

## Tilburg University

### Traveling with cognitive tests

Malda, M.; van de Vijver, F.J.R.; Srinivasan, K.; Sukumar, P.

*Published in:*  
Assessment

*Publication date:*  
2010

*Document Version*  
Publisher's PDF, also known as Version of record

[Link to publication in Tilburg University Research Portal](#)

*Citation for published version (APA):*  
Malda, M., van de Vijver, F. J. R., Srinivasan, K., & Sukumar, P. (2010). Traveling with cognitive tests: Testing the validity of a KABC-II adaptation in India. *Assessment*, 17, 107-115.

#### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

#### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Assessment

<http://asm.sagepub.com/>

---

## **Traveling With Cognitive Tests: Testing the Validity of a KABC-II Adaptation in India**

Maike Malda, Fons J. R. van de Vijver, Krishnamachari Srinivasan, Catherine Transler and Prathima Sukumar

*Assessment* 2010 17: 107 originally published online 10 September 2009

DOI: 10.1177/1073191109341445

The online version of this article can be found at:

<http://asm.sagepub.com/content/17/1/107>

---

Published by:



<http://www.sagepublications.com>

**Additional services and information for *Assessment* can be found at:**

**Email Alerts:** <http://asm.sagepub.com/cgi/alerts>


**Subscriptions:** <http://asm.sagepub.com/subscriptions>

**Reprints:** <http://www.sagepub.com/journalsReprints.nav>

**Permissions:** <http://www.sagepub.com/journalsPermissions.nav>

**Citations:** <http://asm.sagepub.com/content/17/1/107.refs.html>

# Traveling With Cognitive Tests: Testing the Validity of a KABC-II Adaptation in India

Assessment  
17(1) 107–115  
© The Author(s) 2010  
Reprints and permission: <http://www.sagepub.com/journalsPermissions.nav>  
DOI: 10.1177/1073191109341445  
<http://asmnt.sagepub.com>  


Maike Malda<sup>1</sup>, Fons J. R. van de Vijver<sup>2</sup>, Krishnamachari Srinivasan<sup>3</sup>,  
Catherine Transler<sup>4</sup> and Prathima Sukumar<sup>5</sup>

## Abstract

The authors evaluated the adequacy of an extensive adaptation of the American Kaufman Assessment Battery for Children, second edition (KABC-II), for 6- to 10-year-old Kannada-speaking children of low socioeconomic status in Bangalore, South India. The adapted KABC-II was administered to 598 children. Subtests showed high reliabilities, the Cattell–Horn–Carroll model underlying the original KABC-II was largely replicated, and external relations with demographic characteristics and an achievement measure were consistent with expectations. The subtests showed relatively high loadings on the general cognitive factor, presumably because of the high task novelty and, hence, cognitive complexity of the tests for the children. The findings support the suitability and validity of the KABC-II adaptation. The authors emphasize that test adaptations can only be adequate if they meet both judgmental (qualitative) and statistical (quantitative) adaptation criteria.

## Keywords

Kaufman Assessment Battery for Children, cognitive test, adaptation, bias, culture, validity

Many tests that originate from the United States or Europe have been used extensively in developing and emerging countries without practically and scientifically examining the suitability of these instruments outside their country or culture of origin (Misra, Sahoo, & Puhan, 1997). The use of the original or closely translated instruments saves costs and time; yet optimizing an instrument for a specific cultural context implies the need for an adaptation, in which cultural knowledge, values, and practices are taken into account (see Abubakar et al., 2007; Holding et al., 2004). A close translation may then not be sufficient. The validity of adapted instruments cannot be inferred from the original Western instruments and has to be demonstrated in the new cultural context. Adaptations of the Kaufman Assessment Battery for Children (K-ABC; A. S. Kaufman & Kaufman, 1983) have been described before (e.g., Holding et al., 2004; Moon, McLean, & Kaufman, 2003); however, its successor, the Kaufman Assessment Battery for Children, second edition (KABC-II), which differs from the first version in several ways (see A. S. Kaufman & Kaufman, 2004), has not been adapted. The KABC-II is an individually administered measure of cognitive ability that can be used for children from 3 to 18 years of age. It measures long-term storage and retrieval, short-term memory, visual processing, fluid reasoning, and crystallized ability. In a previous report, we described the extensive adaptation of the KABC-II for use among 6- to 10-year-old Kannada-speaking children of low socioeconomic

status in Bangalore, South India (Malda et al., 2008). The present study statistically tested the adequacy of the resulting instrument by examining its reliability and validity.

## Quality Criteria

The use of Western cognitive instruments in non-Western contexts may lead to bias, which refers to factors that may make direct comparisons of test constructs or scores between groups invalid (e.g., Van de Vijver & Tanzer, 2004). Bias makes culture-fair testing impossible (Verney, Granholm, Marshall, Malcarne, & Saccuzzo, 2005). Instrument adaptation has been proposed as a strategy to reduce bias, optimize the suitability of an instrument for a cultural context, and facilitate cross-cultural test transfer (e.g., Van de Vijver &

<sup>1</sup>Tilburg University

<sup>2</sup>Tilburg University

<sup>3</sup>North-West University

<sup>4</sup>St. John's Research Institute

<sup>5</sup>Unilever Food and Health Research Institute

<sup>5</sup>St. John's Research Institute

## Corresponding Authors:

Please address correspondence to Maike Malda, Department of Psychology, Tilburg University, PO Box 90153, 5000 LE Tilburg, Netherlands  
E-mail: [maikemalda@gmail.com](mailto:maikemalda@gmail.com).

Poortinga, 2005). A properly adapted cognitive instrument meets both qualitative and quantitative criteria.

The main qualitative criteria involve cultural and linguistic appropriateness of the instrument in the target context. The judgmental procedures that were applied to meet these criteria in the Indian adaptation of the KABC-II were described before (see Malda et al., 2008). Quantitative criteria to assess the quality of an adaptation include the instrument's reliability and validity. Various criteria have been proposed for demonstrating the construct validity of an instrument (Messick, 1989). In the absence of cross-cultural comparative data, our validity test was based on three criteria that examine whether theoretical expectations are borne out. First, the underlying theoretical model should be well represented in the data; second, relations of test scores with background characteristics such as children's age and sex should be according to expectations; third, relations of test scores with presumably related psychological constructs (such as scholastic achievement) should be according to expectations. The current study tested whether the adapted version of the KABC-II meets these three criteria that are described below in more detail. Other criteria for construct validity, such as a test's predictive validity, were not addressed.

### *Generalizability of Cognitive Structure*

The KABC-II is based on the Cattell–Horn–Carroll (CHC) model of cognitive abilities (A. S. Kaufman & Kaufman, 2004; see also Reynolds, Keith, Goldenring-Fine, Fisher, & Low, 2007). The CHC theory integrates the Cattell–Horn Gf-Gc theory (distinguishing between various fluid and crystallized cognitive abilities) and Carroll's three-stratum theory (Carroll, 1993; McGrew, 2005). The construct validity of our adaptation is statistically supported if our Indian data confirm the original CHC structure, assuming that the model can be generalized to non-Western groups. The generalizability of the CHC model has been shown with exploratory as well as confirmatory factor analyses across age (Bickley, Keith, & Wolfle, 1995; Taub & McGrew, 2004) and sex (Reynolds, Keith, Ridley, & Patel, 2008). Furthermore, the CHC structure is found with many cognitive test batteries even when these were not originally designed to represent this structure (for an overview, see McGrew, 2005). Is the CHC model, besides being generalizable across ages, sexes, and tests, also generalizable across cultures?

Models underlying cognitive test batteries have shown cross-cultural stability (Berry, Poortinga, Segall, & Dasen, 2002; Irvine, 1979; Van de Vijver, 1997). A large study was conducted with culturally adapted versions of the Wechsler Intelligence Scale for Children—third edition (WISC-III; Georgas, Van de Vijver, Weiss, & Saklofske, 2003). Using exploratory factor analyses, it was found that the cross-cultural equivalence of the underlying structure was high. The K-ABC (the predecessor of the KABC-II), based on the

sequential versus simultaneous processing distinction, was applied in many non-Western countries (e.g., Boivin et al., 1996; Conant et al., 1999; Jansen & Greenop, 2008; Mardell-Czudnowski, 1995). Holding et al. (2004) and Moon et al. (2003) found the underlying model to be present in adapted versions in Kenya and Korea, respectively. Not many studies on the CHC model have been conducted in non-Western contexts; still, there is no reason to doubt the universality of the structure of a well-established cognitive model. We consider the CHC model to be a good starting point for statistically evaluating the validity of our KABC-II adaptation.

### *Age and Sex Effects*

Test scores are expected to increase with age in our study sample. Although there tends to be considerable overlap between male and female cognitive test score distributions (Born, Bleichrodt, & Van der Flier, 1987; Fairweather, 1976), males generally score higher on tests measuring visual abilities and mathematical reasoning, whereas females do better on verbal (memory) tasks and numerical calculation (e.g., Reynolds et al., 2008). Sex differences on the original KABC-II are small for school-age children (7-18 years); boys perform better on the visual processing tasks and girls on the learning and fluid reasoning tasks (A. S. Kaufman & Kaufman, 2004).

### *Relations With School Performance: Arithmetic*

Various mechanisms behind the positive relation between cognitive abilities and arithmetic skills in middle childhood have been proposed, such as phonological and/or visuospatial memory, speed of processing, number processing, and spatial and nonverbal ability (Durand, Hulme, Larkin, & Snowling, 2005). The mechanisms are likely to vary across ages and tasks. Correlations between broad ability factors of the original KABC-II and arithmetic scores of other cognitive batteries (A. S. Kaufman & Kaufman, 2004) show that for the younger age group (Grades 2-5) the highest correlations were found for the fluid reasoning factor, possibly because arithmetic processes are not yet differentiated and automated, and hence, their solution still requires complex, integrated cognitive abilities. We expect similar findings in the present study.

### *Hypotheses*

Our study adds to the literature in that it (a) evaluates the validity of an adaptation of the relatively new KABC-II in a non-Western context and (b) examines the CHC model in this non-Western (Indian) context to accomplish this. As a prerequisite for any hypothesis testing in a research context (rather than a clinical context), the internal consistencies of

the subtests should be at least .70 (Cicchetti et al., 2006). The appropriateness of our adaptation is tested using the following hypotheses:

#### Theoretical structure

*Hypothesis 1:* The factor structure of the KABC-II adaptation is in line with the CHC model.

#### Psychometric properties

*Hypothesis 2:* The underlying cognitive structure is similar across sexes (Hypothesis 2a) and ages (Hypothesis 2b).

*Hypothesis 3:* Test scores increase significantly with age.

*Hypothesis 4:* If sex differences in scores are found, boys outperform girls on visual processing tasks, and girls outperform boys on fluid reasoning and learning tasks.

*Hypothesis 5:* All broad ability factors correlate significantly with arithmetic scores.

*Hypothesis 6:* Arithmetic scores show the highest correlations with the fluid reasoning factor.

## Method

### Participants and Study Context

The sample included 598 Kannada-speaking children (293 boys and 305 girls) of low socioeconomic status in Bangalore (state of Karnataka, South India). The children were between 6 and 10 years old ( $M = 8.71$ ,  $SD = 1.17$ ) and from Grade 2 to 5 of two primary schools ( $N = 370$  and  $N = 228$ , respectively).

The children in our sample came from poor families with an average monthly income of 2,700 Indian rupees (US\$56). Many adults were illiterate or had only a few years of education. Houses were crowded; most had one or two rooms, and the average number of people in a household was 5.81 (2.71 adults and 3.11 children). Children had very few toys to play with and very limited access to books. Rote learning was widely used in the schools of the study; it is a commonly applied method in Indian education (Mishra, 1997), which is well applicable with large numbers of children and with a collectivistic style of teaching because the children do not need to be addressed individually.

### Instruments

*Kaufman Assessment Battery for Children, second edition.* An adapted version of eight core subtests of the KABC-II was administered; a description of the extensive adaptation procedure can be found in Malda et al. (2008). Here, we confine ourselves to describing the abilities measured (see also

Carroll, 1993; A. S. Kaufman & Kaufman, 2004; J. C. Kaufman, Kaufman, Kaufman-Singer, & Kaufman, 2005; McGrew, 2005). Two subtests were selected to assess the broad ability of fluid reasoning. Pattern Reasoning measures inductive reasoning and visualization, and our Indian version of WISC-R-III Picture Arrangement (which is a replacement of the KABC-II subtest Story Completion) mainly measures pattern recognition, reasoning, and planning (Wechsler, 1949, 1974, 1991). Two subtests were selected to cover the broad ability factor of short-term memory, namely, Number Recall and Word Order, both measuring memory span. The subtests Rover and Triangles were assumed to reflect the broad ability of visual processing. Rover is a measure of spatial scanning, general sequential or deductive reasoning, and math achievement; Triangles measures spatial relations and visualization. For the broad ability factor of long-term storage and retrieval, Atlantis was selected, which is a measure of associative memory. The second test was an auditory Verbal Learning Test, which was a replacement of the KABC-II subtest Rebus. In this study, we use the recall score of the Verbal Learning Test: the number of words (out of 15) correctly recalled after a 20-minute delay. For the purposes of the larger study to which our study contributes, another subtest that is not part of the KABC-II battery was added, namely, Verbal Fluency (a measure of associational fluency). This addition aimed to ensure proper coverage of long-term storage and retrieval processes. Also, two subtests reflecting the broad ability of cognitive speediness were included, namely, Coding B and Number Cancellation; cognitive speediness was not covered by the core subtests of the KABC-II but is part of the CHC model. Coding B is taken from the Wechsler scales and is (mainly) a measure of attention and concentration; Number Cancellation measures perceptual speed and more specifically scanning (McGrew, 2005).

*Arithmetic test.* Measures of crystallized abilities and school achievement (such as reading, spelling, and arithmetic tasks) provide an external validation criterion for our adapted test. The only available local test of school achievement that seemed sensitive enough to discriminate between children within the same school grade was the Arithmetic Diagnostic Test for Primary School Children (Ramaa, 1994). We based our arithmetic test on this instrument and on information about the curriculum of the schools included in the current study. A test consisting of two parts was composed: 36 addition and 34 subtraction items of increasing difficulty. The sum score of these two parts was used in the analyses.

### Procedure

All children and their parents gave consent for participation in the study, in accordance with Indian ethical rules. Seven test examiners were trained jointly by a Dutch psychologist (first author) and a local psychologist (fifth author), after

which the study commenced. Each test examiner assessed two children every day; the administration was split up in three test sessions (two sessions of 30 minutes and one of 45 minutes). All administered the tests to more or less the same number of girls and boys and to children of all grades. The arithmetic test was administered to batches of about 60 children who had all received KABC-II testing in the same week.

## Analyses

First, reliabilities were calculated by the split-half technique, Cronbach's alpha, or correlations, depending on the characteristics of the subtests. Second, structural equation modeling in Amos 6 (Arbuckle, 2005) was used to test the validity of the CHC model for the study sample (Hypothesis 1), followed by multigroup analyses to test for equivalence of the model across sexes (Hypothesis 2a) and ages (Hypothesis 2b). Multivariate Analyses of Variance (MANOVAs) were conducted to test for the effects of age (Hypothesis 3) and sex (Hypothesis 4) on cognitive test scores. Last, Hypotheses 5 and 6 (dealing with the arithmetic test) were addressed by correlating broad ability scores with arithmetic scores, using Pearson correlation coefficients.

## Results

### Reliability

For the subtests Atlantis, Rover, Number Recall, Pattern Reasoning, Word Order, Triangles, and Picture Arrangement, the internal consistency was measured by the split-half technique. Values of Cronbach's alpha could not be computed because of the discontinuation rules of these subtests. For each subtest, the sum scores of the odd and even items were correlated and the Spearman-Brown formula (Thurstone, 1931) was applied to adjust this reliability estimate for test length. Reliabilities of our adapted subtests were acceptable to very good (Pattern Reasoning, .94; Picture Arrangement, .72; Number Recall, .70; Word Order, .82; Triangles, .89; Rover, .90; Atlantis, .96) and largely in accordance with the reliabilities of the original KABC-II.

The Verbal Learning Test comprised a 15-word list that was read out loud to the child, after which immediate recall was measured. This procedure was repeated twice and after a 20-minute delay, recall was measured for the fourth time. Cronbach's alpha was calculated for the number of correctly recalled words in each of these four trials; the median alpha across the five age groups (i.e., ages 6-10) was high: .84 (range .75-.86). In further analyses, we only used the recall score of the Verbal Learning Test (i.e., number of correctly recalled words after a 20-minute delay).

For Verbal Fluency, Number Cancellation, and Coding B, an indication of reliability could only be obtained by item

or test correlations because these tests consist of one or two items. All correlations were controlled for age. The Verbal Fluency test first required the children to call out as many animals as possible, and then as many first names as possible. The correlation between the two numbers was positive and significant,  $r(598) = .31$ ,  $p < .01$ , according to our expectations. The value was not very high, presumably because most children named their classmates one by one when they generated first names, but they did not use a common strategy in generating animal names. The correlation between Number Cancellation time and Coding B was  $r(598) = -.45$  ( $p < .01$ ), indicating that the faster the child finished the Number Cancellation task, the more correct items were obtained on Coding B.

### Validity

*CHC model.* Structural equation modeling was used to test the validity of the CHC model. The subtests (i.e., specific abilities) were expected to cover five broad abilities, namely, fluid reasoning, short-term memory, visual processing, long-term storage and retrieval, and cognitive speediness. A general cognitive ability factor (called Mental Processing Index for the KABC-II) was expected to underlie these five factors. The fit of the original CHC model (Model 1) to our data was acceptable, however, the modification indices suggested two improvements: (a) linking Verbal Fluency to the cognitive speediness factor rather than the long-term storage and retrieval factor (because Verbal Fluency also involves speed) and (b) combining the visual processing subtests with the fluid reasoning subtests in one factor (because all these subtests involve aspects of reasoning). We tested the first alternative (Model 2), then the second alternative (Model 3), followed by a combination of the two (Model 4). Fit statistics are presented in Table 1. Differences between the fit of the models were small; however, Table 1 suggests that Model 3 showed a slightly better fit than the other models (particularly given its relatively low Akaike's information criterion value). Model 3 is displayed in Figure 1.

There is considerable overlap between the structures of the CHC model and our model; most important, the hierarchy of cognitive abilities is supported. Most expected factors were found: a short-term memory factor, a long-term storage and retrieval factor, and a cognitive speediness factor. Our final model differed from the CHC model in two ways. First, Triangles and Rover, hypothesized to represent a separate visual processing factor, loaded on the same factor as Pattern Reasoning and Picture Arrangement. Second, the patterning of the loadings of the broad abilities on general cognitive functioning is different from the CHC model. This model would predict that fluid reasoning has the highest loading on the general cognitive factor, followed

**Table 1.** Summary Statistics for Various Structural Equation Models Related to the CHC Model

Model	$\chi^2$	df	$\chi^2/df$	<i>p</i>	CFI	RMSEA	AIC
1. CHC	122.711	39	3.146	.001	.941	.060	176.711
2. VF to CS	124.312	39	3.187	.001	.940	.061	178.312
3. <i>VP with FR</i>	<i>103.792</i>	40	<i>2.595</i>	<i>.001</i>	<i>.955</i>	<i>.052</i>	<i>155.792</i>
4. VF to CS and VP with FR	111.406	40	2.785	.001	.950	.055	163.406

Note: Preferred model is in italics. CHC = Cattell–Horn–Carroll; VF = Verbal Fluency; CS = Cognitive Speediness; VP = Visual Processing; FR = Fluid Reasoning; CFI = comparative fit index; RMSEA = root mean square error of approximation; AIC = Akaike's information criterion.

by short-term memory, long-term storage and retrieval, and cognitive speediness. However, the long-term storage and retrieval factor showed the highest loading in our study, followed by fluid reasoning, short-term memory, and cognitive speediness. We estimated the confidence intervals of the factor loadings using a bootstrapping procedure and found that the confidence interval of the long-term storage and retrieval loading did not overlap with the confidence intervals of the other loadings. This indicated that only the long-term storage and retrieval loading differed significantly from the others ( $p < .05$ ).

**Sex and age.** Multigroup analysis was applied to test for equivalence of our final model across sexes and ages. Sex and age could not be combined in one single multigroup analysis because the children were not equally distributed across all possible combinations of sex and age; therefore, we tested multigroup invariance of the CHC model separately for these two variables. A good fit was found for a model in which all parameters were constrained to be equal for boys and girls,  $\chi^2(106, N = 598) = 213.64, p < .01, \chi^2/df = 2.02$ , goodness-of-fit index (GFI) = .94, adjusted goodness-of-fit index (AGFI) = .93, Tucker–Lewis index (TLI) = .95, comparative fit index (CFI) = .95, root mean square error of approximation (RMSEA) = .04). The model testing the invariance across age groups (age groups 6 and 7 are taken together because of the relatively small number of children with age 6) showed a good fit when all parameters were identical, except for equality of measurement residuals,  $\chi^2(205, N = 598) = 298.58, p < .01, \chi^2/df = 1.46$ , GFI = .92, AGFI = .89, TLI = .93, CFI = .94, RMSEA = .03). Widaman and Reise (1997) argue that the latter constraint is rather irrelevant.

A MANOVA was computed with sex, Wilks's  $\lambda = .88, F(11, 578) = 7.23, p < .01$ , partial  $\eta^2 = .12$ , and age group, Wilks's  $\lambda = .57, F(44, 2213.24) = 8.05, p < .01$ , partial  $\eta^2 = .13$ , as independent variables and the sum scores for all cognitive tests as dependent variables. There was no significant interaction between sex and age group, Wilks's  $\lambda = .92, F(44, 2213.24) = 1.10, p = .31$ , partial  $\eta^2 = .02$ . Girls scored significantly higher than boys on Verbal Learning recall,  $F(1, 588) = 18.82, p < .01, d = .42$ , Number Cancellation time,  $F(1, 588) = 36.04, p < .01, d = -.47$ , Coding,  $F(1, 588) = 13.90, p < .01, d = .38$ , and Verbal

Fluency,  $F(1, 588) = 8.34, p < .01, d = .30$ . Boys outperformed girls on Rover,  $F(1, 588) = 5.31, p < .05, d = -.18$ . Scores on all subtests increased significantly with age ( $p < .01$ ). Age explained more variance on the subtests measuring fluid reasoning and cognitive speediness as compared with the subtests measuring short-term memory and retrieval ability. For example, for Rover, partial  $\eta^2$  was .17,  $F(4, 588) = 30.61, p < .01$ , whereas for Number Recall, the value was .04,  $F(4, 588) = 6.05, p < .01$ . These findings indicate that children's reasoning and speed abilities are more age dependent than the other abilities.

**Arithmetic test.** The reliability of the arithmetic test was very high; the median value of Cronbach's alpha for the five age groups was .95 (range .93–.96). The sum scores on the arithmetic test correlated significantly with the broad ability factors, with  $r(598)$  ranging from .37 to .49,  $p < .01$ , and the Mental Processing Index,  $r(598) = .61, p < .01$ . The correlation was significantly lower ( $p < .01$ ; tested using procedures described by Dunn & Clark, 1969) for cognitive speediness than for all other factors (the difference between the cognitive speediness and short-term memory factor was bordering on significance). Arithmetic scores did not show the expected stronger correlation with the fluid reasoning factor than with the other factors.

## Discussion

The KABC-II was extensively adapted for 6- to 10-year-old Kannada-speaking children of low socioeconomic status in Bangalore, South India (Malda et al., 2008). The current study statistically evaluated the adequacy of our adaptation. Most hypotheses were confirmed. The adapted subtests showed high reliabilities; the cognitive CHC model underlying the original KABC-II was largely replicated (Hypothesis 1); the CHC model was valid across sexes (Hypothesis 2a) and age groups (Hypothesis 2b); cognitive test scores increased with age (Hypothesis 3); the small sex differences in some of the subtest scores were in line with expectations (Hypothesis 4); the arithmetic test correlated significantly with all broad ability factors (Hypothesis 5); the arithmetic sum score showed similar correlations with all but one of the broad

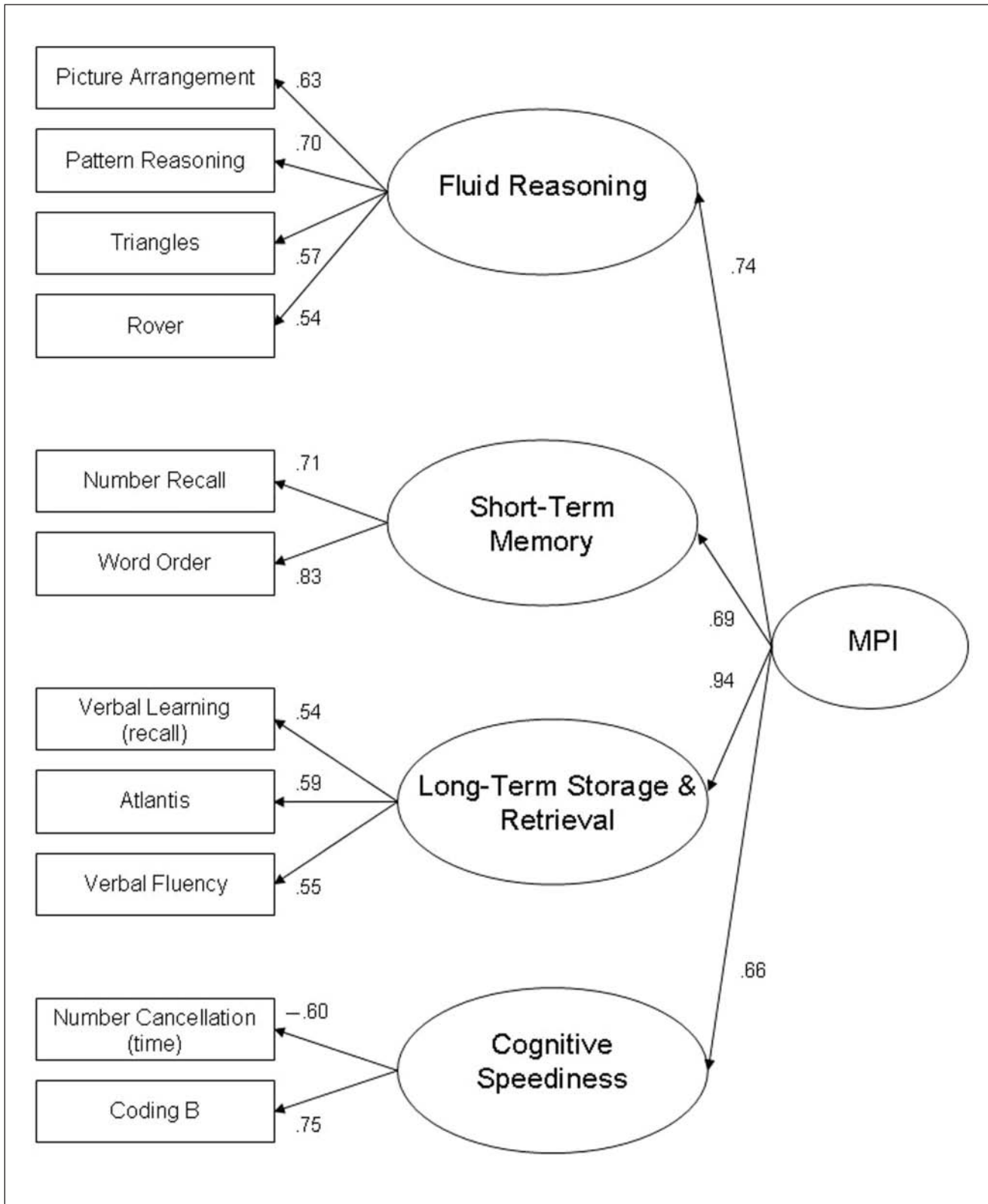


Figure 1. Structural Equation Model of All Cognitive Data



ability factors (not confirming Hypothesis 6, which predicted a higher correlation with fluid reasoning). We can conclude that our adaptation of the KABC-II is a valid cognitive measure for the target sample. This is an important finding in light of the prevalence of bias in applications of Western cognitive instruments in a non-Western context.

Our final model differed from the CHC model in two ways. First, Triangles and Rover, hypothesized to represent a separate visual processing factor, loaded on the same factor as Pattern Reasoning and Picture Arrangement. These four subtests were among the subtests with the highest intercorrelations (varying from .33 to .48;  $p < .01$ ). This finding most probably relates to a combination of the nature of the tasks (all are figural) and their complexity; fluid abilities usually depend on an integration of many distinct cognitive abilities, whereas visual processing abilities require fewer cognitive resources. However, because both the stimulus mode and the response mode of our visual processing tasks were very unfamiliar to the Indian children (despite the test adaptations that were performed to increase familiarity and suitability), it stands to reason that Triangles and Rover reflect fluid reasoning abilities rather than merely visual processing. Training children on various cognitive tasks could reduce the cognitive complexity and cause a better differentiation between tests (and factors) and a smaller reliance on a more general cognitive ability.

Second, the patterning of the loadings of the broad abilities on general cognitive functioning is different from the CHC model. Only the loading of the long-term storage and retrieval factor differed significantly from the others. The limited differentiation between the loadings may, again, be a consequence of the task unfamiliarity. The high loading of the long-term storage and retrieval factor may be caused by the diversity of the subtests that belong to this factor. The factor consists of tests that tap a wide range of abilities, in addition to their common factor. The Verbal Learning Test is a test of short- and long-term memory and learning, whereas Verbal Fluency measures free recall, and Atlantis adds a strong visual component (besides its memory and learning aspects). Because together these tests measure a broad range of abilities, it is not surprising that the loading of the long-term storage and retrieval factor on the Mental Processing Index (MPI, reflecting general cognitive functioning) is so high. The CHC model might have been replicated more closely, if we had included all (instead of merely the core) subtests of the KABC-II, providing a better coverage of the broad ability factors. Adding subtests would however have led to prohibitively long test administration times for our study sample.

Familiarity could also play a role in explaining the lack of differentiation in correlations between the broad ability factors and the arithmetic score. In the introduction, we suggested that arithmetic processes might not yet be differentiated and automated for young children, and hence, their

solution still requires complex, integrated cognitive abilities. This suggestion is in line with our expectation of a higher correlation between the arithmetic score and the (cognitively complex) fluid reasoning factor. Our findings, however, show that the arithmetic score correlates similarly to all but one of the broad ability factors (cognitive speediness), which might indicate that the arithmetic test measures an even more cognitively complex ability in our sample than anticipated. The high correlation with the MPI (.61) confirms the relevance of general cognitive processes in arithmetic performance, possibly because of the unfamiliarity of the sample with such tests and test situations.

The task unfamiliarity that we observed in our sample shows the profound influence of both home environment and educational characteristics on cognitive test performance. These Indian children of low socioeconomic status are provided with suboptimal stimulation (few play materials) at home, and the educational system is mainly focused on collective rote learning, which explains the child's lack of experience with individual test situations, and with materials such as puzzles and (geometrical) figures. This implies that issues with testing in non-Western contexts could be related to differences in socioeconomic status, in addition to cultural differences. It is possible that a sample of children from the same geographic area and same language but with high socioeconomic status would have shown a closer match with the original CHC structure compared with our present sample.

The combination of the evidence obtained in the qualitative adaptation process of the KABC-II (Malda et al., 2008) and the quantitative process discussed here supports the suitability and validity of our adaptation for Kannada-speaking children of low socioeconomic status in India; the current study offers further evidence for the generalizability of the CHC model in developing, non-Westernized countries. Both the qualitative and quantitative parts are prerequisites for ensuring an instrument's adequacy. Many studies omit a detailed test adaptation, which could lead to the use of culturally inappropriate stimuli. However, the current study shows that after an extensive qualitative adaptation process, quantitative analyses are needed to demonstrate its success. Cognitive data can only be interpreted validly when the tests meet both judgmental (qualitative) and statistical (quantitative) adaptation criteria.

### Authors' Note

We would like to thank Unilever Food and Health Research Institute (Vlaardingen, the Netherlands) for sponsoring the study. Our gratitude goes to Kamala, Sapna, Amitha, Ashwini, Krishna Veni, Puneetha, Asha, Naveena, Prathiba, Shashi Rekha, and Sri Vidya for administering the tests to the children. We thank Aruna Thyagarajan for coordinating the last phase of data collection. We are very grateful to Anura Kurpad, Sumithra Muthayya,

and Ans Eilander for their indispensable support. Last, but not the least, this study would not have been possible without the enthusiasm and dedication of the two participating schools and of all children.

## References

- Abubakar, A., Van de Vijver, F. J. R., Mithwani, S., Obiero, E., Lewa, N., Katana, K., et al. (2007). Assessing developmental outcomes in children from Kilifi, Kenya, following prophylaxis for seizures in cerebral malaria. *Journal of Health Psychology, 12*, 417-430.
- Arbuckle, J. L. (2005). Amos (Version 6.0) [Computer Program]. Chicago: SPSS.
- Berry, J. W., Poortinga, Y. H., Segall, M. H., & Dasen, P. R. (2002). *Cross-cultural psychology: Research and applications*. Cambridge, UK: Cambridge University Press.
- Bickley, P. G., Keith, T. Z., & Wolfle, L. M. (1995). The three-stratum theory of cognitive abilities: Test of the structure of intelligence across the life span. *Intelligence, 20*, 309-328.
- Boivin, M. J., Chounramany, C., Giordani, B., Xaisida, S., Choulamounry, L., Pholsena, P., et al. (1996). Validating a cognitive ability testing protocol with Lao children for community development applications. *Neuropsychology, 4*, 588-599.
- Born, M. P., Bleichrodt, N., & Van der Flier, H. (1987). Cross-cultural comparison of sex-related differences on intelligence tests. *Journal of Cross-Cultural Psychology, 18*, 283-314.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, UK: Cambridge University Press.
- Cicchetti, D., Bronen, R., Spencer, S., Haut, S., Berg, A., Oliver, P., et al. (2006). Rating scales, scales of measurement, issues of reliability: Resolving some critical issues for clinicians and researchers. *Journal of Nervous and Mental Disease, 194*, 557-564.
- Conant, L. L., Fastenau, P. S., Giordani, B., Boivin, M. J., Opel, B., & Nseyila, D. D. (1999). Modality specificity of memory span tasks among Zairian children: A developmental perspective. *Journal of Clinical and Experimental Neuropsychology, 21*, 375-384.
- Dunn, O. J., & Clark, V. A. (1969). Correlation coefficients measured on the same individuals. *Journal of the American Statistical Association, 64*, 366-377.
- Durand, M., Hulme, C., Larkin, R., & Snowling, M. (2005). The cognitive foundations of reading and arithmetic skills in 7- to 10-year-olds. *Journal of Experimental Child Psychology, 91*, 113-136.
- Fairweather, H. (1976). Sex differences in cognition. *Cognition, 4*, 231-280.
- Georgas, J., Van de Vijver, F. J. R., Weiss, L. G., & Saklofske, D. H. (2003). A cross-cultural analysis of the WISC-III. In J. Georgas, L. G. Weiss, F. J. R. Van de Vijver, & D. H. Saklofske (Eds.), *Culture and children's intelligence: Cross-cultural analysis of the WISC-III* (pp. 277-313). San Diego, CA: Academic Press.
- Holding, P. A., Taylor, H. G., Kazungu, S. D., Mkala, T., Gona, J., Mwamuye, B., et al. (2004). Assessing cognitive outcomes in a rural African population: Development of a neuropsychological battery in Kilifi District, Kenya. *Journal of the International Neuropsychological Society, 10*, 246-260.
- Irvine, S. H. (1979). The place of factor analysis in cross-cultural methodology and its contribution to cognitive theory. In L. H. Eckensberger, W. J. Lonner, & Y. H. Poortinga (Eds.), *Cross-cultural contributions to psychology* (pp. 300-341). Lisse, Netherlands: Swets & Zeitlinger.
- Jansen, P., & Greenop, K. (2008). Factor analyses of the Kaufman Assessment Battery for Children assessed longitudinally at 5 and 10 years. *South African Journal of Psychology, 38*, 355-365.
- Kaufman, A. S., & Kaufman, N. L. (1983). *Kaufman Assessment Battery for Children*. Circle Pines, MN: AGS Publishing.
- Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Assessment Battery for Children, second edition: Manual*. Circle Pines, MN: AGS Publishing.
- Kaufman, J. C., Kaufman, A. S., Kaufman-Singer, J., & Kaufman, N. L. (2005). The Kaufman Assessment Battery for Children-Second Edition and the Kaufman Adolescent and Adult Intelligence Test. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 344-370). New York: Guilford Press.
- Malda, M., Van de Vijver, F. J. R., Srinivasan, K., Transler, C., Sukumar, P., & Rao, K. (2008). Adapting a cognitive test for a different culture: An illustration of qualitative procedures. *Psychology Science Quarterly, 50*, 451-468.
- Mardell-Czudnowski, C. (1995). Performance of Asian and White children on the K-ABC: Understanding information processing differences. *Assessment, 2*, 19-29.
- McGrew, K. S. (2005). The Cattell-Horn-Carroll theory of cognitive abilities: Past, present, and future. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 136-181). New York: Guilford Press.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-103). New York: American Council on Education and Macmillan Publishing.
- Mishra, R. C. (1997). Cognition and cognitive development. In J. W. Berry, P. R. Dasen, & T. S. Saraswathi (Eds.), *Handbook of cross-cultural psychology: Vol. 2. Basic processes and human development* (pp. 143-175). Needham Heights, MA: Allyn & Bacon.
- Misra, G., Sahoo, F. M., & Puhan, B. N. (1997). Cultural bias in testing: India. *European Review of Applied Psychology, 47*, 309-316.
- Moon, S., McLean, J. E., & Kaufman, A. S. (2003). A cross-cultural validation of the sequential-simultaneous theory of intelligence in children. *School Psychology International, 24*, 449-461.
- Ramaa, S. (1994). *Arithmetic diagnostic test for primary school children*. Mysore, India: Chethana Book House.

- Reynolds, M. R., Keith, T. Z., Goldenring-Fine, J., Fisher, M. E., & Low, J. A. (2007). Confirmatory factor structure of the Kaufman Assessment Battery for Children—Second edition: Consistency with Cattell–Horn–Carroll theory. *School Psychology Quarterly*, *22*, 511-539.
- Reynolds, M. R., Keith, T. Z., Ridley, K. P., & Patel, P. G. (2008). Sex differences in latent general and broad cognitive abilities for children and youth: Evidence from higher-order MG-MACS and MIMIC models. *Intelligence*, *36*, 236-260.
- Taub, G. E., & McGrew, K. S. (2004). A confirmatory factor analysis of Cattell–Horn–Carroll theory and cross-age invariance of the Woodcock–Johnson Tests of Cognitive Abilities III. *School Psychology Quarterly*, *19*, 72-87.
- Thurstone, L. L. (1931). *The reliability and validity of tests: Derivation and interpretation of fundamental formulae concerned with reliability and validity of tests and illustrative problems*. Ann Arbor, MI: Edwards Brothers.
- Van de Vijver, F. J. R. (1997). Meta-analysis of cross-cultural comparisons of cognitive test performance. *Journal of Cross-Cultural Psychology*, *28*, 678-709.
- Van de Vijver, F. J. R. (2003). Test adaptation/translation methods. In R. Fernández-Ballesteros (Ed.), *Encyclopedia of psychological assessment* (pp. 960-963). Thousand Oaks, CA: Sage.
- Van de Vijver, F. J. R., & Poortinga, Y. H. (2005). Conceptual and methodological issues in adapting tests. In R. K. Hambleton, P. F. Merenda, & C. D. Spielberger (Eds.), *Adapting educational and psychological tests for cross-cultural assessment* (pp. 39-63). Mahwah, NJ: Lawrence Erlbaum.
- Van de Vijver, F. J. R., & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: An overview. *European Review of Applied Psychology*, *54*, 119-135.
- Verney, S. P., Granholm, E., Marshall, S. P., Malcarne, V. L., & Saccuzzo, D. P. (2005). Culture-fair cognitive ability assessment: Information processing and psychophysiological approaches. *Assessment*, *12*, 303-319.
- Wechsler, D. (1949). *Wechsler Intelligence Scale for Children*. New York: Psychological Corporation.
- Wechsler, D. (1974). *Manual for the Wechsler Intelligence Scale for Children—revised*. San Antonio, TX: Psychological Corporation.
- Wechsler, D. (1991). *Wechsler Intelligence Scale for Children—third edition*. San Antonio, TX: Psychological Corporation.
- Widaman, K. F., & Reise, S. P. (1997). Exploring the measurement equivalence of psychological instruments: Applications in the substance use domain. In K. J. Bryant, M. Windle, & S. G. West (Eds.), *The science of prevention* (pp. 281-324). Washington, DC: American Psychological Association.