



## Assessment of cognitive abilities in multiethnic countries: The case of the Wolof and Mandinka in the Gambia

Matthew C. H. Jukes<sup>1,2,\*</sup> and Elena L. Grigorenko<sup>3,4</sup>

<sup>1</sup>Harvard Graduate School of Education, Cambridge, Massachusetts, USA

<sup>2</sup>Partnership for Child Development, Imperial College, London, UK

<sup>3</sup>Yale University, New Haven, Connecticut, USA

<sup>4</sup>Moscow State University, Moscow, Russia

**Background.** The use of cognitive tests is increasing in Africa but little is known about how such tests are affected by the great ethnic and linguistic diversity on the continent.

**Aim.** To assess ethnic and linguistic group differences in cognitive test performance in the West African country of the Gambia and to investigate the sources of these differences.

**Samples.** Study 1 included 579 participants aged 14–19 years from the Wolof and Mandinka ethnic groups of the Gambia. Study 2 included 41 participants aged 12–18 years from the two ethnic groups.

**Methods.** Study 1 assessed performance on six cognitive tests. Participants were also asked about their history of education, residence in the city, parental education, and family socio-economic status. Study 2 assessed performance on two versions of the digit span test. Recall of the numbers 1–5 were compared with recall of numbers 1–9 for both the Wolof (who count in base 5) and the Mandinka (who count in base 10).

**Results.** Study 1 established that Wolof performance was lower than that of the Mandinka on five out of six cognitive tests. In four of these tests, group differences were partially mediated by participation in primary school and migration to the city. Group differences were substantial for the digit span test and were not attenuated by mediating variables. Study 2 found that digit span among the Wolof was shorter than that of the Mandinka for numbers 1–9 but not for numbers 1–5.

**Conclusions.** Several suggestions are made on how to consider the ethnicity, language, education, and residence (urban vs. rural) of groups when conducting comparative cognitive assessments or collecting normative data.

\*Correspondence should be addressed to Dr Matthew C. H. Jukes, Harvard Graduate School of Education, 6 Appian Way, Gutman 457, Cambridge, MA 02138, USA (e-mail: jukesma@gse.harvard.edu).

2 Matthew C. H. Jukes and Elena L. Grigorenko

The adaptation of cognitive tests to new populations and settings presents a number of challenges. One issue that has received relatively little attention is the adaptation of tests for use in multiethnic countries (Carter *et al.*, 2005; van de Vijver & Phalet, 2004). Many assessment tools originated in the countries of Europe or North America, where only one or two languages are spoken. The situation is different in most African countries where cultural and linguistic diversity abound, with an estimated 3,000 spoken languages and 8,000 dialects (Owino, 2002). The Gambia, where our research was conducted, provides a good illustration of such diversity. The country is a thin (~40 km wide) strip of land on either side of the River Gambia with a population of around 1.7 million (UNPD, 2005) and yet there are eight indigenous languages spoken in the Gambia (Gordon, 2005) in addition to the colonial language of English. The aim of this article is to outline the factors affecting cognitive test performance in such multiethnic settings and to examine these factors empirically using data from the Gambia collected from two ethnic groups, the Mandinka and the Wolof. We use the term ‘ethnic group’ in this article to refer to people who identify with each other on the basis of a common ancestry and a preference for endogamy and who share common cultural, linguistic, religious, behavioural, and biological traits (Banks, 1996). In this article, we discuss ‘ethnic group differences’ in cognitive performance. In theory, these differences may arise from any of the dimensions of ethnicity outlined above, broadly classified as environment or genetic in origin. However, the focus of this paper is on the environmental determinants of cognitive test performance, both because selective hereditary pressures tentatively hypothesized to influence the evolution of cognitive abilities (Rushton & Jensen, 2005) are unlikely to differ between the Mandinka and Wolof, given highly similar population histories of the two tribes, to our knowledge, and also because even the staunchest supporters of the hereditary position for origins of group differences in cognition suggest that environmental influences are substantial (Rushton & Jensen, 2005).

### ***The growing need for mother tongue assessment***

The challenges faced by cognitive testing in multiple ethnic groups in Africa have received little attention. Most assessments of children’s mental abilities in sub-Saharan African have involved school-based tests of educational achievement typically conducted in a limited number of colonial languages or other lingua franca. However, there are several trends in assessment on the continent requiring the development of measures in children’s mother tongue. First, the use of cognitive assessments other than educational achievement tests is increasing in Africa. There has been widespread use of cognitive tests to assess the impact of health and nutrition interventions in early childhood and in school age (Clarke *et al.*, 2008; Grigorenko *et al.*, 2006; Holding, Stevenson, Peshu, & Marsh, 1999; Jukes, Drake, & Bundy, 2008; Jukes *et al.*, 2002). Similarly, the expansion of early childhood care and education programme in Africa (UNESCO, 2007) has led to an increase in the use of school readiness tests and other measures of cognitive development. In all such cases, the use of mother tongue as the medium of assessment produces a more valid measure of children’s abilities (Carter *et al.*, 2005; Claassen, 1997; Foxcroft, 1997; Grieve, 2005; Owen, 1991). Second, there is an increasing recognition of the importance of mother tongue as a medium of instruction for the early grades of primary school (Alexander, 1995; Chumbow, 2005; Heugh, 1995; Lockett, 1995; Trudell & Schroeder, 2007). Scholars and writers from across Africa stated in the Asmara Declaration (2000) that ‘all African children have the

unalienable right to attend school and learn in their mother tongues'. At the same time, there is a push for more monitoring and assessment of early literacy and numeracy skills (RTI, 2008). Such educational assessments also require adaptation to the many languages in multiethnic countries. Thus, an increase in the use of mother tongue assessments in Africa and the need for clear guidelines on how to ensure validity and reliability of such assessments across multiple ethnic groups are anticipated on several fronts.

Before considering how these issues play out among the ethnic groups of the Gambia, we first describe the culture, history, and language of the two groups in our study.

### **The Mandinka and the Wolof**

Mandinka are the largest ethnic group in the Gambia with 453,500 members. There are 165,000 Wolof, making it the third largest group in the country, behind the Fula. However, the Wolof predominate in Banjul, the capital city of the Gambia, constituting around half of the population there.

The Mandinka language (also referred to as Mandingo or Malinke) is one of the most prevalent languages in West Africa and is spoken in several countries including Senegal and the Gambia (Camara, 1999). Wolof is spoken in central and western Gambia (Gordon, 2005). It is a lingua franca in the Gambia, Senegal, and Mauritania (Ngom, 2003). There is little dialectal variation; the differences that exist between dialects result mainly from the influence of English, most notably in loanwords. Both Wolof and Mandinka are part of the Niger-Congo family, but Mandinka is from the Mande branch and Wolof is from the Atlantic-Congo branch (Gordon, 2005).

Mandinka (Rowlands, 1959) has a five vowel system with both long and short vowels. There are 21 consonants including: voiced and voiceless stops, nasals, voiceless fricatives, laterals, and approximants. The syllable structure is (C)V(N).<sup>1</sup> Mandinka is a pitch accent language, where pitch is contrastive. There are two pitches called *level* and *moving*. Mandinka has a subject-object-verb sentence structure and has post-positions.

Wolof is quite distinct from Mandinka (Ngom, 2003). The vowel system of Wolof has short and long vowels. Diphthongs also occur as the result of phonological processes. Wolof has both voiced and voiceless consonants. Its consonant inventory includes stops, nasals, fricatives, affricates, laterals, and approximates. The syllable structure is CV(C).<sup>1</sup> Sentences in Wolof are subject-verb-object. Wolof is an agglutinating language with rich inflectional morphology, including eight noun classes. Wolof is a pro-drop language.

Religion and society have many apparent similarities in the two groups. Both groups have been almost exclusively Muslim for centuries. Both groups have hierarchical social stratification with hereditary nobility and a paramount chief at its head. Both groups are farmers, growing groundnuts (peanuts) as a cash crop, on which the country's economy is dependent, and millet and sorghum as staples. Across the Gambia, 70% of the population is engaged in agriculture.

### **Differential influences on cognitive test performance among the Wolof and Mandinka**

From our understanding of the Wolof and Mandinka peoples what challenges might we expect in conducting cognitive ability assessments in these two groups? How might the

---

<sup>1</sup> C, consonant; V, vowel; N, nasal; ( ), optional.

## 4 Matthew C. H. Jukes and Elena L. Grigorenko

two groups differ in performance? We have been careful to talk of ‘differences in cognitive test performance’ throughout this article. This phrase is intended to cover two categories of difference. First, performance may differ because of testing bias. Q1 Mellenbergh (1989) considers a test item to be biased when ‘it differs in difficulty between subjects of identical ability from different groups’ (p. 128). If the Mandinka and Wolof have the same working memory capacity, but experience different levels of difficulty on a test designed to measure this trait, such as the digit span test, this test can be said to be biased in these two populations. Alternatively, cognitive test performance may differ because of genuine differences in ability between the two groups. Both of these two categories of difference are evident in the following discussion.

Lessons learned from testing minority students in the West (Brown, Reynolds, & Whitaker, 1999) suggest that test performance can be influenced by culture, education, socio-economic factors, language, and tester identity. We consider each of these in turn. First, cognitive assessment must take into account cultural differences between groups. In our discussion above, we noted that there are many similarities in the cultures of the Mandinka and the Wolof. However, individuals in the two groups may be differentially exposed to other cultural influences. One such influence results from contact with urban culture. Although the study area in our research consisted of rural villages, it was commonplace for villagers to spend time with friends or relatives living in the one major urban area in the country, known as ‘the Kombos’. This consists of the relatively small island-city of Banjul, the capital city of the Gambia, its surrounding conurbation and the larger diffuse urban area of Serekunda. Villagers spend time in the Kombos in order to study, work, or help relatives. A large proportion of small businesses in the Kombos are run by people originating from our study area, and help is needed from their relatives living up-country. Employment is also often in connection with the tourist industry, a main driver of the city’s economy. Movement around the country is frequent and many temporary urban residents return to their villagers on a regular basis, for example, to provide labour at harvest time or to attend ceremonies. Such mobility reportedly varies by ethnic group. The Mandinka are more populous in the Gambia than the Wolof; they have more dense social networks and are more likely to have friends or relatives to stay with around the country. The Mandinka from our study area, around Farafenni town, trace a history of temporary city residence to pre-independence political movements (O. Mboob, personal communication, 13 November 2008) and today many run small businesses in the city. The Wolof, by contrast, are less likely to move to the city. Although the Wolof are relatively more common in urban areas, the urban and rural Wolof are somewhat distinct (Greenfield, 1966; Greenfield, Reich, & Oliver, 1966) and Wolof parents cite potential maltreatment by prospective urban guardians as a reason for not sending their children to stay with relatives in the city. If this is the case, the rural Mandinka in our research may be more influenced by urban culture.

How might such urban culture influence cognitive test performance? It may be that urban culture assigns different values (Ardila, 2005; Greenfield, 1997) to cognitive skills compared to rural culture. It is observed across the continent (Bissiliat, Laya, Pierre, & Pidoux, 1967; Dasen *et al.*, 1985; Fortes, 1938; Grigorenko *et al.*, 2001; Serpell, 1993; Super, 1983) that Africans value social roles and the development of social responsibility and place less emphasis on some aspects of cognitive proficiency. Some have argued that cooperative production in subsistence agriculture leads to communities valuing compliance (Berry, 1967), respect, and obedience (LeVine *et al.*, 1993). Conversely, traits required for proficiency in cognitive tests may be less valued in African culture. For example, in Zambia, children performed speeded tasks more slowly than children

in the USA and were less responsive to requests to increase speed (Mulenga, Ahonen, & Aro, 2001). In urban Africa, where cooperative agricultural production gives way to forms of employment more reliant on cognitive skills, values may differ. For example, one study in Uganda found that slowness of cognitive performance was linked to intelligence by villagers in rural areas but not by primary school teachers nor by the Western educated elite (Wober, 1972). Similarly, urban culture may engender different conventions of communication (Greenfield, 1997) whereby dialogue dominated by information probes are more commonplace, making participants more at ease when taking part in cognitive assessments. Certainly, evidence suggests that children in urban environments perform better in cognitive tests than children from rural areas in Africa (Weisner, 1976) and in other developing countries (Mwamwenda, 1992; Rosselli & Ardila, 2003; Sinha, 1988). Similar evidence in the USA suggests that black-white differences in cognitive test performance are partially moderated by the extent to which African-Americans are acculturated (Manly *et al.*, 1998) to the majority culture of their country which values individualism and speeded performance. On the basis of this evidence, and on our observations in the Gambia, we would expect the Mandinka group who are more likely to travel in the country to be more exposed to urban culture and to have higher scores on tests of cognitive function.

A second source of differences in cognitive test performance may be the different levels of schooling between the two ethnic groups. Similar to the arguments presented above for the effects of urban culture, schooling may encourage individuals to value skills that promote performance in cognitive tests. Schooling may also help students become more familiar with the test-taking environment. In addition, schooling directly develops the cognitive abilities being assessed (Ceci, 1991). Evidence suggests that education can mediate differences in cognitive test performance between ethnic groups. Previous research among the Wolof (Greenfield, 1966) found that performance on a set of Piagetian cognitive tests was similar to Western norms among educated Wolofs but not among those who had not been to school. This supports the view that education is particularly effective in promoting cognitive abilities among communities where compliance is valued (Munroe, Munroe, & LeVine, 1972). Similarly, black-white differences in cognitive test performance in the USA are attenuated when adjustments are made for levels of education of the two groups particularly when educational quality is taken into account (Manly, Byrd, Touradji, & Stern, 2004; Manly, Jacobs, Touradji, Small, & Stern, 2002).

Families in the region of our research typically send their children either to Koranic schools (Madrassas) or to secular primary schools. In 1990, when participants in our study were beginning to attend primary school, only 59% of school-age boys and 41% of school-age girls were attending secular primary school (UNESCO, 2003). The rest of the children attended Madrassas or were out of school. From our observations, inhabitants of our study site report that the Wolof have a greater tendency to send their children to Koranic school whereas the Mandinka are more likely to send their children to secular schools so that the children can learn English. Based on these observations, we would expect the Mandinka group to have higher levels of secular education and consequently higher cognitive test scores.

A third potential source of differences between cultural groups is their socio-economic circumstances. The impact of poverty on children's development (Parker, Greer, & Zuckerman, 1988) is invoked as an explanation for poorer performance by minorities in cognitive tests in the USA. It is possible that socio-economic differences, other than those related to education and exposure to urban culture, may exist between

the two ethnic groups sampled in this research. However, neither the ethnographic literature nor our own informal observations suggested any clear differentiation in socio-economic circumstances between Mandinka and Wolof and, correspondingly, we did not formulate any clear hypotheses related to socio-economic variables.

A fourth influence on the comparability of cognitive test scores concerns the different languages spoken by the two groups, discussed above. There are several issues here. The development of equivalent tests of verbal abilities in the two languages requires more than translation of test items (Artiola i Fortuny *et al.*, 2005; Grigorenko *et al.*, 2009; Pena, 2007). Issues such as word frequency and the equivalence of linguistic constructs need to be considered in order for tests to be matched for difficulty and to ensure validity. In addition, the two languages in our studies use different counting systems. The Mandinka count in base 10 (as in English) whereas the Wolof count in base 5. This means that Wolof numbers 6–9 take longer to say than the Mandinka equivalent. For example, the Wolof word for *eight* is *juróom-ñett* (five-three) whereas the Mandinka word for eight is *sey*. This is likely to affect performance in the digit span test where the number of digits recalled is related to the length of the number words in the language of the test (Georgas, Weiss, van de Vijver, & Saklofske, 2003; Murray & Jones, 2002; Naveh-Benjamin & Ayres, 1986; Shebani, van de Vijver, & Poortinga, 2005). Compared to English, more digits can be remembered in languages such as Chinese where number words are shorter (Elliott, 1992). Correspondingly, we expected to see lower levels of performance among Wolof participants on the digit span test that involved digits higher than five.

Fifth, the identity and expectations of testers can affect cognitive performance. Participants may perform better in cognitive tests when the tester is from their own ethnic group (Terrell, Terrell, & Taylor, 1981), because they feel more comfortable with the tester or because they are better able to understand instructions. Bias may also result when testers have lower expectations of an individual's performance based on their ethnic identity. Thus, we expected to see lower levels of performance when participants were tested by testers from a different ethnic group.

The above discussion suggests that there are many potential challenges to developing equivalent cognitive tests in multiethnic settings. Test performance can be influenced by genuine differences in ability between groups, for example, through the influence of schooling and urban culture on the values attached to different cognitive skills, or by measurement bias resulting from familiarity with testing, language differences, or tester effects. The aim of our two research studies was to assess the difference in cognitive test performance between the Wolof and Mandinka groups in one region of the Gambia and to identify the source of these differences. Study 1 involved a large survey in 10 villages and used individual variability in background characteristics to examine hypotheses presented above. To recap, two key hypotheses were that higher education levels and greater exposure to urban culture would lead to higher test scores among the Mandinka. We assessed this hypothesis in Study 1 by examining the relationship between ethnic group membership, individual performance in cognitive tests and individual levels of education and history of urban residence. We also assessed other socio-economic status variables and tester identity as a source of ethnic group differences in test performance. The role of language in cognitive test performance could not be assessed in such an analysis of individual variability (because there is no individual variability in language spoken within one group). This was the issue we addressed in Study 2. Here, we aimed to test a hypothesis relating to one cognitive test: the Wolof have a poorer digit span because number-words are longer in Wolof than Mandinka for numbers 6–9.

## STUDY I

The aim of Study 1 was to assess ethnic group differences in performance of six cognitive tests, and to identify the extent to which education, place of residence, and other socio-economic factors mediate (Baron & Kenny, 1986) the relationship between ethnic group and cognitive test performance.

## Methods

### Participants

Participants were part of a study of the long-term educational impact of early childhood malaria (Jukes *et al.*, 2006). There were 579 participants, 296 male and 283 female, of mean age 17.1 years (range = 14 - 19). Participants were from 10 villages; 6 of which are predominantly inhabited by the Mandinka ethnic group and 4 by the Wolof group. The villages were located within 35 km of Farafenni, a small town with around 30,000 inhabitants and a busy ferry transporting Trans-Gambia highway traffic across the river. Farming was the main occupation among participants' fathers ( $N = 391$ ; 67%), with some engaged in waged employment ( $N = 52$ ; 9%) and trading ( $N = 35$ ; 6%).

### Cognitive tests

A battery of cognitive tests was adapted to the Gambia (Jukes *et al.*, 2006) and involved the following subtests.

#### *Visual search*

Visual search is a timed test that involved examining a row of pictures and identifying those that match a target picture to the left of the row (Baddeley, Gardner, & Grantham McGregor, 1995).

#### *Raven's coloured progressive matrices*

Raven's coloured progressive matrices (Raven, Styles, & Raven, 1998) required participants to identify the missing piece from a partially complete visual pattern or sequence. The coloured version of the test has good validity across cultures (Irvine, 1969; Owen, 1992).

#### *Digit span*

Digit span involved immediate recall of a string of digits in order (Wechsler, 1997).

#### *Categorical fluency*

Categorical fluency required children to name as many animals as they could in 1 min (Ardila, Ostrosky-Solis, & Bernal, 2006).

#### *Vocabulary test*

In this test a word is read out to participants, who then have to identify a synonym from among four options.

8 Matthew C. H. Jukes and Elena L. Grigorenko

#### *Proverb understanding*

This test was adapted from the Wechsler Adult Intelligence Scale (Wechsler, 1997) and measured participants' ability to explain the meaning of local proverbs.

All tests were translated from English into Wolof and Mandinka. The vocabulary test and the proverb understanding test were developed in parallel in the two languages. Wherever possible, equivalent words or proverbs were used in the two versions of these tests. All items were matched approximately for difficulty. All cognitive measures were piloted in the local population and had test-retest reliabilities over 7 days of more than 0.65.

All four cognitive testers were from the Mandinka ethnic group; three of them could speak fluent Wolof. Testers took part in 4 months of training and test piloting led by both authors and two research assistants experienced in psychometric assessment. Training covered the principles of assessment, procedures for administering the test battery and considerations for adapting the test battery to the local population. Test-retest reliabilities were assessed for all testers over a period of 7 days. Where mean reliability over the battery was less than 0.65, further training was conducted and reliability re-assessed. Three testers met the quality criterion at first assessment and the fourth tester met the criterion after one period of additional training.

#### **Socio-economic status, education, and migration**

A questionnaire was administered to participants to obtain information about their history of education and periods of living in other parts of the country. Information was also gathered on parental occupation and education and indicators of wealth such as the materials used to construct the family home and quantity of agricultural products sold.

#### **Analysis**

The main analyses proceeded in three steps, all conducted using Stata software (StataCorp, 2007). In each step, six multiple regression equations assessed the impact of independent variables (listed in Table 2) on standardized scores for the six cognitive tests. First, all independent variables were included in the regression equation to assess which had a significant impact on performance. Interactions between independent variables and ethnic group were included in regression equations; only terms significant in one or more of the six models were retained. Second, mediation analyses (MacKinnon, 2008) assessed whether each significant determinant of test performance mediated the relationship between ethnic group and test performance. In a final step, ethnic group differences in test performance were estimated controlling for significant mediators. Fifty-two participants had some missing data (see Table 2) and values were imputed from other socio-economic status variables using multiple imputation (Royston, 2004).

#### **Results**

We describe first the ethnic group differences in cognitive test performance and background variables before testing hypotheses about these differences. Of the 579 participants, analyses were restricted to the 562 participants who provided information on their history of education and migration. Table 1 describes the cognitive test scores for the two ethnic groups. In five of the six tests the Mandinka group outperformed the Wolof group - they were faster in the visual search test and had higher scores in four other tests ( $p < .01$  in all cases). These differences remained significant ( $p < .05$ ) even

**Table 1.** Cognitive test scores in Mandinka and Wolof ethnic groups

	Mandinka (N = 418)		Wolof (N = 144)	
	Mean	SD	Mean	SD
Visual search (secs) <sup>***</sup>	57.51	28.25	65.30	28.53
Raven's matrices (max 18) <sup>***</sup>	9.29	2.78	8.34	2.53
Digit span (number of digits) <sup>***</sup>	5.34	1.06	4.12	0.83
Categorical fluency (number of items)	14.93	3.76	14.58	3.58
Proverb understanding (max 15) <sup>***</sup>	6.93	3.71	4.44	3.47
Vocabulary (max 20) <sup>***</sup>	13.85	3.11	11.94	2.55

Q5 \* $p < .1$ ; \*\* $p < .05$ ; \*\*\* $p < .01$ .

after corrections for multiple comparisons. There was no difference between groups in categorical fluency.

Table 2 describes potential determinants of cognitive test performance broken down by ethnic group. Results of one-way analyses of variance (for age and number in-household) and chi-squared tests (for all other variables) show that the Wolof and

**Table 2.** Potential determinants of cognitive test performance in Mandinka and Wolof ethnic groups

Variable	Mandinka (N = 418)		Wolof (N = 144)	
	Mean	SD	Mean	SD
Age	17.09	1.34	16.98	1.34
Number in-household	9.35	5.35	9.92	5.89
	N	%	N	%
Female	205	49	68	47
Tester 1 <sup>***</sup>	90	22	68	47
Tester 2 <sup>***</sup>	79	19	67	47
Tester 3 <sup>***</sup>	111	27	9	6
Tester 4 <sup>***</sup>	138	33	0	0
Primary school <sup>***</sup>	169	40	38	26
Upper primary school <sup>**</sup>	139	33	31	22
Secondary school	60	14	13	9
Madrassa <sup>***</sup>	249	60	106	74
Madrassa > 4 years <sup>***</sup>	42	10	33	23
Lived in city <sup>***</sup>	202	48	33	23
Left village to study	68	16	26	18
Left village for work <sup>**</sup>	18	4	13	9
Father went to primary school <sup>a ***</sup>	82	20	77	54
Mother went to primary school <sup>***</sup>	17	4	50	35
Father alive <sup>*</sup>	354	85	113	78
Mother alive	404	97	136	94
Family own cart <sup>b ***</sup>	277	73	80	62
Family sell groundnuts <sup>b ***</sup>	174	46	19	15
Household has cement/brick wall <sup>b</sup>	360	95	126	97
Household has tin roof <sup>b ***</sup>	88	23	70	54

<sup>a</sup> Missing data for 13 Mandinka and 2 Wolof.

<sup>b</sup> Missing data for 38 Mandinka and 14 Wolof.

\* $p < .1$ ; \*\* $p < .05$ ; \*\*\* $p < .01$ .

10 Matthew C. H. Jukes and Elena L. Grigorenko

Mandinka differed significantly on many variables including education, migration, parental education, and measures of family wealth. Testers were allocated to the two ethnic groups differently. Testers 1 and 2 conducted most of the testing among the Wolof communities, supported by Tester 3 in one village. Testers 1–3 spoke both languages equally well. Tester 4 did not speak Wolof and so conducted testing only in Mandinka villages, supported by the other three testers.

### **Determinants of cognitive ability**

Table 3 presents the results of six multiple regression equations. In each case, coefficient estimates represent the effect of independent variables on standardized cognitive test scores. The difference between Wolof and Mandinka remained significant in the case of two tests. We return to investigate these differences further below. Tester identity affected all test scores. The significant interaction between Tester 1 and Wolof ethnic group indicates that the Wolof performed better on the digit span test with Tester 1. All six tests were related to participants' attendance in formal school. For the visual search test, each level of education (primary school, upper primary, and secondary) was associated with an improvement in scores. For other tests, significant improvements in test scores were only seen once participants had reached a certain level of schooling. For the Raven's matrices test, there was a significant interaction between attending formal school and ethnic group. Attending school was associated with a greater improvement in scores for the Wolof than for the Mandinka.

Migration also affected test scores. Among those who had migrated to the city (the Kombos), the mean length of stay there was 2.2 years ( $SD = 3.5$ ). Participants who had lived in the city at some point in their lives scored higher on five of the six tests. Proverb understanding was improved among those who had additionally left home to work or study and categorical fluency was higher among those who had left to work. Three household-level socio-economic status variables had a positive impact on three or more tests. Scores were higher among those with educated fathers, those from larger households and from households who own a cart (an indication of wealth). Participants whose families sold groundnuts (the main indication of commercial agriculture in the region) scored more poorly on three cognitive tests.

### **Mediators of the relationship between ethnic group and cognitive test performance**

We turn now to the central question of the analyses: to what extent can ethnic group variables be explained by differences in education, migration, and background variables? Such variables which account for the relationship between predictor and criterion are known as mediators (Baron & Kenny, 1986). The above analyses suggest two candidate mediators: education and city residence. Table 4 shows the results of a mediation analysis to assess the contribution of both of these variables<sup>2</sup> to ethnic group differences in cognitive test performance. The first set of regression equations estimate ethnic group differences adjusted only for age and sex. Differences were significant in five of the six tests with the effect size ranging from 0.36 in the Raven's matrices test to 1.09 in the digit span test. When controlling for primary school attendance and city residence, the estimates of ethnic group difference were reduced substantially for the

<sup>2</sup>Analyses were also conducted on all other variables, none of which were found to be statistically significant positive mediators of the relationship between ethnic group and cognitive test performance.

**Table 3.** Unstandardized coefficient estimates from six multivariate regression equations with standardized cognitive test score as dependent variable in each case

	Visual search	Raven's matrices	Digit span	Categorical fluency	Proverb understanding	Vocabulary
Wolof	0.05 (0.28)	-0.36 (0.30)	-1.49*** (0.32)	0.25 (0.32)	-0.37 (0.31)	-0.60** (0.29)
Age (years)	-0.02 (0.02)	-0.02 (0.03)	-0.03 (0.03)	0.04 (0.03)	0.17*** (0.03)	0.01 (0.03)
Sex	-0.37*** (0.07)	-0.33*** (0.08)	-0.25*** (0.08)	-0.70*** (0.09)	0.20** (0.08)	-0.31*** (0.08)
Tester 1 <sup>a</sup>	0.23** (0.10)	-0.09 (0.11)	0.27** (0.12)	0.34*** (0.12)	-0.32*** (0.12)	-0.48*** (0.11)
Tester 2 <sup>a</sup>	0.02 (0.11)	-0.67*** (0.12)	-0.06 (0.12)	0.17 (0.12)	-0.38*** (0.12)	-0.18 (0.11)
Tester 3 <sup>a</sup>	0.37*** (0.09)	-0.74*** (0.10)	0.38*** (0.11)	0.70*** (0.11)	-0.66*** (0.11)	-0.17* (0.10)
Tester 1 × Wolof	-0.19 (0.29)	0.09 (0.32)	0.77** (0.33)	-0.42 (0.34)	-0.06 (0.33)	0.49 (0.31)
Tester 2 × Wolof	-0.26 (0.29)	0.07 (0.32)	0.42 (0.33)	-0.54 (0.34)	-0.24 (0.33)	-0.01 (0.31)
School <sup>b</sup>	0.31** (0.14)	0.18 (0.15)	-0.18 (0.16)	-0.22 (0.16)	-0.24 (0.16)	0.03 (0.15)
School × Wolof	0.03 (0.16)	0.38** (0.18)	-0.26 (0.19)	0.06 (0.19)	0.24 (0.19)	0.22 (0.17)
Upper primary	0.68*** (0.14)	0.25 (0.16)	0.45*** (0.16)	0.31* (0.17)	0.30* (0.16)	0.52*** (0.15)
Secondary	0.50*** (0.12)	0.76*** (0.13)	0.12 (0.14)	0.00 (0.14)	0.14 (0.14)	0.41*** (0.13)
In Madrasa > 4 years	0.17 (0.10)	-0.02 (0.11)	-0.01 (0.12)	0.17 (0.12)	-0.12 (0.12)	0.03 (0.11)
Ever in city	0.20*** (0.07)	0.13* (0.08)	0.20** (0.08)	0.11 (0.08)	0.30*** (0.08)	0.26*** (0.07)
Left to study	-0.04 (0.10)	0.06 (0.11)	0.03 (0.11)	0.14 (0.12)	0.25** (0.11)	0.17 (0.11)
Left for a job	-0.13 (0.15)	0.05 (0.16)	-0.25 (0.17)	0.37** (0.18)	0.43** (0.17)	0.13 (0.16)
Mother in school	-0.07 (0.11)	0.01 (0.13)	-0.17 (0.13)	-0.02 (0.13)	-0.25* (0.13)	-0.03 (0.12)
Father in school	-0.05 (0.08)	-0.07 (0.09)	0.17* (0.09)	0.21** (0.10)	0.20** (0.09)	0.03 (0.09)
Household size	0.02*** (0.01)	0.00 (0.01)	0.00 (0.01)	0.02** (0.01)	0.01** (0.01)	0.01 (0.01)
Owens cart	0.06 (0.08)	-0.03 (0.09)	0.16* (0.09)	0.03 (0.09)	0.19** (0.09)	0.25*** (0.08)
Sells groundnuts	0.03 (0.08)	-0.17** (0.08)	-0.18** (0.09)	-0.01 (0.09)	0.02 (0.09)	-0.18** (0.08)
Tin roof	-0.02 (0.08)	-0.05 (0.09)	-0.04 (0.09)	0.02 (0.09)	-0.09 (0.09)	-0.16* (0.08)
Cement/brick wall	-0.10 (0.17)	0.05 (0.18)	0.21 (0.19)	0.27 (0.19)	-0.33* (0.19)	-0.19 (0.17)
Constant	-0.21 (0.47)	0.64 (0.51)	0.38 (0.53)	-1.24** (0.55)	-2.75*** (0.53)	0.07 (0.49)
R <sup>2</sup>	.473	.378	.335	.287	.341	.388
F(23,538)	21.03	14.24	11.77	9.42	12.09	14.84

<sup>a</sup> Reference – Tester 4.

<sup>b</sup> Primary school versus Madrasa.

\*p < .1; \*\*p < .05; \*\*\*p < .01; standard errors are in brackets.

**Table 4.** Unstandardized coefficient estimates for equations estimating effect of ethnic group on cognitive test performance before (Equation (1)) and after (Equation (2)) mediators are included in analyses

	Visual search	Raven's matrices	Digit span	Categorical fluency	Proverb understanding	Vocabulary
<b>Equation (1)</b>						
Wolof	-0.40*** (0.09)	-0.36*** (0.09)	-1.09*** (0.09)	-0.10 (0.09)	-0.63*** (0.09)	-0.61*** (0.09)
Age (years)	-0.03 (0.03)	0.00 (0.03)	-0.03 (0.03)	0.05* (0.03)	0.23*** (0.03)	0.03 (0.03)
Female	-0.67*** (0.08)	-0.47*** (0.08)	-0.25*** (0.07)	-0.82*** (0.08)	0.11 (0.08)	-0.53*** (0.08)
Constant	0.95* (0.50)	0.32 (0.52)	0.89* (0.48)	-0.47 (0.49)	-3.81*** (0.49)	-0.08 (0.49)
R <sup>2</sup>	.146	.080	.237	.177	.180	.150
F(3,558)	31.79	16.07	57.74	40.01	40.85	32.86
<b>Equation (2)</b>						
Wolof	-0.20** (0.08)	-0.20** (0.09)	-1.02*** (0.09)	-0.07 (0.09)	-0.50*** (0.09)	-0.42*** (0.08)
Age (years)	-0.00 (0.03)	0.01 (0.03)	-0.03 (0.03)	0.05 (0.03)	0.21*** (0.03)	0.04 (0.03)
Female	-0.40*** (0.07)	-0.30*** (0.08)	-0.22*** (0.08)	-0.82*** (0.08)	0.14* (0.08)	-0.36*** (0.07)
School	1.01*** (0.07)	0.69*** (0.08)	0.15* (0.08)	0.02 (0.08)	0.20** (0.08)	0.72*** (0.08)
Ever in city	0.20*** (0.07)	0.23*** (0.08)	0.17** (0.08)	0.11 (0.08)	0.41*** (0.08)	0.31*** (0.07)
Constant	-0.19 (0.44)	-0.38 (0.50)	0.83* (0.49)	-0.41 (0.51)	-3.75*** (0.49)	-0.76 (0.46)
R <sup>2</sup>	.375	.191	.248	.180	.226	.289
F(5,556)	66.74	26.22	36.62	24.40	32.50	45.22
<b>Mediation effects</b>						
School	37.9%***	29.4%***	2.1%	4.00%	5.1%**	18.0%***
Ever in city	14.3%**	17.23%**	3.9%	26.8%**	16.2%***	13.6%***

\*p < .1; \*\*p < .05; \*\*\*p < .01; standard errors are in brackets.

visual search and Raven's matrices tests, reduced somewhat for the proverb understanding and vocabulary test but were not significantly reduced for the digit span test. Statistics at the bottom of Table 4 show the percentage of ethnic group difference in test performance attributable to the two mediators. School attendance was a substantial mediator of the effect of ethnic group on visual search and Raven's matrices test performance accounting, respectively, for 37.9 and 29.4% of the effect. Schooling also significantly mediated the effect of ethnic group on proverb understanding and vocabulary. For all four of these tests, the mediation effect of living in the city remained at around 15%, statistically significant in each case.

## Discussion

Results show substantial differences in performance on five of six cognitive tests between the two ethnic groups in this study, with the Mandinka group outperforming the Wolof group in each case. Analyses suggested that two variables, indicators of education and city residence, partly mediated the effect of ethnic group on performance in four of the five tests. These results fit our hypothesis that the cognitive abilities of the two groups differ *because of* their different experiences. However, it is impossible to establish causal relationships from our cross-sectional data. Other interpretations are possible. For example, it may be that participants with higher cognitive abilities are more likely to be enrolled in, and succeed in education. They may also be more likely to seek out opportunities afforded by urban life. Consequently, we note that these data are consistent with, but do not confirm our hypotheses.

The fifth test, the digit span test, differed from other assessments in two respects. The ethnic group difference was largest for this test than for the other tests and was not attenuated after controlling for potential mediators. This observation is examined more closely in Study 2.

## STUDY 2

One potential explanation for the poorer short-term recall for numbers among the Wolof (captured by the digit span test) is that they use a base five counting system. This means that the words used for numbers 6–9 are longer and thus fewer can be rehearsed in the phonological loop (Baddeley, 1992). For example, the Wolof equivalent of *three, one, five-three* (*ñett, benn, juróom-ñett*) takes longer to rehearse than the Mandinka equivalent of *three, one, eight* (*saba, kiling, sey*). The aim of Study 2 was to assess if this hypothesis explained ethnic differences in digit span performance. The hypothesis implies that ethnic group differences in recall would be greater for sequences of numbers ranging from 1 to 9 than for number ranging 1–5.

## Methods

### Participants

Forty-one adolescents took part in this study, with mean age 14.4 years (range = 12–18). Twenty-one participants were Mandinka, 20 were Wolof. There were 30 females and 11 males. Participants were recruited in Farafenni town where they all lived except for 3 participants from nearby villages.

14 Matthew C. H. Jukes and Elena L. Grigorenko

### Procedure

All testing was conducted in the Medical Research Centre compound in Farafenni by the four cognitive testers from Study 1. All participants were given two versions of the digit span test. One version contained the numbers 1–9 and was the same test used in Study 1. A second version contained only numbers 1–5. The order in which participants received the two versions of the tests was varied randomly. There were three items at each level of difficulty, from a three digit span to a nine digit span. Digit span is calculated as  $(\text{total correct})/3 + 2$ .

### Analysis

All analyses were conducted in Stata (StataCorp, 2007). A multiple regression equation estimated the impact of language and test version on digit span. Dummy variables were entered for Mandinka recall of the nine-number version, Wolof recall of the five-number version and Wolof recall of the nine-number version, with Mandinka recall of the five-number version acting as the comparison group.

### Results and discussion

As illustrated in Figure 1, digit span was at a similar level for both versions of the test among the Mandinka group and for the five-number version in the Wolof group. The Wolof group's digit span was smaller for the nine-number version of the test. These observations were supported by the results of the regression analysis (Table 5). Wolof recall of numbers 1–9 was the only recall category to differ significantly from the reference category (Mandinka recall of numbers 1–5). The difference was substantial with digit span reduced by almost two digits.

These results are consistent with the notion that the digit span test was biased in favour of the Mandinka. However, mean differences in performance are not sufficient by

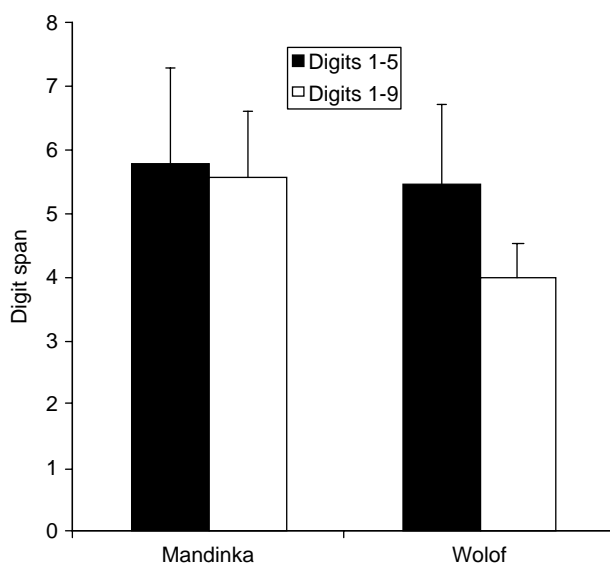


Figure 1. Digit span for recall of digits 1–5 or 1–9 by Wolof and Mandinka ethnic groups.

**Table 5.** Unstandardized coefficient estimates for the effect of digit range and ethnic group on digit span

	Digit span
Age (years)	0.07 (0.13)
Female	0.10 (0.42)
Mandinka digits 1–9 <sup>a</sup>	–0.21 (0.16)
Wolof digits 1–5 <sup>a</sup>	–0.33 (0.38)
Wolof digits 1–9 <sup>a</sup>	–1.83*** (0.38)
Constant	4.56** (2.18)
R-squared	
Within	.68
Between	.18
Overall	.29
Wald chi-squared (5)	89.26***

<sup>a</sup> Reference Mandinka digits 1–5; \* $p < .1$ ; \*\* $p < .05$ ; \*\*\* $p < .01$ ; standard errors are in brackets.

themselves to establish bias. Performance differences may alternatively arise from genuine differences in memory abilities, although this explanation does not fit well with the comparable recall of the digits 1–5 by Wolof and Mandinka. Nevertheless, it is possible that the mean differences in recall of this study exaggerated the measurement bias of the digit span test. The degree of measurement bias cannot be established without further work to model the latent trait of working memory in the two groups.

Notwithstanding the limitations of our study in identifying testing bias, the overall results are consistent with the hypothesis that numeric representation differences help explain the larger digit span of the Mandinka group in comparison with the Wolof group, because recall among the Wolof was only poorer for items containing numbers 6–9. These numbers contain two number-words in Wolof but not in Mandinka.

## GENERAL DISCUSSION

In this article, we presented evidence of significant differences in cognitive test performance between the Wolof and Mandinka ethnic groups in the Gambia. We hypothesized that these differences would be partly mediated by exposure to urban culture, education, language, and socio-economic background. Study 1 provided support for the first two of these factors, with education and city residence being found to be significant mediators of the relationship between ethnic group and cognitive test performance. Study 2 provided support for the hypothesis that difference in the languages was an explanation for group differences in cognitive test performance. Findings related only to the digit span test and suggested that the increased word length found in the base-5 Wolof counting system helped explain why recall was poorer in this group. We found no evidence of other socio-economic variables mediating the effect of ethnic group on cognitive test performance.

Our findings highlight many problems of comparing cognitive test scores in different populations. The results of the digit span test indicate that linguistic factors bias test scores for some populations. This is one reason why the analysis of IQ scores across populations and countries (Lynn & Vanhanen, 2001; Rushton & Jensen, 2005) is unlikely to constitute a valid comparison.

Our results also have implications for the argument that the reported mean IQ in sub-Saharan Africa of around 70 (Lynn, 1991) is a causal factor in these countries' low economic productivity and that differences from IQ levels of Europeans and Asians are 'largely genetically determined' (Lynn & Vanhanen, 2001, p. 431). Our study did not compare participants across continents but our finding that group differences in cognitive ability can be attributed to the environmental variables of schooling and urban residence points to the potential importance of these factors in explaining group-level IQ differences. The finding lends credence to the view that economic productivity may, in turn, influence IQ (Morse, 2008) through expanded access to schooling and increased urbanization and emphasizes that the correlations found between IQ and economic productivity at the national level tell us little about whether IQ affects productivity or vice versa. It is also interesting to note that our findings for the Raven matrices test, the test used in many studies reporting IQ scores for sub-Saharan Africans, suggest that almost 50% of intergroup differences could be attributed to the effects of schooling and urban residence. Our interpretation of these findings is that cognitive abilities are higher in arenas, such as in school and in cities, where these abilities are adaptive and their development is fostered. Our results raise the possibility that sub-Saharan Africa's historically low levels of formal schooling (e.g. 64% gross enrolment ratio in primary schools in 1990; UNESCO Institute of Statistics, 2009) and low levels of urbanization (estimated 37% urban in 2010 compared to 75% in more developed countries; UN, 2007) are leading explanations for the continent's poor performance on Western IQ tests, which are designed in and for the context of schooling in industrialized societies.

There are further implications of these results which differ for each of the cognitive tests used. We will consider first the non-verbal tests. There is a tradition of viewing non-verbal tests as being 'culture free' or 'culture fair' (Rosselli & Ardila, 2003). However, our findings support other recent evidence suggesting that education level, cultural background, and urban residence are all associated with performance on non-verbal tests (Rosselli & Ardila, 2003). Thus, there are no 'culture free' or 'culture fair' tests.

Notwithstanding the explanatory power of schooling and place of residence, significant differences in ethnic group performance remained after controlling for these two factors. Several explanations are possible. Unobserved cultural or socio-economic differences may have influenced cognitive test performance. Another possible explanation for difference in test performance is tester identity. Because all testers were Mandinka, testing may have been biased in favour of participants from the same ethnic group. Our results show that performance varied from one cognitive tester to another and that tester identity affected the two ethnic groups differently. However, without any Wolof testers we were unable to test the hypothesis that test scores were higher when the tester was from the participant's ethnic group.

For the digit span test our findings suggest that language, and particularly the counting system, influenced performance to the greatest extent. We found no evidence of either education or city residence significantly affecting this test, despite these factors being documented as influences on the digit span test elsewhere (Ostrosky-Solis & Lozano, 2006). Perhaps, the magnitude of the language effect (the Wolof recalled almost 2 fewer digits in Study 2) masked other influences on test performance.

Categorical fluency was the only test for which no ethnic group differences were found. This was in part because neither education nor city residence was significantly associated (at the 5% level) with performance on this test. Previous studies have found that the relationship between education and categorical fluency is dependent on the semantic criteria for the task. da Silva, Petersson, Faisca, Ingvar, and Reis (2004) found

that there was no relationship between education and categorical fluency for target items with which literate and illiterate participants had similar experience. Perhaps, the target items in our test (i.e. animals) were equally familiar to both groups and provided the most 'culture fair' assessment.

The two language-based tests were developed in parallel in the two languages. Like the non-verbal tests discussed above, the proverb understanding test and the vocabulary test were both associated with education and city residence, with both factors partly mediating ethnic group differences in this test. In addition, the proverb understanding was, uniquely among the six tests, associated with leaving one's village to study or to work, and by paternal education and family wealth, perhaps reflecting the range of situations outside school where one encounters the use of proverbs. However, ethnic group differences in both language-based tests were substantial and significant even after controlling for explanatory variables. In all likelihood these tests were not well matched for difficulty during their development in the two languages. This reflects the difficulty in using cognitive tests for countries and languages where normative data have not been collected. Inherent properties of the language, unobserved cultural factors and tester identity may have also played a part in these differences.

What relevance do these results have for the assessment of cognitive abilities in Africa and other multiethnic regions of the world? The answer to this question depends in part on the importance of comparability across ethnic groups in assessment exercises. For some uses of cognitive assessments, comparability of outcomes is not essential. This is the case when cognitive tests are used as intervention evaluations, such as the early childhood malaria prevention study giving rise to the data reported in this article. Here, the focus is less on whether tests are culture-fair but whether improvements found in cognitive test scores represent development of cognitive abilities which influence life-relevant outcomes. Such questions of external validity were not directly addressed in the current study.

When the comparability of outcomes is important, for example, when regional analyses of cognitive impairment are carried out, our findings suggest a number of possible approaches to conduct valid comparisons. First, sampling strategies could involve stratification by urban residence and education levels in order to minimize the impact of these variables on cognitive outcomes. Second, assessment batteries could focus on cognitive tests that are less sensitive to levels of education and urban residence. Results from this study and others (Ardila, 2005; Ardila *et al.*, 2006) suggest that the categorical fluency test, particularly when requiring participants to recite the names of animals, meets these criteria. Third, where assessments are conducted among rural (and/or illiterate) groups, they should aim to assess competences valued and nurtured by these communities. This has proved challenging in the past, but some progress has been made in developing reliable measures of the development of social responsibility in the Gambia (Jukes, Grigorenko, & Sternberg, in preparation) requiring community members to make comparisons amongst people they know in terms of their competence in this domain (Grigorenko *et al.*, 2001). Fourth, careful piloting and psychometric analysis is required in order to develop comparable tests of verbal ability in different languages (Grigorenko *et al.*, 2009; Stemler *et al.*, 2009). Fifth, a careful analysis of differences in language structure is required. This article provides one example - in the digit span test - where language differences can affect the comparability of cognitive test performance. Sixth, it is important to consider the ethnic group to which testers belong. A requirement that they are from the same group as those being assessed presents great challenges. There are difficulties in identifying and

18 Matthew C. H. Jukes and Elena L. Grigorenko

training cognitive testers in countries with little history of psychometric assessment. Developing a team of testers which an ethnic profile representative of the target population adds to these difficulties.

In conclusion, we believe that cognitive assessment has many important roles to play in the development of education in Africa and in other regions of the world, from needs assessment to intervention evaluation. There is much work to do, but the results of this study and their implications suggest a possible way forward.

### Acknowledgements

This research was funded by the Partnership for Child Development and by grants from the Medical Research Council and the Wellcome Trust. We would like to thank Margaret Pinder, Gijs Walraven, and the field research team at Medical Research Council, Farafenni, The Gambia: Elisa Bariau, Helen Baños Smith, Alasan Camara, Dodou Trawally, Sheriff Kanyi, Lamin Janneh, Abass Sillah, Lamin Colley, Saja Sey, and Jarai Bangura for their hard work and Omar Mboob and Momodou Touray for many insights into the Wolof and Mandinka of our study region. We would also like to thank the people of Alkali Kunda, Bambali, Bassik, Dibba Kunda, Jajari Kunda, Jumansare Ba, Kani Kunda Wharftown, Kani Kunda Suba, Ker Bamba Lowe, and Kumbija for their participation in the study. Finally, we would like to thank Jodi Reich and Kelly Nedwick of Yale University for their linguistic advice.

### References

- Alexander, N. (1995). Models of multilingual schooling in South Africa. In K. Heugh, A. Siegruhn, & P. Pluddemann (Eds.), *Multilingual education for South Africa* (pp. 79–82). Johannesburg: Heinemann.
- Ardila, A. (2005). Cultural values underlying psychometric cognitive testing. *Neuropsychology Review*, *15*(4), 185–195.
- Ardila, A., Ostrosky-Solis, F., & Bernal, B. (2006). Cognitive testing toward the future: The example of semantic verbal fluency (ANIMALS). *International Journal of Psychology*, *41*(5), 324–332.
- Artiola i Fortuny, L., Garolera, M., Romo, D. H., Feldman, E., Barillas, H. F., Keefe, R., *et al.* (2005). Research with Spanish-speaking populations in the United States: Lost in the translation: A commentary and a plea. *Journal of Clinical and Experimental Neuropsychology*, *27*(5), 555–564.
- Asmara Declaration (2000). *The Asmara declaration on African languages and literatures*. Paper presented at the Against All Odds: African Languages and Literatures into the 21st Century Conference, Asmara, Eritrea.
- Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556–559.
- Baddeley, A., Gardner, J. M., & Grantham McGregor, S. (1995). Cross-cultural cognition: Developing tests for developing countries. *Applied Cognitive Psychology*, *9*, S173–S195.
- Banks, M. (1996). *Ethnicity: Anthropological constructions*. London: Routledge.
- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173–1182.
- Berry, J. W. (1967). Independence and conformity in subsistence-level societies. *Journal of Personality and Social Psychology*, *7*(4), 415–418.
- Bissiliat, J., Laya, D., Pierre, E., & Pidoux, C. (1967). La notion de lakkal dans la culture Djerma-Songhai. *Psychopathologie Africaine*, *3*, 207–264.
- Brown, R. T., Reynolds, C. R., & Whitaker, J. S. (1999). Bias in mental testing since bias in mental testing. *School Psychology Quarterly*, *14*(3), 208–238.
- Camara, M. (1999). *Parlons Malinké*. Montreal: L'Harmattan.

- Carter, J. A., Lees, J. A., Murira, G. M., Gona, J., Neville, B. G. R., & Newton, C. R. J. C. (2005). Issues in the development of cross-cultural assessments of speech and language for children. *International Journal of Language and Communication Disorders*, 40(4), 385–401.
- Ceci, S. J. (1991). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Developmental Psychology*, 27(5), 703–722.
- Chumbow, B. S. (2005). The language question and national development in Africa. In T. Mkandawire (Ed.), *African intellectuals: Rethinking politics, language, gender and development* (pp. 165–192). Dakar and London: CODE SRIA and Zed Books.
- Claassen, N. C. W. (1997). Cultural differences, politics and test bias in South Africa. *European Review of Applied Psychology*, 47, 297–307.
- Clarke, S. E., Jukes, M. C. H., Njagi, J. K., Khasakhala, L., Cundill, B., Otido, J., et al. (2008). Effect of intermittent preventive treatment of malaria on health and education in schoolchildren: A cluster-randomised, double-blind, placebo-controlled trial. *Lancet*, 372(9633), 127–138.
- da Silva, C. G., Petersson, K. M., Faisca, L., Ingvar, M., & Reis, A. (2004). The effects of literacy and education on the quantitative and qualitative aspects of semantic verbal fluency. *Journal of Clinical and Experimental Neuropsychology*, 26(2), 266–277.
- Dasen, P., Barthelemy, D., Kan, E., Kouame, K., Douada, K., Adjei, K., et al. (1985). N'glouele, l'intelligence chez la Baoule. *Archives de Psychologie*, 53, 295–324.
- Elliott, J. M. (1992). Forward digit span and articulation speed for Malay, English, and two Chinese dialects. *Perceptual and Motor Skills*, 74(1), 291–295.
- Fortes, M. (1938). Social and psychological aspects of education in Taleland. *Africa*, 11(4), 1–64.
- Foxcroft, C. D. (1997). Psychological testing in South Africa: Perspectives regarding ethical and fair practices. *European Journal of Psychological Assessment*, 13, 229–235.
- Georgas, J., Weiss, L. G., van der Vijver, F. J. R., & Saklofske, D. H. (Eds.), (2003). *Culture and children's intelligence: Cross-cultural analysis of the WISC-III*. London: Academic Press.
- Q2 Gordon, R. G., Jr. (Ed.), (2005). *Ethnologue: Languages of the world* (15th ed.), (Vol. 2008).
- Q2 Dallas, TX: SIL International.
- Greenfield, P. M. (1966). On culture and conservation. In J. Bruner, R. Oliver, & P. Greenfield (Eds.), *Studies in cognitive growth* (pp. 327–334). New York: Wiley.
- Greenfield, P. M. (1997). You can't take it with you - why ability assessments don't cross cultures. *American Psychologist*, 52(10), 1115–1124.
- Greenfield, P. M., Reich, L. C., & Oliver, R. (1966). On culture and equivalence. In J. Bruner, R. Oliver, & P. Greenfield (Eds.), *Studies in cognitive growth* (pp. 270–318). New York: Wiley.
- Grieve, K. (2005). Factors affecting assessment results. In C. Foxcroft & G. Roodt (Eds.), *An introduction to psychological assessment in the South African context* (2nd ed. pp. 224–241). Cape Town: Oxford University Press.
- Grigorenko, E. L., Geissler, P., Prince, R., Okatcha, F., Kenny, D., Bundy, D., et al. (2001). The organisation of Luo conceptions of intelligence: A study of implicit theories in a Kenyan village. *International Journal of Behavioural Development*, 25(4), 367–378.
- Grigorenko, E. L., Nedwick, K., Kwadade, D., Boro, E., Hart, L., Newman, T., et al. (2009). Assessing mother tongue in the era of globalization: Promise and challenge. In E. L. Grigorenko (Ed.), *Multicultural psychoeducational assessment* (pp. 187–212). New York: Springer.
- Grigorenko, E. L., Sternberg, R. J., Jukes, M. C. H., Alcock, K., Lambo, J., Ngorosho, D., et al. (2006). Effects of antiparasitic treatment on dynamically and statically tested cognitive skills over time. *Journal of Applied Developmental Psychology*, 27(6), 499–526.
- Heugh, K. (1995). The multilingual school: Modified dual medium. In K. Heugh, A. Siegruhn, & P. Pluddemann (Eds.), *Multilingual education for South Africa* (pp. 83–88). Johannesburg: Heinemann.
- Holding, P. A., Stevenson, J., Peshu, N., & Marsh, K. (1999). Cognitive sequelae of severe malaria with impaired consciousness. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 93(5), 529–534.
- Q3 Irvine, S. H. (1969). Figural tests of reasoning in Africa: Studies in the use of Raven's progressive matrices across cultures.

20 Matthew C. H. Jukes and Elena L. Grigorenko

- Jukes, M. C. H., Drake, L. J., & Bundy, D. A. P. (2008). *School health, nutrition and education for all: Levelling the playing field*. Wallingford, CT: CABI Publishing.
- Jukes, M. C. H., Grigorenko, E. L., & Sternberg, R. J. (in preparation). The influence of urban migration and schooling on parents' development goals in the Gambia.
- Q4 Jukes, M. C. H., Nokes, C. A., Alcock, K. J., Lambo, J., Kihamia, C., Mbise, A., *et al.* (2002). Heavy schistosomiasis associated with poor short-term memory and slower reaction times in Tanzanian school children. *Tropical Medicine and International Health*, 7(2), 104-117.
- Jukes, M. C. H., Pinder, M., Grigorenko, E. L., Smith, H. B., Walraven, G., Bariau, E. M., *et al.* (2006). Long-term impact of malaria chemoprophylaxis on cognitive abilities and educational attainment: Follow-up of a controlled trial. *PLoS Clinical Trials*, 1(4), e19.
- LeVine, R. A., Dixon, S., LeVine, S., Richman, A., Leiderman, P. H., Keefer, C. H., *et al.* (1993). *Child care and culture: Lessons from Africa*. Cambridge: Cambridge University Press.
- Luckett, K. (1995). National additive bilingualism: Towards a language plan for South African education. In K. Heugh, A. Siegruhn, & P. Pluddemann (Eds.), *Multilingual education for South Africa* (pp. 73-78). Johannesburg: Heinemann.
- Lynn, R. (1991). Race differences in intelligence: A global perspective. *Mankind Quarterly*, 31, 255-296.
- Lynn, R., & Vanhanen, T. (2001). National IQ and economic development: A study of eighty-one nations. *Mankind Quarterly*, 41, 415-435.
- MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. New York: Erlbaum.
- Manly, J. J., Byrd, D. A., Touradji, P., & Stern, Y. (2004). Acculturation, reading level, and neuropsychological test performance among African American elders. *Applied Neuropsychology*, 11(1), 37-46.
- Manly, J. J., Jacobs, D. M., Touradji, P., Small, S. A., & Stern, Y. (2002). Reading level attenuates differences in neuropsychological test performance between African American and White elders. *Journal of the International Neuropsychological Society*, 8(3), 341-348.
- Manly, J. J., Miller, S. W., Heaton, R. K., Byrd, D., Reilly, J., Velasquez, R. J., *et al.* (1998). The effect of African-American acculturation on neuropsychological test performance in normal and HIV-positive individuals. *Journal of the International Neuropsychological Society*, 4(3), 291-302.
- Morse, S. (2008). The geography of tyranny and despair: Development indicators and the hypothesis of genetic inevitability of national inequality. *Geographical Journal*, 174, 195-206.
- Mulenga, K., Ahonen, T., & Aro, M. (2001). Performance of Zambian children on the NEPSY: A pilot study. *Developmental Neuropsychology*, 20, 375-383.
- Munroe, R. H., Munroe, R. L., & LeVine, R. A. (1972). Africa. In F. L. K. Hsu (Ed.), *Psychological anthropology*. Cambridge: Schenkman Publishing Company.
- Q2 Murray, A., & Jones, D. M. (2002). Articulatory complexity at item boundaries in serial recall: The case of Welsh and English digit span. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(3), 594-598.
- Mwamwenda, T. S. (1992). Cognitive development in African children. *Genetic Social and General Psychology Monographs*, 118(1), 5-72.
- Naveh-Benjamin, M., & Ayres, T. J. (1986). Digit span, reading rate, and linguistic relativity. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, 38(4), 739-751.
- Ngom, F. (2003). *Wolof*. Munich: Lincom Europa.
- Ostrosky-Solis, F., & Lozano, A. (2006). Digit span: Effect of education and culture. *International Journal of Psychology*, 41(5), 333-341.
- Owen, K. (1991). Test bias: The validity of the junior aptitude tests (JAT) for various population groups in South Africa regarding constructs measured. *South African Journal of Psychology*, 21, 112-118.
- Owen, K. (1992). The suitability of Raven standard progressive matrices for various groups in South Africa. *Personality and Individual Differences*, 13(2), 149-159.
- Q2 Owino, F. R. (Ed.), (2002). *Speaking African: African languages for education and development*. Cape Town: CASAS.

- Parker, S., Greer, S., & Zuckerman, B. (1988). Double jeopardy - the impact of poverty on early child development. *Pediatric Clinics of North America*, 35(6), 1227-1240.
- Pena, E. D. (2007). Lost in translation: Methodological considerations in cross-cultural research. *Child Development*, 78(4), 1255-1264.
- Raven, J. C., Styles, I., & Raven, M. A. (1998). *Raven's progressive matrices: CPM parallel test booklet*. Oxford: Oxford Psychologists Press.
- Rosselli, N., & Ardila, A. (2003). The impact of culture and education on non-verbal neuropsychological measurements: A critical review. *Brain and Cognition*, 52(3), 326-333.
- Rowlands, E. C. (1959). *A Grammar of Gambian Mandinka*. London: University of London.
- Royston, P. (2004). Multiple imputation of missing values. *Stata Journal*, 4, 227-241.
- RTI (2008). *Early grade reading assessment (EGRA)*. 2nd Workshop, Washington, DC, USA.
- Rushton, J. P., & Jensen, A. R. (2005). Thirty years of research on race differences in cognitive ability. *Psychology, Public Policy, and Law*, 11(2), 235-294.
- Serpell, R. (1993). *The significance of schooling: Life-journeys in an African society*. Cambridge: Cambridge University Press.
- Shebani, M. F. A., van de Vijver, F. J. R., & Poortinga, Y. H. (2005). A strict test of the phonological loop hypothesis with Libyan data. *Memory and Cognition*, 33(2), 196-202.
- Sinha, G. (1988). Exposure to industrial and urban environments and formal schooling as factors in psychological differentiation. *International Journal of Psychology*, 23(6), 707-719.
- StataCorp (2007). *Stata statistical software: Release 10*. College Station, TX: StataCorp LP.
- Stemler, S. E., Chamvu, F., Chart, H., Jarvin, L., Jere, J., Hart, L., et al. (2009). Assessing competencies in reading and mathematics in Zambian children. In E. L. Grigorenko (Ed.), *Multicultural psychoeducational assessment* (pp. 157-186). New York: Springer.
- Super, C. M. (1983). Cultural variation in the meaning and uses of children's intelligence. In J. B. Deregowski, S. Dziurawiec, & R. C. Annis (Eds.), *Expiscations in cross-cultural psychology* (pp. 199-212). Lisse: Swets and Zeitlinger.
- Terrell, F., Terrell, S. L., & Taylor, J. (1981). Effects of race of examiner and cultural mistrust on the WAIS performance of Black students. *Journal of Consulting and Clinical Psychology*, 49, 750-751.
- Trudell, B., & Schroeder, L. (2007). Reading methodologies for African languages: Avoiding linguistic and pedagogical imperialism. *Language, Culture and Curriculum*, 20(3), 165-180.
- UN (2007). *World population prospects: The 2006 revision and World Urbanization Prospects*. Retrieved May 5, 2009, from <http://esa.un.org/unup/>
- UNESCO (2003). *EFA global monitoring report 2003/4: Gender and education for all. The leap to equality*. Paris: UNESCO.
- UNESCO (2007). *EFA global monitoring report 2007: Strong foundations. Early childhood care and education*. Paris: UNESCO.
- UNESCO Institute of Statistics (2009). *UNESCO's world education indicators*. Retrieved May 5, 2009, from <http://www.unesco.org/education/information/wer/htmlENG/tablesmenu.htm>
- UNPD (2005). *Population, resources, environment and development*. Retrieved from <http://unstats.un.org/pop/>
- van de Vijver, F. J. R., & Phalet, K. (2004). Assessment in multicultural groups: The role of acculturation. *Applied Psychology - An International Review - Psychologie Appliquee-Revue Internationale*, 53(2), 215-236.
- Wechsler, D. (1997). *Wechsler Adult Intelligence Scale-III*. San Antonio, TX: The Psychological Corporation.
- Weisner, T. S. (1976). Urban-rural differences in African children's performance on cognitive and memory tasks. *Ethos*, 4(2), 223-250.
- Wober, M. (1972). Culture and the concept of intelligence: A case in Uganda. *Journal of Cross-Cultural Psychology*, 3, 327-328.

22 Matthew C. H. Jukes and Elena L. Grigorenko

## Author Queries

*JOB NUMBER:* 746

*JOURNAL:* BJEP

- Q1** Reference Mellenbergh (1989) has been cited in text but not provided in the list. Please supply reference details or delete the reference citation from the text.
- Q2** Please supply the page ranges for the references Georgas *et al.* (2003), Gordon (2005), Munroe *et al.* (1972), and Owino (2002).
- Q3** Please provide complete details for reference Irvine (1969).
- Q4** Please update the publication details for reference Jukes (in preparation).
- Q5** Please provide the significances of '\*' and '\*\*' in Table 1.
- Q6** Equations (1) and (2) are cited in the Table 4 but have not been provided. Please remove the citation or provide the equation.
- Q7** Please provide the significance of '\*' in Table 5.