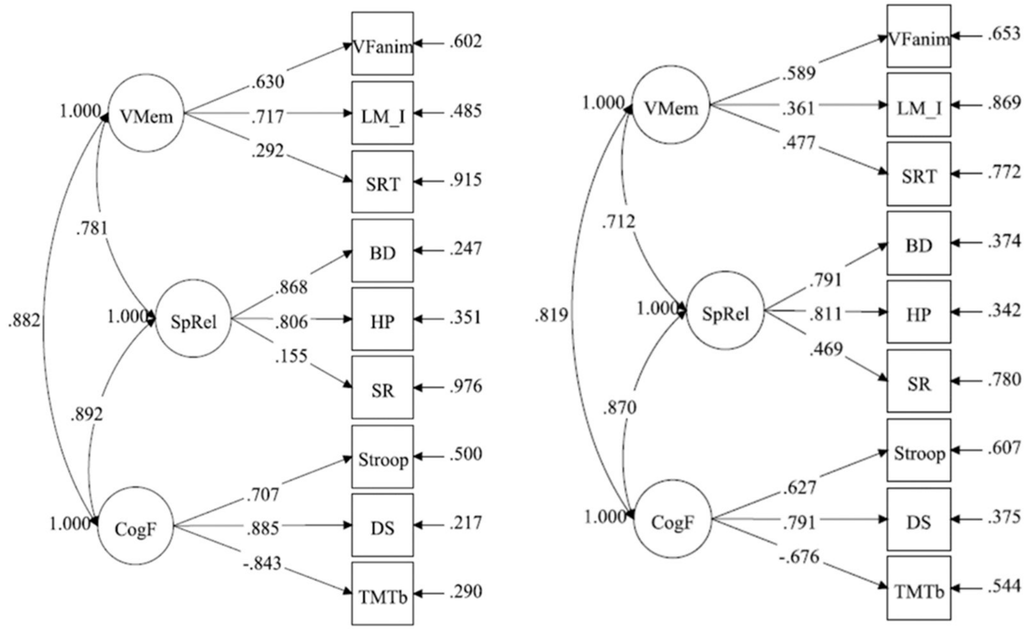

Figure 1. CFA baseline model with three cognitive latent factors and nine neuropsychological tests.

a. Rural sample, b. Urban sample

Note. Latent variables: VMem = Verbal memory; SpRel = Spatial relations; CogF = Cognitive flexibility. Neuropsychological tests: VFanim = Verbal fluency animals; LM_I = Logical memory immediate recall; SRT = Selective reminding test total score; BD = Block Design; HP = Hidden Pattern; SR = Spatial relations; Stroop = Stroop Color-word interference; DS = Digit Symbol; TMTb = Trail Making Test b time.

Table 1.

Demographic Characteristics of the EDAD Sample

Sample	Urban (N = 181)	Rural[a] (N = 114)
Characteristic, mean (SD)		
Age, years	67.6 (5.54)	70.51 (6.34)**
Education	13.96 (6.52)	10.2 (5.57)**
Female (%)	136 (75.1%)	91 (79.8%)[b]
MMSE	28.98 (1.45)	27.70 (1.92)**

[a] Asterisks indicate p-values for t-test comparisons of urban vs. rural.

[b] The X^2 statistic was used to compare the number of female vs. male participants by region, $X^2 = .87$, $p > .05$

* $p < .05$.

** $p < .001$

Table 2

Descriptive Statistics (means, SD) of the Neuropsychological Measures in the Baseline Model Across Urban/Rural Groups

Sample	Urban (N = 181)	Rural[a] (N = 114)	Skewness (SE)	Kurtosis (SE)	Shapiro-Wilk p-value	
					Urban	Rural
LM – I	9.88 (3.24)	8.34 (2.88) **	.587 (.142)	.725 (.283)	.001	< .001
VF – Ani	19.48 (4.35)	18.57 (5.10)	.107 (.142)	−.228 (.284)	.153	.159
SRT	46.81 (2.30)	46.99 (2.43)	−1.724 (.142)	7.095 (.283)	< .001	< .001
TMT-B	140.09 (68.9)	172.8 (67.7) **	.884 (.153)	.079 (.306)	< .001	< .001
DS	35.70 (10.18)	24.23 (11.4) **	−.005 (.142)	−.593 (.283)	.314	.012
BD	26.17 (10.22)	19.20 (9.29) **	.663 (.142)	.691 (.284)	< .001	.007
Stroop	29.91 (9.71)	23.78 (9.18) **	.774 (.142)	2.919 (.284)	< .001	.020
SR	8.74 (3.56)	6.75 (2.95) **	.312 (.144)	−.244 (.287)	.013	.005
HP	45.82 (23.53)	32.43 (22.3) **	.854 (.145)	.880 (.288)	< .001	< .001

* p <.05.

** p <.001

Table 3

Statistics and Indices for Invariance Models Across Urban/Rural Groups

Invariance model	χ^2	*df*	*p*	RMSEA [CI 90%]	TLI	CFI	$\Delta\chi^2$ *(df, p)*	ΔRMSEA	ΔTLI	ΔCFI	?
Measurement Invariance Model											
Model 1 *Configural*	56.30	48	0.192	0.034 [0.000-0.066]	0.984	0.989	-	-	-	-	YES
Model 2 *Full Metric*	71.55	54	0.055	0.047 [0.000-0.074]	0.970	0.978	15.25 (6, p < 0.019)	0.013	0.014	0.011	NO[a]
Model 2.1 *Partial metric*	60.60	52	0.193	0.033 [0.000-0.065]	0.985	0.989	4.3 (2, p < 0.116)	0.001	0.001	0.000	YES[b]
Model 3 *Full Scalar*	96.45	59	0.002	0.066 [0.041-0.089]	0.942	0.952	35.85 (7, p < 0.001)	0.033	0.043	0.037	NO[a]
Model 3.1 *Partial scalar*	71.22	57	0.098	0.041 [0.000-0.069]	0.977	0.982	10.62 (5, p = 0.06)	0.008	0.008	0.007	YES[c]
Structural Invariance Model											
Model 4 *Variance Invariance*	77.77	60	0.061	0.045 [0.000-0.071]	0.973	0.977	6.55 (3, p = 0.088)	0.004	0.004	0.005	YES[d]
Model 5 *Mean Invariance*	155.9	63	0.001	0.100 [0.080-0.120]	0.865	0.882	78.1 (3, p < 0.001)	0.055	0.108	0.095	NO[e]
Model 5.1 *VMem Mean invariance*	79.10	61	0.06	0.045 [0.000-0.071]	0.973	0.977	1.33 (1, p = 0.249)	0.000	0.000	0.000	YES[f]

Note. ? = Can measurement invariance be established? Model 5.1. constrained Verbal Memory only.

[a]Model fit is not acceptable, invariance cannot be established.

[b]Model fit for the partial metric model is acceptable, invariance can be established. The partial metric model fits significantly better than the full metric model. When compared to the full metric model, model 2.1 improved its model fit, $\Delta\chi^2$ (df = 2) = 10.95, p < 0.004 and increased its fit indices (ΔRMSEA = 0.014, ΔTLI = 0.015; ΔCFI = 0.011).

[c]Model fit for the partial scalar model is acceptable, invariance is established. The partial scalar model fits significantly better than the full scalar model. When compared to the full scalar model, model 3.1 improved its model fit, $\Delta\chi^2$ (df = 2) = 25.23, p < 0.001 and fit indices (ΔRMSEA = 0.025, ΔTLI = 0.035; ΔCFI = 0.03).

[d]Model fit for the structural variance invariance model is acceptable, invariance is established.

[e]The mean invariance model is significant. It cannot assume that latent means are equal across groups.

[f]Model fit for the Verbal Memory mean invariance model is acceptable, invariance holds. The model fits significantly better than the full mean invariance model. When compared to the full mean invariance model, model 5.1 improved its model fit, $\Delta\chi^2$ (df = 2) = 76.77, p < 0.001 and fit indices (ΔRMSEA = 0.055, ΔTLI = 0.108; ΔCFI = 0.095).

Table 4

Statistics and Indices for Invariance Models Across Urban/Rural Groups with Age and Education as Covariates

Invariance model	χ^2	***df***	***p***	**RMSEA [CI 90%]**	**TLI**	**CFI**	**$\Delta\chi^2$ *(df, p)***	**ΔRMSEA**	**ΔTLI**	**ΔCFI**	**?**
Measurement Invariance Model											
Model 1 *Configural*	82.52	72	0.19	0.032 [0.000-0.060]	0.982	0.988	-	-	-	-	YES
Model 2 *Full Metric*	99.25	78	0.053	0.043 [0.000-0.067]	0.967	0.976	16.73 (6, p < 0.01)	0.011	0.015	0.012	NO[a]
Model 2.1 *Partial metric*	88.00	76	0.16	0.033 [0.000-0.060]	0.981	0.987	5.478 (4, p < 0.24)	0.001	0.001	0.001	YES[b]
Model 3 *Full Scalar*	126.2	83	0.002	0.060 [0.037-0.080]	0.938	0.952	38.21 (7, p < .0001)	0.027	0.043	0.035	NO[a]
Model 3.1 *Partial scalar*	96.21	81	0.12	0.036 [0.000-0.061]	0.978	0.983	8.21 (5, p = .145)	0.003	0.003	0.004	YES[c]
Structural Invariance Model											
Model 4 *Variance Invariance*	90.47	83	0.27	0.024 [0.000-0.054]	0.989	0.992	5.74 (2, p = 0.06)	0.012	0.011	0.009	YES[d]
Model 5 *Mean Invariance*	96.55	86	0.205	0.029 [0.000-0.056]	0.985	0.988	6.07 (3, p = 0.108)	0.005	0.004	0.004	YES[e]

Note. All measurement (full and partial) and structural models include age and education as covariates. ? = Can measurement invariance be established?

[a]Model fit is not acceptable, invariance does not hold.

[b]Model fit for the partial metric model is acceptable, invariance holds. The partial metric model fits significantly better than the full metric model. When compared to the full metric model, model 2.1 improved its model fit, $\Delta\chi^2$ (df = 2) = 11.25, p = 0.004 and increased its fit indices (ΔRMSEA = 0.010, ΔTLI = 0.014; ΔCFI = 0.011).

[c]Model fit for the partial scalar model is acceptable, invariance holds. The partial scalar model fits significantly better than the full scalar model. When compared to the full scalar model, model 3.1 improved its model fit, $\Delta\chi^2$ (df = 2) = 11.25, p < 0.001 and fit indices (ΔRMSEA = 0.024, ΔTLI = 0.040; ΔCFI = 0.031).

[d]Model fit for the structural variance invariance model is acceptable, invariance holds.

[e]The mean invariance model is not significant. It assumes that latent means are equal across groups. Latent mean invariance across regions is established.

Table 5

Factor Loadings, Intercepts (standard errors in parentheses) and Effect Sizes (R^2) from the Final Structural Invariance Model Tested Across Regions, with Age and Education as Covariates

Latent variable	Observed variable	Urban			Rural		
		λ (SE)	**τ (SE)**	**R^2**	**λ (SE)**	**τ (SE)**	**R^2**
Verbal Memory	LM-I	0.472 (0.056)	**4.208 (0.513)**	0.22	0.676 (0.065)	4.425 (0.595)	0.46
	VF-anim	0.609 (0.072)	5.970 (0.667)	0.37	0.630 (0.061)	4.974 (0.571)	0.41
	SRT	0.308 (0.062)	21.826 (1.086)	0.16	0.353 (0.068)	19.858 (1.297)	0.12
	α_{VMem}		0.000			0.000	
	ψ_{VMem}		1.000			1.000	
Spatial Reasoning	BD	0.769 (0.036)	4.427 (0.624)	0.59	0.890 (0.030)	4.364 (0.623)	0.79
	HP	0.768 (0.037)	3.864 (0.629)	0.59	0.840 (0.033)	3.600 (0.593)	0.71
	SR	0.448 (0.064)	3.597 (0.406)	0.20	**0.115^{ns}(0.116)**	**2.596(0.358)**	0.01^{ns}
	α_{SpR}		0.000			0.000	
	ψ_{SpR}		1.000			1.000	
Cognitive Flexibility	Stroop	0.526 (0.043)	4.960 (0.447)	0.28	0.722 (0.041)	4.790 (0.441)	0.52
	DS	0.803 (0.033)	6.112 (0.624)	0.65	0.933 (0.020)	4.988 (0.546)	0.87
	TMT-b	−0.651 (0.042)	0.052^{ns} (0.510)	0.42	−0.835 (0.036)	0.047^{ns} (0.460)	0.70
	α_{CogF}		0.000			0.000	
	ψ_{CogF}		1.000			1.000	

Note. The final structural model was the mean invariance model with all latent means constrained equal across groups. Factor loadings (λ) and intercepts (τ) that were freely estimated are bolded. All latent variances (ψ) were set to 1.0 as they were tested to be invariant.

^{ns}Not significant parameters ($p > 0.05$). All other parameters were significant at $p < .001$

Table 6

Regression Coefficients (SE) and Effect Sizes (R^2) of Latent Cognitive Constructs on Age and Education, by Urban/Rural Region.

	Age		**Education**		**R^2**	
	Urban	**Rural**	**Urban**	**Rural**	**Urban**	**Rural**
Verbal Memory	−0.239 (0.080)	−0.237 (0.109)	0.283 (0.098)	0.562 (0.098)	0.152	0.464
Spatial Reasoning	−0.233 (0.063)	−0.212 (0.091)	0.247 (0.078)	0.551 (0.079)	0.128	0.368
Cognitive Flexibility	−0.324 (0.058)	−0.272 (0.068)	0.374 (0.070)	0.688 (0.055)	0.272	0.640

Note. All regression coefficients were significant at a p-value < .05. The effect sizes of each latent cognitive construct include age and education, as they were included simultaneously in the model.