

Measurements of Intelligence in sub-Saharan Africa: Perspectives Gathered from Research in Mali

Cissé Dramé¹ · Christopher J. Ferguson²

© Springer Science+Business Media New York 2017

Abstract One of the most controversial debates around intelligence testing regards how tests are used to measure intelligence among non-Western populations. Studies conducted since the 1930's consistently indicate significant sub-average intelligence among African populations. The purpose of this study is to examine whether commonly used intelligence tests such as the Ravens Progressive Matrices are valid indices of cognitive functioning among children in Mali, Africa. Participants in the current study were 206 children from Mali attending French-language schools. The Woodcock-Johnson II math assessment was used to measure participants' academic achievement. The Vineland II- Adaptive Behavior Scale (VABS) was used to indicate their adaptive functioning level. In this study, tests of IQ were compared against adaptive functioning and academic achievement, to examine whether IQ scores measured among African populations are artificially low or are an accurate measure of performance. IQ scores as measured by the Ravens were discrepant with standardized scores on math achievement and adaptive functioning. Results indicate that use of the Ravens may substantially underestimate the intelligence of children in Mali. This can be particularly problematic when comparisons are made across cultures using the same test and norms. . It is recommended that tests developed with local normative samples be used to assess for IQ.

Keywords Intelligence · Child development · Cross-cultural · Testing · Africa

Introduction

Issues of race and intelligence remain extremely controversial, particularly whether intelligence (IQ) tests developed in Western countries can be meaningfully used to infer intellectual functioning among non-Western cultures. The DSM-V discusses the nature of Intellectual Disability (Intellectual Developmental Disorder; IDD) at length. Under "Specifiers" it notes, "the various levels of severity are defined on the basis of adaptive functioning, and not IQ scores, because it is adaptive functioning that determines the level of supports required." This specification of intellectual disability is crucial in understanding the role of testing in sub-Saharan Africa for two reasons: this specification may help clarify debates about the role of IQ in understanding intellectual functioning among sub-Saharan cultures; and establishing a more effective measurement model for intellectual and adaptive functioning in such cultures. Although most psychological evaluators understand the importance of using local, culturally similar norms when evaluating tests, such norms have not always been available in underdeveloped nations. This may have caused some confusion when considering comparisons of IQ between cultures using similar tests of intelligence with western norms. (e.g. Lynn et al. 2005; Rushton, 2008). The current study in working with Malian children in West Africa, seeks to examine the degree of concordance between measurement of IQ scores and adaptive functioning.

The interplay between race and intelligence has been controversial for some time (e.g. Cooper 2005; Rowe and Rodgers 2005). As part of this issue is consideration of whether perceived differences in IQ are genetic as well as

✉ Christopher J. Ferguson
CJFerguson1111@aol.com

¹ Texas A&M International University, 5201 University Blvd, Laredo, TX, USA

² Department of Psychology, Stetson University, 421 N. Woodland Blvd, DeLand, FL 32729, USA

whether differences in IQ scores among populations indicate differential adaptive functioning (Jensen 2013). Individuals with a score of 65–75, or two standard deviations below the population mean, are classed as having Intellectual Disability Disorder. This criterion applies to both adaptive behavior and IQ. Adaptive behavior refers to the social and practical skills that people use to effectively function in their everyday lives. These skills include communication, social interactions, and daily living skills. IQ can be considered a general capacity to learn, although the specific nature of IQ or intelligence continues to be debated (e.g. Nisbett et al. 2012).

In 1992, the American Association on Intellectual and Developmental Disabilities (AAIDD) changed the basis of classification for Intellectual Disability based not on IQ scores, but instead on the level of support individuals require to function successfully in everyday society. AAIDD concluded low IQ (65–75) alone is insufficient for a diagnosis of intellectual disability (Maulik and Harbour 2010). Leading professional organizations use the approach of AAIDD to include IQ *and* adaptive behavior: International Classification of Diseases – 10th Revision (World Health Organization 1992), Diagnostic Statistical Manual of Mental Disorders – 4th Edition, Text Revision (APA, 2000) and International Classification of Functioning, Disability, and Health (World Health Organization 2001). An accurate diagnosis of intellectual disability thus requires *two* components: An IQ score of 65–75 or below *and* a determination of deficits in adaptive behavior. Nevertheless, much of the research studies in Africa, if not all, continue to prioritize IQ scores obtained from western-developed tests over adaptive measures. In many respects, this test procedure has not changed since the inception of intelligence testing beginning around 1900. Given research in the field as it has developed in the past hundred years, it would seem overdue to ask, is this procedure for measuring intelligence across-cultures valid?

Measuring Intelligence across Cultures

The Raven's Progressive Matrices (RPM) is one of the most widely used non-verbal IQ tests precisely because it has been presumed to be culturally fair. The test is also considered to be one of the best predictors of Spearman's *g* factor (Court 1983; Jensen 1998). Research has consistently reported significantly sub-average IQ scores in sub-Saharan Africa on the Raven's Progressive Matrices. For instance, Lynn et al. (2008) have found younger Libyan children aged 6 and 7 performed better on Progressive Matrices than the older children in relation to British norms, although Libyan scores were low compared to British averages. In Ghana, an important study conducted by Anum (2014) standardized the Raven's Colored Progressive Matrices (RPCM) in rural and

urban schools. The results indicated that when compared to the British published norm (Raven and Court 1989), the range of scores for the Ghanaian norms was much lower. Indeed, scores that correspond to the 95th percentile rank in the Ghana data were comparable to scores at the 50th percentile on the British norms. Furthermore, children in the urban group performed higher than the rural group on the RPCM. Similarly, Constenbader and Ngari (2001) have found that norms generated on a sample of Kenyan children differ substantially from the norms reported for other groups (Raven et al. 1990).

Although most of this paper is focused on Africa, there is evidence that suggests that the need to develop appropriate norms to interpret Western-standardized IQ tests is not limited only to African populations (Lynn et al. 2005; Pind et al. 2003; Pullmann et al. 2004). One solution to reduce the potential for Eurocentric bias in administering and interpreting standardized tests in sub-Saharan Africa would be to collect data using culture-specific samples and use those norms to adapt and change specific test items. However, re-standardization cannot alter that the original test items were developed with Eurocentric samples (Sternberg 2000). While the RPM's validity with non-European populations continues to be questioned by some in the scientific community, the current study is designed to help fill in gaps in the area of cross-cultural testing by examining the relationship between IQ scores and adaptive functioning scores among a sample of youth in Mali. Adaptive behavior outcomes related to communication, daily living skills and socialization will all be considered. The designs are meant to test the hypothesis that significantly low intelligence functioning will be related to significant limitations in adaptive functioning in one or more activities of daily life. Additionally, this study will examine the relationship among math scores, IQ scores and adaptive functioning.

Methods

Participants

Participants in the study were 206 children from Mali (city of Bamako) attending French language schools. Students were enrolled in an urban private school located in an urban area. 92 (44.7%) of the participants were male. The majority (95.6%) were from the Bambara ethnic group, with small numbers of Soninke, Sonrhail and Senoufo. Participants ranged between ages 6 and 14 ($M = 9.85$, $SD = 1.61$). The students tested were in the second, third, fourth and fifth grade. Testing took place during school time; teachers filled out their surveys during downtime. Students were recruited from classes in which teachers agreed to participate. Consent was then obtained from parents and all students with

parental consent from participating classes were included. No incentives for participation were offered. All scores were converted to age normed standard scores. One teacher declined to complete the Vineland, thus sample size for analyses with this measure is 151. All students with family and teacher consent from a participating school were included.

There are three types of schooling available in Mali: 1) public schools run by the national government 2) private schools run by private entities and 3) Koranic schools which focus on the study of the Koran. In this study, children were recruited in a private nonreligious school. For both private and public schools in Mali, classes are taught in French, but based on the Mali national education program.

There is relatively little research on the diagnosis of intellectual disability in Mali. Children are typically referred for testing if problems are noted with academic or adaptive functioning. Children may be referred for testing for IDD by parents, teachers, general doctors, or pediatricians. For children, the Wechsler Intelligence Schedule for Children (WISC) is often used for diagnosis. International norms are commonly used and testers may adapt certain items to the local culture.

Measures

Raven's Progressive Matrices (Raven et al. 1990; RPM)

The RPM is a commonly used scale of intelligence that does not require verbalizations. Participants are asked to identify the missing piece of a visual pattern from among several potential options. The Ravens is administered as 5 sets of 12 items for a total number of 60 items. The RPM specifically avoids verbally loaded items in attempt to avoid the cultural loading inherent in such items. Our interest is in comparing functioning on the RPM using British norms in order to examine the validity of cross-racial intelligence claims. Although norms for the RPM have been developed for some African countries, we were unaware of any norms developed in Mali, and some evidence suggest that IQ test norm development in African countries has not always produced valid norms (Shuttleworth-Edwards 2016). Use of British norms here also allows us to examine the issue of cross-national comparisons of IQ using a set group of norms which has become controversial.

Vineland Adaptive Behavior Scales (Sparrow et al. 1984)

In the case of the adaptive behavior criterion, the Vineland Adaptive Behavior Scales (VABS) was administered in order to measure conceptual, social and practical skills. The VABS consists of a normed interview checklist to assess whether children are able to perform appropriate adaptive behaviors (e.g. tie shoes, dress selves, etc.) appropriate for their age level. Separate scores were calculated for communication (COM), daily living skills (DLS) and socialization

(SOC). The Vineland includes both starting points with questions appropriate to age group (example items range from "Says month and day of birthday when asked" through "describes a realistic long-range goal that can be done in 6 months or more") as well as discontinue "ceiling" rules for when four consecutive zero-score answers are achieved. As such, the specific number of questions applicable to a given child may vary slightly due to the start and discontinue rules applicable for that child.

Woodcock-Johnson III (McGrew et al. 2007 WJ)

The WJ is a commonly employed test of academic achievement across reading, writing and math. The WJ is commonly employed to test for learning disabilities and other academic deficits. For the current study, math achievement scores were calculated as these were not language dependent. The Calculation subscale was solely used, with raw scores converted to standard scores. This subscale consists of 45 math problems of varying difficulty, ranging from simple addition through high school level algebra. Administration is discontinued after six consecutive wrong answers. Although there are multiple ways achievement could have been assessed, the use of math achievement allowed for a fairly culturally-free approach, particularly for a sample of children for whom well-developed, normed tests are fairly few. Use of this tests with basic norms allowed for a comparison of how Mali children were achieving on a standard subject relative both to international expectations as well as their performance on RPM.

Note that, for comparison, all scores were converted to IQ standard scores ($M = 100$; $SD = 15$.)

Procedure

Questionnaire data were recorded from students and teachers, in 2012. Students were administered two scales: the RPM and the Woodcock-Johnson III. The RPM was administered in French following the directions for individual administration suggested in the test manual (Raven and Court 1989). The WJ III was administered following the instructions for group administration in the test manual (McGrew et al. 2007 WJ). Teachers completed the Vineland Adaptive Behavior Scales (Teacher Rating Form). All scales were translated into French by the lead author. These translation were then checked for accuracy by an independent member of the university IRB. Prior to being administered, each item of the VABS was evaluated by a panel of Psychologists from Mali using the CBT/McGraw-Hill review form in order to avoid item bias. As a result, a total of 12 items were either eliminated or reconstructed in order to make sense in the Malian culture and ensure the culture fairness of the test.

Results

Mean Differences

For RPM to be valid indices of intellectual functioning among children in Mali, standard scores ($M = 100$, $SD = 15$) should not differ substantially between the RPM, VABS and WJ. In order to use the RPM to make cross-ethnic IQ comparisons, low scores on the RPM should not only correlate with the VABS and WJ, but generally correspond in scale score means. Significant differences between IQ means and those on adaptive functioning and academic achievement could indicate cultural artifacts in the RPM resulting in artificially low scores.

RPM score correlations with the adaptive behavior and math achievement are presented in Table 1. As expected, RPM scores correlate with all outcomes, although correlations with VABS (SOC) are fairly weak. This indicates that the RPM may be reasonably effective in predicting outcomes among children in Mali. However, such correlations are not sufficient to document the use of the RPM to make cross-racial comparisons using a single set of westernized norms. The RPM may still correlate with outcomes despite the presence of cultural artifacts that truncate mean RPM scores among Mali children.

To examine concordance between RPM scores with VABS and WJ scores, paired samples t-tests were employed. Table 2 presents the mean, standard deviation and t scores for the differences between RPM and VABS and WJ scores. As can be seen, scaled scores for the RPM were far lower, corresponding perhaps to frequent claims made about low IQ scores in African populations. However, these scores did not correspond to standardized scores on math achievement or adaptive functioning. Differences, in fact, were rather large. These differences suggest a worrying discrepancy between RPM scores and adaptive functioning in African youth.

In some cases, the predictive accuracy of a test can be presented in table form to illustrate proportion of children whose criterion scores are accurately predicted (see Lichtwark et al. 2013.) However, in this case, the RPM scores were so uniformly low for our sample (only 40.8% of the sample scored above at standard score of 74 on the RPM and only 1% above 84) and few criterion scores fell within this range (for instance, only a single child (0.5%) scored

Table 1 Correlations between RPM and WJ and VABS scores

	WJ Math	VABS (COM)	VABS (DLS)	VABS (SOC)
RPM	.697**	.478*	.406**	.191*

** $p < .01$

* $p < .05$

Table 2 Means and differences between standardized scores on the RPM and WJ and VABS

Scale	Mean score	Standard deviation	t Diff
RPM	65.85	7.14	
WJ	106.67	17.90	42.15***
VABS (COM)	114.07	8.62	69.69***
VABS (DLS)	106.62	8.92	53.38***
VABS (SOC)	114.77	10.77	49.15***

t Diff = test for difference between RPM and other scales;

*** $p < .001$

below 70 on the WJ and only 3.4% below 85.) As such, the predictive accuracy of the RPM for predicting standard deviation bands (below 70, 70–84, 85–100, etc.) was clearly near to 0%.

Discussion

Previous studies have suggested that IQ test norms for African samples are considerably lower than for European or North American samples. Such observed test differences are often used to conclude that real cognitive differences exist between African and European samples (e.g. Rindermann et al. 2014; Rushton and Skuy 2000). Although our results appear to support the existence of relatively low scores among African youth on the Raven's Progressive Matrices, our results also suggest that such scores do not necessarily reflect lower academic performance or adaptive living skills among African children. Thus, it is advised that caution be employed when interpreting results from the Raven's Progressive Matrices as indicating lower general intelligence among African samples.

The question of why African samples perform relatively poorly on the Raven's is a curious one. At present, we can only speculate, but one possibility is that the issue is one of culturally specific transfer-of-learning. That is to say, it may be that European and North American children may spend more time with play tasks such as jigsaw puzzles or connect the dots that have similarities with the Ravens and, thus, train on similar tasks more than do African children. If African children spend less time on similar tasks, they would have fewer opportunities to train for the Ravens (however unintentionally) reflecting in poorer scores. In this sense, verbal ability need not be the only pitfall in selecting culturally sensitive IQ testing approaches. Thus, differences in Ravens scores may be a cultural artifact rather than an indication of true intelligence differences.

Our results suggest that these low scores on the Ravens do not translate to lower academic functioning or adaptive living skills. As noted above, the tasks of the RPM may simply be less familiar to children in Mali than to children in the West, resulting in spuriously lower scores. It is also possible that children in Mali may be less familiar with the format of the RPM questions (i.e. multiple choice) than are children in the West. Thus, we advise caution in using the Ravens to imply lower intelligence or greater cognitive dysfunction among African samples. It is possible that the RPM could conceivably be renormed with a sample of children from Mali. This would require a significant investigation of the validity of such norms that was beyond the scope of this paper, which focused on the issue of cross-cultural comparisons of IQ based on a particular set of normative assumptions.

It is worth noting that our concerns mainly focus on using the RPM to make comparisons of IQ across cultures. Our analyses suggest that such comparisons using the RPM are fraught with potential sources of misinformation. However, RPM scores were correlated with achievement and adaptive functioning outcomes. Thus, the RPM may retain some basic value use in predicting academic outcomes among youth in Mali. However, it is essential that any cognitive testing undertaken with the RPM ought to use norms developed in Mali. Even norms developed with other African cultures could differ and it should not be assumed that norms developed in other African countries could apply to Mali. It may also be desirable to consider development of “ground up” strategies for IQ test development among children in Mali. That is to say, rather than converting tests developed in Western countries for use in Mali, developing cognitive tests based on tasks that are culturally familiar to children in Mali rather than those in the west. Such locally developed IQ tests may have greater validity in predicting performance in schools in Mali.

As with all studies, our analyses have limitations. Our study was conducted with children enrolled in French-language schools. As indicated by their relatively high adaptive living skills scores, such children are unlikely to be representative of all children in Mali or other regions of Africa, particularly those for whom poverty or social strife prevent their attending school. Also, our study is correlational in nature, so causality cannot be asserted.

We hope that our study provides further data for discussions of cognitive functioning and intelligence testing among African samples. These issues have often been contentious, but further data will help in illuminating them.

Compliance with Ethical Standards

Conflict of Interest None to declare for either author.

References

- American Psychiatric Association. (2000). *Diagnostic and statistical manual – IV – TR*. Arlington, VA: American Psychiatric Association.
- Anum, A. (2014). A standardization study of the Raven’s Coloured Progressive Matrices in Ghana. *IFE Psychologia*, 22(2), 27–35.
- Constenbader, V., & Ngari, S. M. (2001). A Kenya standardization of the Raven’s colored Progressive Matrices. *School Psychology International*, 22(3), 258–268.
- Cooper, R. S. (2005). Race and IQ: Molecular genetics as deus ex machina. *American Psychologist*, 60(1), 71–76. doi:10.1037/0003-066X.60.1.71.
- Court, J. H. (1983). Sex differences in performances on Raven’s Progressive Matrices: A review. *Alberta Journal of Educational Research*, 29, 54–74.
- Jensen, A. R. (1998). *The g factor*. Westport: Praeger.
- Jensen, A. R. (2013). Rushton’s contributions to the study of mental ability. *Personality and Individual Differences*, 55(3), 212–217. doi:10.1016/j.paid.2012.03.008.
- Lichtwark, I. T., Starkey, N. J., & Barker-Collo, S. (2013). Further validation of the New Zealand Test of Adult Reading (NZART) as a measure of premorbid IQ in a New Zealand sample. *New Zealand Journal Of Psychology*, 42(3), 75–83.
- Lynn, R., Abdalla, S., & Al-Shahomee, A. (2008). Norms for the Progressive Matrices for Libya and Tunisia. *Mankind Quarterly* 49(3), 71–77.
- Lynn, R., Backhoff, E., & Contreras, L. A. (2005). Ethnic and racial differences on the standard Progressive Matrices in Mexico. *Journal of Biosocial Science*, 37(1), 107–113.
- Maulik, P. K., Harbour, C. K. (2010). Epidemiology of intellectual disability. *International Encyclopedia of Rehabilitation*, 2. <http://cirrie.buffalo.edu/encyclopedia/en/article/144/>. Access 3 April 2017.
- McGrew, K., Schrank, F., & Woodcock, R. (2007). *Woodcock-Johnson III: Normative update*. Rolling Meadows: Riverside Publishing.
- Nisbett, R. E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D. F., & Turkheimer, E. (2012). Intelligence: New findings and theoretical developments. *American Psychologist*, 67(2), 130–159. doi:10.1037/a0026699.
- Pind, J., Gunnarsdottir, E. K., & Johansson, H. S. (2003). Raven’s standard Progressive Matrices: new school age norms and a study of the test’s validity. *Personality and Individual Differences*, 34, 375–386.
- Pullmann, H., Allik, J., & Lynn, R. (2004). The growth of IQ among Estonian schoolchildren from ages 7 to 19. *Journal of Biosocial Science*, 36(6), 735–740.
- Raven, J., & Court, J. H. (1989). *Manual for Raven’s Progressive Matrices and vocabulary scales*. London: Lewis.
- Raven, J., et al. (1990). *Manual for Raven’s Progressive Matrices and vocabulary scales: Research supplement no. 3: American and international norms*. Oxford: Oxford Psychologists Press.
- Rindermann, H., Falkenhayn, L., & Baumeister, A. E. (2014). Cognitive ability and epistemic rationality: A study in Nigeria and Germany. *Intelligence*, 4723–4733. doi:10.1016/j.intell.2014.08.006.
- Rowe, D. C., & Rodgers, J. (2005). Under the skin: On the impartial treatment of genetic and environmental hypotheses of racial differences. *American Psychologist*, 60(1), 60–70. doi:10.1037/0003-066X.60.1.60.
- Rushton, J. P. (2008). Testing the genetic hypothesis of group mean IQ differences in South Africa: Racial admixture and cross-situational consistency. *Personality and Individual Differences*, 44(3), 768–776. doi:10.1016/j.paid.2007.10.011.
- Rushton, J. P., & Skuy, M. (2000). Performance on Raven’s Matrices by African and white university students in South Africa. *Intelligence*, 28(4), 251–265. doi:10.1016/S0160-2896(00)00035-0.
- Shuttleworth-Edwards, A. B. (2016). Generally representative is representative of none: commentary on the pitfalls of IQ test standardization

- in multicultural settings. *The Clinical Neuropsychologist*, 30(7), 975–998. doi:10.1080/13854046.2016.1204011.
- Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (1984). *Vineland adaptive behavior scales*. Circle Pines: American Guidance Service.
- Sternberg, R. J. (2000). Implicit theories of intelligence as exemplar stories of success: why intelligence test validity is in the eye of the beholder. *Psychology, Public Policy, and Law*, 6, 159–167.
- World Health Organization. (1992). *International classification of diseases—tenth Revision*. Geneva: World Health Organization.
- World Health Organization. (2001). *International classification of functioning, disability, and Health*. Geneva: World Health Organization.