The measurement of expressive vocabulary in school-age children: Development and application of the Kilifi Naming Test (KNT)

The measurement of expressive vocabulary in school-age children: Development and application of the Kilifi Naming Test (KNT)

Abstract

The dearth of locally developed measures of language makes it difficult to detect language and communication problems among school-age children in sub-Saharan African settings. We sought to describe variability in vocabulary acquisition as an important element of global cognitive functioning. Our primary aims were to establish the psychometric properties of an expressive vocabulary measure, examine sources of variability, and investigate the measure's associations with non-verbal reasoning and educational achievement. The study included 308 boys and girls living in a predominantly rural district in Kenya. The developed measure, the Kilifi Naming Test (KNT), had excellent reliability and acceptable convergent validity. However, concurrent validity was not adequately demonstrated. In the final regression model, significant effects of schooling and area of residence were recorded. Contextual factors should be taken into account in the interpretation of test scores. There is need for future studies to explore the concurrent validity of the KNT further.

Key words: expressive vocabulary, Kilifi Naming Test, school-age children, resource-limited settings, reliability, contextual factors

The measurement of expressive vocabulary in school-age children: Development and application of the KNT

Few studies report the measurement of expressive vocabulary in school-age children in resource-limited settings. Furthermore, only a small number of locally developed standardised norm-referenced measures of language functions have been published for use with the multiple language groups of sub-Saharan Africa (SSA). This makes it difficult to detect language and communication problems especially among school-age children who may be wrongly diagnosed as having a general learning disability. Our current understanding of influences on vocabulary acquisition is generally limited to those linguistic and cultural contexts where standardised tests of vocabulary are available.

The multi-directional interactions between biological (internal) factors and environmental (external) inputs, couched within Bronfenbrenner's bioecological model (Bronfenbrenner, 1995), have a strong influence on children's vocabulary acquisition (Apiwattanalunggarn & Luster, 2005; Hamadani et al., 2010; Hart & Risley, 1995; Hoff, 2003; Weizman & Snow, 2001). Internal (i.e. child attributes such as age and gender) and external (e.g. availability of household resources, neighbourhood of residence and school exposure) factors, may underlie the substantial variability observed in vocabulary acquisition among children. For instance, several study findings attest to the fact that children show vast improvements in vocabulary acquisition with increasing age (Basilio, Puccini, Silva, & Pedromónico, 2005; Bates, Dale, & Thal, 1995; Vogt, Douglas, & Aussems, 2015). Although there is evidence of gender differences in vocabulary acquisition, some studies have reported that they are small and inconsistent (Hyde & Linn, 1988). On the other hand, other studies have reported consistent and stable gender differences during the early period up to age of six years (Bornstein, Hahn, & Haynes, 2004; Eriksson et al., 2012). However,

the research has not clearly established if these gender differences may be attributed to age, innate biological differences or external environmental and social factors (Bornstein et al., 2004; Burman, Bitan, & Booth, 2008; Leaper, 2002; Maccoby, 1980). The negative effects of poor nutritional status include a shortened attention span, reduced capacity (Sigman, Neumann, Carter, et al., 1989) and little energy to learn (Brown & Pollitt, 1996) resulting in lower scores on various outcomes, including vocabulary tests.

With regard to external factors, at the family level, socioeconomic status (SES) affects the manner in which adults use language with their children. Parents with more socioeconomic resources at their disposal more frequently talk with the aim of eliciting conversation, use longer sentences and a richer vocabulary than those with less (Hoff-Ginsberg, 1991; Hoff, 2003). In contrast, poorer outcomes have been reported for children living in homes with fewer resources at their disposal (Hart & Risley, 1995). For example, poorly educated parents living in crowded homes are less verbally responsive to their children, use less diverse language and their speech more frequently serves the function of directing the child's behaviour (Evans, Maxwell, & Hart, 1999; Hoff, Laursen, & Tardif, 2002). Larger socioeconomic structures such as the neighbourhoods in which children live influence children's outcomes indirectly through various proximal social contexts such as families and schools (Bronfenbrenner, 1979; Leventhal & Brooks-Gunn, 2000; Sampson, Morenoff, & Gannon-Rowley, 2002). This association varies by the extent of neighbourhood advantage (Dupéré, Leventhal, Crosnoe, & Dion, 2010) so children living in neighbourhoods with more resources are likely to have better outcomes. By the time they get to school, most of the words that children encounter in their everyday conversations are already in their vocabulary repertoires (Cunningham & Stanovich, 1998; Hayes & Ahrens, 1988). Children may however pick up new words through incidental exposure; for example, in their play

experiences in school (Connor, Morrison, & Slominski, 2006; Miller & Gildea, 1987). Although some studies suggest that additional years in school do not have a measurable impact on vocabulary growth in children especially during the early school years (Cantalini, 1987; Chall, Jacobs, & Baldwin, 1990; Christian, Morrison, Frazier, & Massetti, 2000; Skibbe, Connor, Morrison, & Jewkes, 2011), other studies have shown that since language is a socially-mediated process, teachers provide children with opportunities for vocabulary learning through their daily oral language discourse (Huttenlocher, Vasilyeva, Cymerman, & Levine, 2002). An examination of these internal and external factors will shed light on the potential influences on vocabulary acquisition.

There have been studies on language development in children in the sub-Saharan African setting; however, they are few in number, have mostly utilised small sample sizes or have relied on 'Western' instruments to measure child outcome (Carter et al., 2006; Carter et al., 2005; Demuth, 1990; Suzman, 1987). Whilst most of these studies have compared the rates of development of speech among different language groups, others have focussed on the influence of illness, nutritional supplementation and various environmental factors on various aspects of language functioning in children. For instance, a study conducted in Madagascar concluded that among 3- to 6-year-old children, there were large differences in language development between those in the top wealth quintile and those in the lowest wealth quintile, and these differences increased as children grew older (Fernald, Weber, Galasso, & Ratsifandrihamanana, 2011). Similarly, as reported in a Mozambican study, being a female, having a mother who was educated to secondary level or higher, living in an urban area and being older were positively associated with the expressive and receptive language scores of infants (Vogt et al., 2015). The findings from these studies provide evidence that while taking into account factors that may influence child development, interventions to promote

child development, and particularly language development, should begin at earlier ages and should target those who are most disadvantaged.

Other studies have reported the adaptation of measures of vocabulary among infants, pre-school and school-age children living in resource-limited settings (Alcock et al., 2015; Bortz, 1995; Childers, Vaughan, & Burquest, 2007; Holding et al., 2004; Pakendorf & Alant, 1997; Sigman, Neumann, Carter, et al., 1989). In a study among infants living in a rural setting, Alcock and others (2015) evaluated the use of an interview format of the Communicative Development Inventory (CDI) in a setting where direct language testing was impractical. The adapted tool was found suitable for use among respondents with low literacy levels, and in a setting where young children had multiple caregivers. In a similar study among infants in Nigeria, Childers and colleagues (2007) relied on a parental checklist to examine whether joint behaviours of toddlers and their caregivers were linked to the acquisition of nouns and verbs. Their study findings highlighted the importance of early contexts for the development of language. In an earlier study within the same socioeconomically-deprived setting as the current study context, Holding and colleagues (2004) administered an adapted Picture Vocabulary Test (PVT) to school-age children to examine the neuropsychological consequences of brain insults. They concluded that the modified tests retained their psychometric properties, and were sensitive to health-related and sociodemographic factors.

The forms of validity that have been tested in earlier applications of vocabulary measures among young children include establishing their relationship with age, SES (Luoni et al., 2015), non-verbal reasoning (Luoni et al., 2015; Storms, Saerens, & Deyn, 2004), reading skills (Luoni et al., 2015; Nation & Snowling, 2004), word production (Vogt et al., 2015) and clinical outcomes (Alcock, Abubakar, Newton, & Holding, 2016; Alcock,

Holding, Mung'ala-Odera, & Newton, 2008). The fact that there are few studies reported from sub-Saharan Africa demonstrates that much more effort is required in the validation of vocabulary measures that will be responsive to the cultural and language diversity in this context.

The advances made in adaptation and validation of vocabulary measures for young children are impressive. However, we are not aware of any efforts to create a standardised assessment of expressive vocabulary for school-age populations in SSA. Considering how varied and complex language is, our study did not seek to distinguish language delays and disorders; rather, we were more interested in describing variability in vocabulary acquisition as an important element of global cognitive functioning. In designing a vocabulary measure for rural school-aged children, context-specific cultural and language differences present translation difficulties. Hence it would not be valid to apply any of the available published measures, such as the Boston Naming Test (BNT; Kaplan et al., 1983) as items on tests that are intended for specific cultures may not be culturally meaningful for or familiar to populations living in other contexts; moreover, specific translations of certain items may not exist, rendering the test inappropriate (van de Vijver & Tanzer, 2004). The BNT is a widely used confrontation naming test that is an informative measure of word retrieval and productive vocabulary among children with learning disabilities, brain-damaged patients, and typically developing populations. Confrontation naming entails showing a subject one picture at a time and requiring him/her to produce the correct verbal label for the item in the picture. The response bias observed in past administration of a picture vocabulary test (Holding et al., 2004) such as children picking a picture from the same position, coupled with the problem of producing sufficient drawings recognisable to the study population, provided the impetus for developing a confrontation naming test for this age group. Moreover,

confrontation naming is particularly sensitive to subtle brain injury (Cheung, Cheung, & Chan, 2004; Jordan & Ashton, 1996), making the BNT suitable for use among an apparently normal population that may be susceptible to the effects of central nervous system infections (for e.g. malaria, meningitis and neonatal sepsis) that are endemic to the study area. And although different versions of the BNT have been widely used to investigate naming or word retrieval (Kim & Na, 2008; Miotto, Sato, Lucia, Camargo, & Scaff, 2010; Storms et al., 2004; Tatsuta et al., 2013), these past studies have highlighted the limitations in cross-cultural applicability of the test. Our primary aims were therefore to establish the psychometric properties of the developed measure of expressive vocabulary, examine sources of variability in vocabulary acquisition, and investigate associations of children's vocabulary scores with non-verbal reasoning and educational achievement.

Method

The data in the current study were derived from a larger cross-sectional study on the development of appropriate methodologies to assess executive functions, motor skills and the home environment in school-age children in a rural district. To satisfy the inclusion criteria for the main study, children had to speak one of the local dialects or Kiswahili as their first language; demonstrate physical ability to perform the tasks; and, be resident within the study area.

In this section, we first describe the development of the confrontation naming test – the Kilifi Naming Test (KNT). We then explain how we established the psychometric properties of the KNT in terms of its reliability, validity and sensitivity.

Study setting

The study was conducted in the former Kilifi District (currently known as Kilifi County) that constituted the former Coast Province in Kenya. Over 80% of the county's

inhabitants belong to the Mijikenda ethnic group, which comprises nine sub-groups with similar dialects closely related to Kiswahili, the lingua franca and country's national language. Approximately 50% of the labour force is engaged in subsistence farming mainly growing maize, cassava and cowpeas. A few engage in livestock farming of cattle, sheep and goats. Coconuts, cashew nuts and mangoes are the main cash crops (Kahuthu, Muchoki, & Nyaga, 2005). The majority of the population (70.8%) experiences absolute poverty (Commission on Revenue Allocation, 2013), manifest as limited access to basic needs, and an inability to meet the minimum cost of food and non-food items essential to sustain life (Kahuthu et al., 2005). High poverty levels are also associated with school drop-out, deteriorating health conditions and poor literacy levels (Kahuthu et al., 2005). The average literacy level in the county is 68.2% (Commission on Revenue Allocation, 2013), which compares poorly with the national rate of 83.9% (UNESCO Institute for Statistics, 2011). About 21% of the population has never attended school, with the greater proportion being women.

Family structure is largely collectivist, with extended families including members of several generations, living together in homesteads comprising four to ten mud-walled thatched houses. More recently brick-walled and iron-roofed houses have become more common. It is imperative to point out some salient features of the setting within which the study was conducted as these may have had a covert influence on the vocabulary development of children. Boys have more unstructured time than girls and most of their time is spent unsupervised outside the household. They thus have more opportunities for social play with same-gender peers (Awiti, 2011; Wenger, 1989). Girls on the other hand spend more time with older females within the homestead engaged in chores such as looking after infants and toddlers (Wenger, 1989). Such activities are unlikely to promote vocabulary

development. It is not unusual for a child growing up in this context to learn three different languages – their mother tongue, Kiswahili the national language and English, the language of instruction in elementary schools. Traditionally as in most African societies, a child is expected to be obedient, quiet and undemanding in the presence of adults and talkativeness is frowned upon (Nyasani, 1997). Children could not initiate conversations and were taught to avoid asking adults questions as it would seem as though they were challenging them and attempting displays of superior knowledge. Adults rarely engage in any play activities with children (Mbise & Kysela, 1990) as most adult-child communication is for the purpose of giving instructions (Wenger, 1989).

Participants

We recruited children from the catchment area of five local schools distributed across neighbourhoods in the district ranging from sparsely populated (64 persons/sq.km compared to the district average of 114 persons/sq.km) to densely populated (325 persons/sq.km) semi-urban areas. Children in school and out of school were included in the sample. A description of the full study sample is presented in Table 1. A total of 308 children (51% girls, n = 160) were included in the tool development (n = 100) and validation process (n = 208) of the current study. Their ages ranged between 5 and 13 years (*M*, *SD* = 9.08, 1.12). Nearly one quarter (n = 74) of the sample had linear growth retardation (or stunting) while 11.4% (n = 35) had no schooling experience. Within household status, more than one third (39.9%, n = 123) fell into the category of 'least wealthy' (Level 1). The derivation of the household wealth index is explained in a later section.

Ethical considerations

The Kenya Medical Research Institute/National Ethics Review Committee (KEMRI/NERC) provided ethical clearance for the study. Permission to visit schools was

obtained from the District Education Office. We explained the purpose of the study to the head teachers of selected schools and then sought their permission to recruit children. We also held meetings with community leaders, elders, and parents (and guardians) of selected pupils to explain the purpose of the study. We presented information regarding the study in the language with which parents were most familiar. After each meeting, a screening questionnaire was administered to parents/guardians to establish if selected children met the study's eligibility criteria. We then obtained written informed consent for their children's participation. We explained the nature of the assessments to the children prior to test administration. All the selected children assented to their participation in the study.

Procedures

Development of the confrontation naming test. We developed a confrontation naming test similar to the BNT in terms of structure, administration and scoring that would be appropriate for school-age children (eight years and above). The BNT also provided an appropriate framework for length, and was used to suggest possible categories of words. In designing the test, we had a number of objectives – that the test would: be simple and quick to administer; require no specialised equipment; and, elicit clear, responses that are easy to record. To develop the test, we followed the 4-step systematic test adaptation procedure outlined by Holding and colleagues (2009).

Step 1 – Construct definition. An extensive review of existing literature did not reveal any studies reporting the concept of confrontation naming within the sub-Saharan African context. We therefore obtained a simple definition – the ability of children to name common words depicted through pictures presented to them – from a reference book.

Step 2 – Item pool creation. We identified a list of words that would be suitable for creating a measure of vocabulary development in a rural community of school-age children.

Figure 1 summarises the procedures followed in creating the item pool and provides details of the number of participants included at each stage. We supplemented words from existing measures of child development based on previous research (Carter et al., 2006; Carter et al., 2005; Holding et al., 2004) with words obtained from surveys of children's language. In one of the surveys, 176 schooling and out-of-school children whose ages ranged between 5 and 17 years were asked to list as many words as they knew in any language (the local language, Kiswahili or English). We also considered the input of 54 community members and an 'expert panel' comprising a psychologist, a nurse, an educationist and a linguistics professional.

Step 3 – Developing the procedure and training examiners. In the next stage, we formulated semantic stimulus cues. These were phrases used to prompt the respondent to produce a correct response if they were not able to name a picture accurately. For example, if the child misperceived a saucepan as a cup, s/he would be given the cue that the item was used "for cooking"). The appropriateness and utility of these stimulus cues was then tested on a group of children by presenting the pictures first, with and then, without stimulus cues. Some of the stimulus cues were rewritten to improve clarity. In most cases, the frequency of correct responses was higher when the items were presented with stimulus cues. We then developed a list of acceptable responses for each item to reduce ambiguity in scoring as multiple possible names would make it difficult to score an item reliably.

We trained four local child development assessors to administer and score the KNT. All the assessors had completed secondary school education and were familiar with the local dialect. At the start of the training, we explained the purpose of the study, and the importance of adhering to the tool administration protocol. This training ensured that the test was consistently administered in a standardised manner by all assessors. The assessors were

provided opportunities for practice among themselves, and with non-study children under the direct supervision of the first two authors. They received feedback on how to improve their test administration techniques until they got to a level where they were well-versed with the procedures. As explained in the procedures followed for the larger study (Kitsao-Wekulo, Holding, Taylor, Abubakar, & Connolly, 2013), observations on tool administration continued until more than 90% agreement with any one of the trainers was reached by each assessor.

Step 4 – Evaluation of developed schedule. Sixty items were tested on 75 non-study children and then ordered according to frequency of correct responses. These items were then administered to the first 100 children (according to the identification numbers assigned to them) enrolled in the current study. Fifteen items that were found to be too easy (more than 90% correct responses) were taken out of the list. Slightly more difficult items (N = 19) were then tested on another 16 non-study children. The easy items that were discarded from the original list were replaced with sixteen of these difficult items. In total, 191 children participated in this evaluation. The final version of the KNT had 61 items ordered according to their difficulty level. The names of the objects range in difficulty from simple, high frequency words (easy) to rare words with low frequency (difficult) of occurrence.

Validation Study. The 61 items of the KNT were administered as part of a neuropsychological battery to 208 children from the main sample. (These children were not included in the process of developing the KNT as described in Step 4 of the adaptation procedures). The full battery (See Appendix for a brief description of the tests) comprised tests of executive function, verbal and working memory, verbal/visual selective reminding, learning, auditory and visual sustained and selective attention, and non-verbal reasoning. The battery did not include any other measures of expressive language. The tests were modified

for cultural and linguistic appropriateness (See Kitsao-Wekulo et al., 2013, for a detailed description of test modifications made and test administration procedures followed) and the battery was on average administered in a single session of approximately 2-hour duration, including two 10-minute breaks. In order to maximize participation, testing was conducted using an interactive play-like style which has been found suitable for children having little experience with standardised testing or where a significant proportion does not attend school (Alcock et al., 2008; Holding et al., 2004; Kitsao-Wekulo et al., 2013; Nampijja et al., 2010).

Administration of the KNT. In the KNT, the child is asked to spontaneously give one-word responses when presented with a black and white line drawing of a familiar object. The assessor pointed to a picture or part of a picture which the child was required to name. Testing was conducted within a room or in a quiet outdoor location at a school near the child's home. All children were tested individually but within sight of other children to minimize test anxiety. The items were administered to children in a standard order beginning with item 1. A stimulus cue was provided when no response was given, the child stated that s/he did not know the name or the item was incorrectly perceived. No time limits were imposed for responding. Similarly to the cut-offs used in the original procedure for the administration of the BNT (Kaplan et al., 1983), if a child failed to correctly name any objects on six consecutive trials, the test was discontinued. Several children (n = 167; 80.3%) met the criteria for discontinuation. The test took between 10 and 20 minutes to administer. A second administration of the KNT was completed within 6-8 weeks of the first assessment.

Scoring. If a child provided the correct response, i.e. the name of the item as indicated on the record sheet, the assessor recorded 'C' on the record sheet. An erroneous response that was spontaneously corrected before any cue was provided was also scored as

correct. In addition, a score was awarded when the child responded correctly after the cue was provided. Credit was given for a correct answer in any one of three languages – English, Kiswahili and the local dialect. Making provisions for responses in different languages is especially important in settings where children grow up using several languages (Alcock et al., 2008) as they are likely to pick up vocabulary terms in more than one language. A score of '1' was awarded for all correct responses. If provision of a stimulus cue did not result in a correct answer, i.e. the child's response differed from the names specified on the record sheet, the word that was produced by the child was recorded verbatim as a non-target word response. All scoring was checked by the assessor who administered the test and then cross-checked by a second assessor. Any disagreements were resolved through discussions. The final score was calculated by summing the number of spontaneously correct items with the number of correct items following a stimulus cue. The maximum score was 61.

Other Variables. Information on child gender, age, school experience (number of years that child has attended school), and household wealth was collected using a structured interview form. Birth records were used, where available, to confirm the child's date of birth. For the purpose of the current study, an age variable in 6-month increments was created. School experience was classified according to three categories – no school (non-schoolers), between one and two years of school (recent attenders), and more than 2 years of school (longer-term attenders). A composite index of household wealth that divided the sample into three approximately equal groups – least wealthy (Level 1), moderately wealthy (Level 2), and the most wealthy (Level 3) – was derived from six socioeconomic indicators: maternal and paternal education, maternal, and paternal occupation, type of windows in the child's dwelling and ownership of small livestock. The wealth index score was calculated by summing the values assigned to each of these indicators as detailed by Kitsao-Wekulo and

colleagues (2013) in an earlier study. Using a stadiometer, we measured children's heights to the nearest centimetre to calculate their nutritional status designated as with or without linear growth retardation. Linear growth retardation, a marker for nutritional status, was defined as height that was more than 2 standard deviations below levels predicted for age according to the World Health Organization (WHO) reference curves for school-aged children (World Health Organization, 2007). Area of residence was characterized as rural or urban according to the most common settlement within the school catchment area.

To quantify reading skills, we administered reading (letters, words and sentences) tests (Bhargava, Jukes, Ngorosho, Khilma, & Bundy, 2005) to a subset of 135 children in our study sample. In the reading task, children were required to select real letters, words and sentences from lists which included fake forms. This test was designed to measure children's comprehension and the scores were adjusted for incorrect answers. We summed the scores across reading tests to obtain the reading score.

A modified version of the Coloured Progressive Matrices (CPM: Raven, Court, & Raven, 1998) test was administered as a measure of non-verbal reasoning. In the CPM, the child is required to complete a pattern by selecting the appropriate missing symbol from a set of six alternatives. A detailed description of this test is provided in the appendix.

Analysis

A descriptive analysis of the background characteristics and distribution of scores was conducted. Item difficulty, defined as the percentage of correct responses for each item, was assessed to determine whether the items included on the test had appropriate difficulty levels (easy, medium, hard). Internal consistency and test-retest reliability of the KNT were quantified using the intraclass correlation coefficient (ICC). The Pearson product moment correlation coefficients (Pearson *r*) were computed to examine the relationship between the

KNT and non-verbal reasoning (CPM), and the KNT and reading skills. These associations were used as measures of convergent and concurrent validity, respectively.

As there were significant effects of maturational changes on performance (Figure 3), the KNT scores were regressed against age to produce age-corrected scores. Using the standardized KNT scores as the dependent variable, hierarchical linear regression analyses were conducted to understand which among the independent (background) variables (age, gender, nutritional status, household wealth, school experience and area of residence) explained the observed variance in KNT scores, and to explore the forms of these relationships. In the first step of the hierarchical regression, we inserted school experience and area of residence in a stepwise fashion as the main predictors. In the second step, we inserted household-level variables (household wealth and nutritional status) as these have been shown to have an influence on vocabulary scores. The third step involved entering the child-level characteristic (gender) as a predictor. Our justification for entering the variables in this order was based on the strength of the associations between the variables and KNT scores. The proportion of variance in naming performance accounted for by each of the background variables was quantified using R^2. Alpha levels were set at <.05 for statistical significance.

Results

Descriptives and item difficulty

Children who completed the final version of the KNT had a mean score of 20.74 (*SD* = 8.37; N = 208). The raw scores were normally distributed. The total number of correct responses was counted for each of the items (Table 2). Examples of some of the pictures that the children were required to name are provided in Figure 2. Overall, 96.7% (n = 59) of test items were named correctly by at least one child. The KNT had six extremely easy items

(which were named accurately by 90% or more of the children), seven very easy items (between 60% and 89% answered correctly), and 10 moderately easy items (named accurately by 40% to 59% of all children). Ten items were moderately difficult (named accurately by 20% to 39%), 21 items were very difficult (named accurately by between 2% and 19%) while two items were extremely difficult (none of the children responded correctly on these items).

Reliability and validity

The KNT had an internal consistency coefficient of .905 and a test-retest reliability level of .957. A moderate correlation, $r = .43$ ($p < .001$) was recorded between language (KNT) and non-verbal reasoning (CPM) scores. There was a weak correlation between the KNT and the reading score, $r = .18$ ($p = .022$).

Hierarchical regression analysis

The results of the hierarchical regression analysis produced four models. All the models were statistically significant: Model 1, $F = 39.450$, $p < .001$; Model 2, $F = 41.797$, $p < .001$; Model 3, $F = 21.952$, $p < .001$; and, Model 4, $F = 18.356$, $p < .001$. The regression analysis in Step 1 showed that school experience, $\beta = .443$, $t(304) = 8.281$, $p < .001$, contributed significantly to vocabulary development, accounting for 11.5% of the variance observed in KNT scores. The more years of schooling children obtained, the better their performance on the KNT. Area of residence was also a significant predictor which accounted for an additional 10.1% of the variance, $\beta = -.335$, $t(304) = -6.261$, $p < .001$. Children living in the rural areas of the study context had higher KNT scores than those living in urban areas. The results from Step 2 showed that nutritional status had a marginally significant effect on KNT scores, $\beta = .102$, $t(303) = 1.931$, $p = .054$. Even though children without linear growth retardation had slightly higher scores than those with linear growth retardation, the effect of

nutritional status was small and explained only 1% of the variance. The addition of gender in Step 3 resulted in the final model which showed that boys performed slightly better than girls. However, gender did not have a significant effect on KNT scores, and explained only an additional 0.8% of the variance. With all variables included in the model in Step 3, school experience, $\beta = .410$, $t(302) = 7.001$, $p < .001$ and area of residence, $\beta = -.338$, $t(302) = -6.362$, $p < .001$, remained as significant predictors and together accounted for 23.4% of the variance observed in KNT scores (Table 3).

Discussion

The primary aim of the current study was to develop a psychometrically-sound measure of expressive vocabulary for use in a resource-limited setting. In order to assess expressive vocabulary, we chose to use confrontation naming. First, the measures used to assess confrontation naming tap cognitive skills such as encoding and retrieval (Cheung et al., 2004; Halperin, Healey, Zeitchik, Ludman, & Weinstein, 1989; Jordan & Ashton, 1996). Secondly, expressive vocabulary tests show strong relationships with other aspects of oral language and therefore more accurately reflect emergent literacy (Malvern & Richards, 1997). For instance, reading vocabulary was found to be highly correlated with BNT performance (Hawkins, Sledge, Orleans, Quinlan, & Huffman, 1993), while other studies have reported that expressive vocabulary (as measured by the BNT) accounted for significant variance in both exception word reading and reading comprehension among both children (Ricketts, Nation, & Bishop, 2007) and adults (Hall, Greenberg, Gore, & Pae, 2014). Expressive vocabulary measures can thus serve as proxies for reading comprehension specifically, and academic achievement more generally (R. C. Anderson & Freebody, 1981; Beck, McKeown, & Kucan, 2002; Cunningham & Stanovich, 1998; Dickinson & Tabors, 2001; Sénéchal, Ouellette, & Rodney, 2006). Thirdly, whereas receptive vocabulary tests do

not require reading, writing or speaking during assessment, they are more costly and complex to produce and require more time to administer than expressive vocabulary tests. Also, the requirement to choose from a selection of available items bears little relation to the way language is used in most real-life situations (Luo & Zhang, 2011). This may make the test format more susceptible to guessing and impulsive responding than tests requiring an open-answer format (Luo & Zhang, 2011). Fourth, we considered confrontation naming a suitable approach because compared to younger children, most school-age children possess naming abilities and are able to verbalise their responses. Furthermore, at this age children have appropriate levels of comprehension and concentration making such a method more sensitive (Clacherty & Kushlik, 2004). This procedure therefore provides a more direct assessment of vocabulary skills than would be obtained using parental reports or observation of communicative interactions. We first formulated an initial set of items based on previous research which were then refined through administrations to small, diverse groups of children. The KNT did not require reading and writing; this feature made it suitable for both schooling and non-schooling children. The test was short, relatively inexpensive and easy to administer without the need for specialised training and equipment.

The KNT scores showed a normal distribution demonstrating sensitivity to within-population variance, and had an adequate difficulty level. The KNT also had excellent internal consistency and test-retest reliability levels (George & Mallery, 2003). The positive and moderate correlation between vocabulary scores (KNT) and non-verbal reasoning (CPM) demonstrated evidence of convergent validity in accordance with earlier reports (Court & Raven, 1995; Luoni et al., 2015; Storms et al., 2004). As evidence of concurrent validity, we expected that vocabulary scores would correlate strongly with reading comprehension as established in earlier reports – however, our tool only demonstrated weak correlations with

reading skills. One possible explanation for the weak correlation with reading skills is the methodological differences in measurement of reading outcomes between the current study and earlier studies. Secondly, the KNT has not previously been normed on a rural school-age population and the findings may indicate that the KNT is not an accurate measure of their expressive vocabulary. A similar limitation was put forward by Hall and colleagues (2014) in their study among struggling adult readers. In addition, there may be some unknown or poorly-understood differences in expressive naming in the children sampled in the current study compared to other samples. These differences may have arisen from individual factors such as age of acquisition, or word attributes such as depictability and word frequency (Massaro & Perlman, 2017), which may influence children's ability to correctly name a target item.

Beyond simply reporting the development of a measure of expressive vocabulary for school-age children, the current study's major contribution is the examination of the sources of variability in naming performance. For one, age-related differences in vocabulary acquisition were similar to what has been reported within similar (Alcock et al., 2008) and different contexts (Storms et al., 2004; Vogt et al., 2015). These age-related differences in vocabulary scores suggest that naming ability improves with maturation as children acquire more vocabulary with increased exposure.

Second, the findings of the current study demonstrated the important role of school exposure in a predominantly non-literate setting. The positive effects of school exposure on vocabulary acquisition have also been reported in other studies within similar contexts (Alcock et al., 2008; Carter et al., 2005; Sigman, McDonald, Neumann, & Bwibo, 1991). The experience of schooling may foster certain cognitive styles (Sigman, Neumann, Carter, et al., 1989; Sinha & Misra, 1982) and promote knowledge of common and uncommon objects

(Strauss, Sherman, & Spreen, 2006). The testing format may also closely mimic the requirement of following instructions that children who go to school are exposed to, thus making them more confident in responding to test items.

Third, our neighbourhood variable, represented by the locality of the school (rural vs urban), also had a significant impact children's vocabulary scores. Although rural areas are likely to be characterised by limited access to resources and services (Greenfield, 2009), which in turn has an indirect negative impact on children's language scores, in the current study, children in rural areas of the district performed better than their counterparts in urban areas. Our findings are in contrast with other studies that have reported overall better performance among children living in urban areas (Kornilova et al., 2017; Schady et al., 2015; Vogt et al., 2015). These findings may be attributed to the likelihood that the drawings on the KNT represented objects that were more familiar to children in rural areas than those living in urban areas.

Fourth, stunting had a marginally significant impact on vocabulary scores – poorer nutritional status was associated with lower language scores. The results from the studies conducted by Wachs (1995) and Grantham-McGregor and others (2007) have shown a strong relation between malnutrition and human development; chronic malnutrition is associated with a variety of cognitive deficits in school-age children. Risk factors related to poverty frequently co-vary and previous studies of poorly-resourced contexts have also reported a strong association between children's nutritional status and socioeconomic conditions (Grantham-McGregor et al., 2007; Kanjilal, Mazumdar, Mukherjee, & Rahman, 2010; Sigman, Neumann, Carter, et al., 1989). A major problem of the current and previous studies is the difficulty of controlling for potential confounders in order to determine the unique

contribution of co-existing adverse environmental factors that have a negative effect on child outcome.

Fifth, although the effects of gender were evident in the current study, the existing body of literature that has examined similar outcomes reports contradictory patterns of performance. The finding that boys in the current study performed slightly better than girls is congruent with other reports of school-age populations in non-similar cultural settings in Spain, Belgium and Egypt where males scored significantly higher than females on tests of verbal abilities (Ardila & Rosselli, 1994; Storms et al., 2004; Wachs et al., 1995). In contrast, other studies have reported superior performance among girls (Hamadani et al., 2010) or the lack of a gender effect on naming performance (Kim & Na, 2008; Luoni et al., 2015). We speculate that boys in the current study performed better than girls because of the influences within the socio-cultural environment. For instance, boys are often encouraged to play with toys which encourage spatial manipulations (Levine, Huttenlocher, Taylor, & Langrock, 1999), and this could have given them an advantage in interpretation of the line drawings. As a result, they may have found it easier to name the items on the KNT.

There was no association between household resources, our representation of SES, and vocabulary scores. Earlier reports have also established lack of significant associations between SES and other cognitive outcomes (Kitsao-Wekulo et al., 2013). However, our finding contrasts other studies which report superior performance among children from households with more resources at their disposal compared to their peers from poorly-resourced households in both non-Western and Western contexts (Hart & Risley, 1995; Hoff, 2003; Noble, McCandliss, & Farah, 2007; Sigman et al., 1991; Sigman, Neumann, Carter, et al., 1989; Whaley, Sigman, Beckwith, Cohen, & Espinosa, 2002). The relatively homogeneous distribution of resources among households within the current study setting or

the manner in which information on household resources was collected may possibly explain our null findings.

Our study demonstrated that while school experience and area of residence accounted for the differences observed in KNT performance in the final analytic model, school experience explained the biggest proportion of this variation. In other studies within similar settings, duration of schooling, physical stature and SES accounted for much of the variability seen (Sigman et al., 1991; Sigman, Neumann, Jansen, & Bwibo, 1989). The overlap between the current study and previous studies seen only in terms of schooling exposure points to region-specific influences of contextual factors.

The current study did not collect data on participants' expressive semantic abilities using another measure, precluding the possibility of adequately validating the KNT against a 'gold standard.' This limits our understanding of the psychometric properties of the tool. Another limitation that arose was that the moderate association of the KNT with non-verbal reasoning, and the weak correlation with reading scores may not adequately inform the development of a language measure for this population. Moreover, the study context poses several challenges to the assessment of vocabulary acquisition in children occasioned by limited skills and resources, limited African empirical literature and the lack of validation studies from elsewhere. Standardised tests are designed to be administered by specially trained professionals and much expertise is required for scoring and interpretation. We put in place a rigorous training programme to ensure that standardised procedures were followed during administration and scoring of the test. And similarly to Barker-Collo's (2001) conclusions in her study of New Zealanders, we can surmise that while the process of development of the KNT resulted in a test that was more culturally valid within a rural

African context, it is unlikely that any single test version will be culturally appropriate to the diverse linguistic groupings within the larger society.

Conclusion

In interpreting our test results, we took various contextual factors which appear to be important influences on performance, into account. However, while an earlier study (Kitsao-Wekulo et al., 2013) demonstrated that the patterns of influence and strength of these relationships may differ even within similar contexts, some of the relationships were common across several outcomes. Others relationships were specific to expressive vocabulary. Furthermore, improving measurement of contextual variables such as household resources is vital to the accurate interpretation of test scores and may elucidate further the contribution of other salient factors.

We suggest that our study findings provide preliminary evidence for the range of scores that we should expect from typically developing school-age children in a predominantly rural setting. These findings are important for the development of normative tables which will be a significant contribution for researchers and professionals in the child development field. However, because a 'gold standard' of expressive language does not exist for school-age populations in sub-Saharan Africa, of value in future studies would be to concurrently administer a second language measure together with the KNT. As has been suggested by Hoffman and colleagues in their investigation of the psychometric properties of two oral language measures, (2011), the assessment of language abilities of school-age children requires systematic collection of data from a variety of sources. Alternative measures of expressive vocabulary using oral language samples include the Test of Language Development-Primary, 3rd Edition (TOLD-P:3; Newcomer & Hammill, 1997) which would be suitable for this process as it has Picture Vocabulary and Oral Vocabulary subtests.

Another possibility, which is considered a more naturalistic and culture-fair assessment, is measuring the Number of Different Words (NDW) produced in a narrative, to determine children's lexical diversity. Such testing could occur in the context of structured tasks (Mills, 2015).

We also suggest the need for further studies to investigate the predictive and clinical validity of the KNT to elucidate cause-effect relationships, and to provide information on the tool's sensitivity and specificity, as has been recommended by Dollagan (2007), before the tool can be used in large-scale screening. This will facilitate the accurate interpretation of assessment results.

Figure 1. Steps in creation of KNT item pool

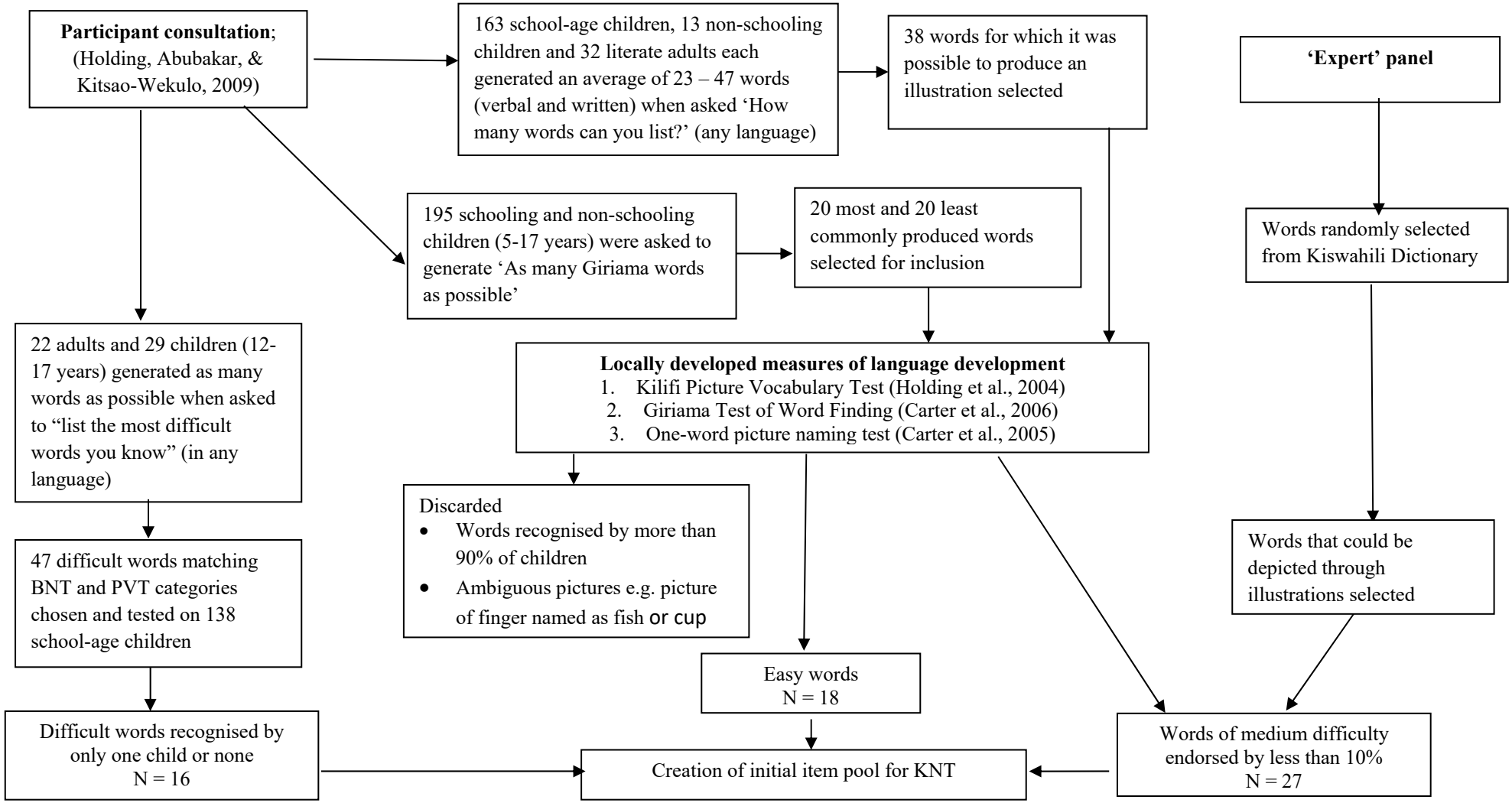

Figure 2: Examples of pictures used on the KNT

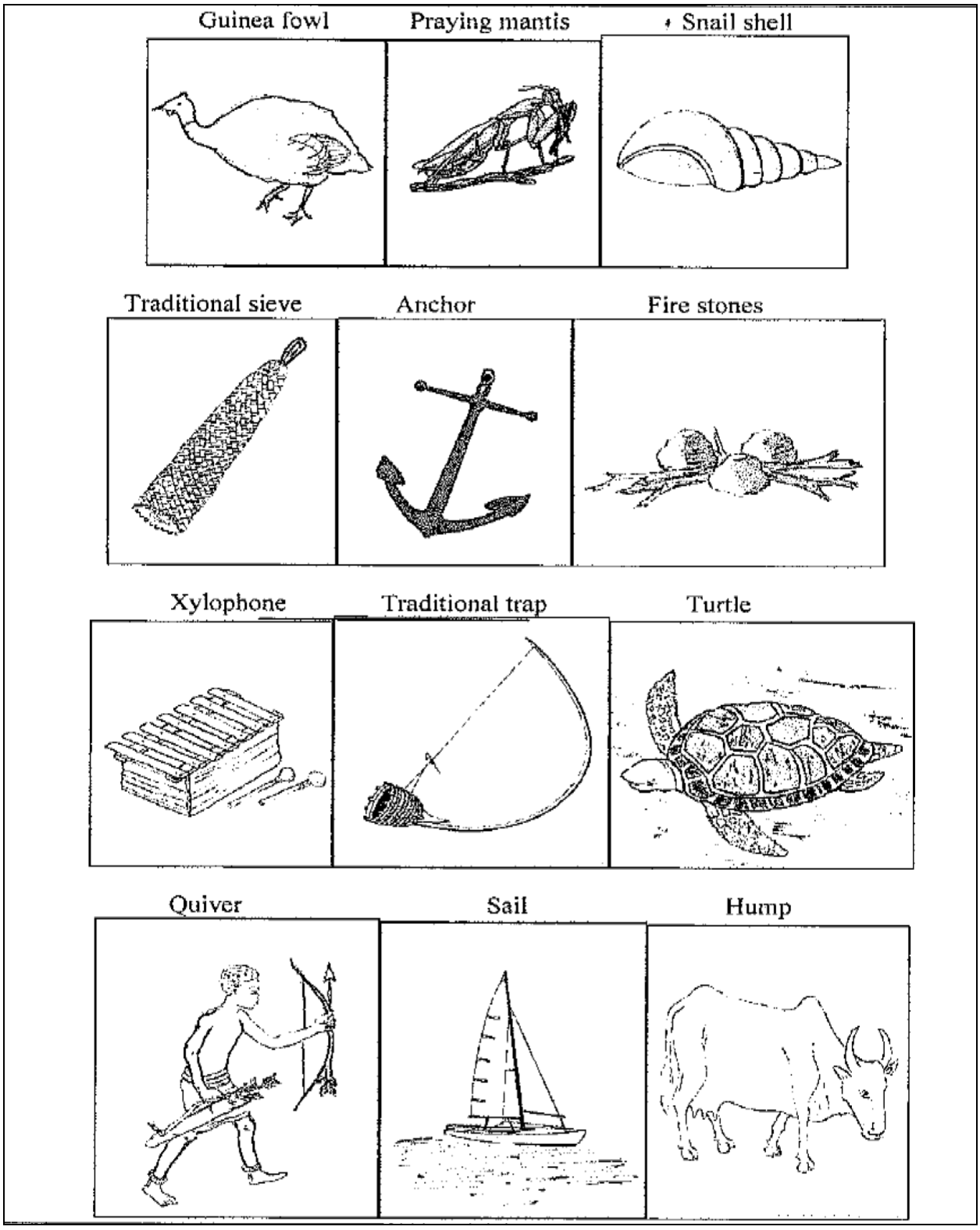

Figure 3. Age effects on KNT performance

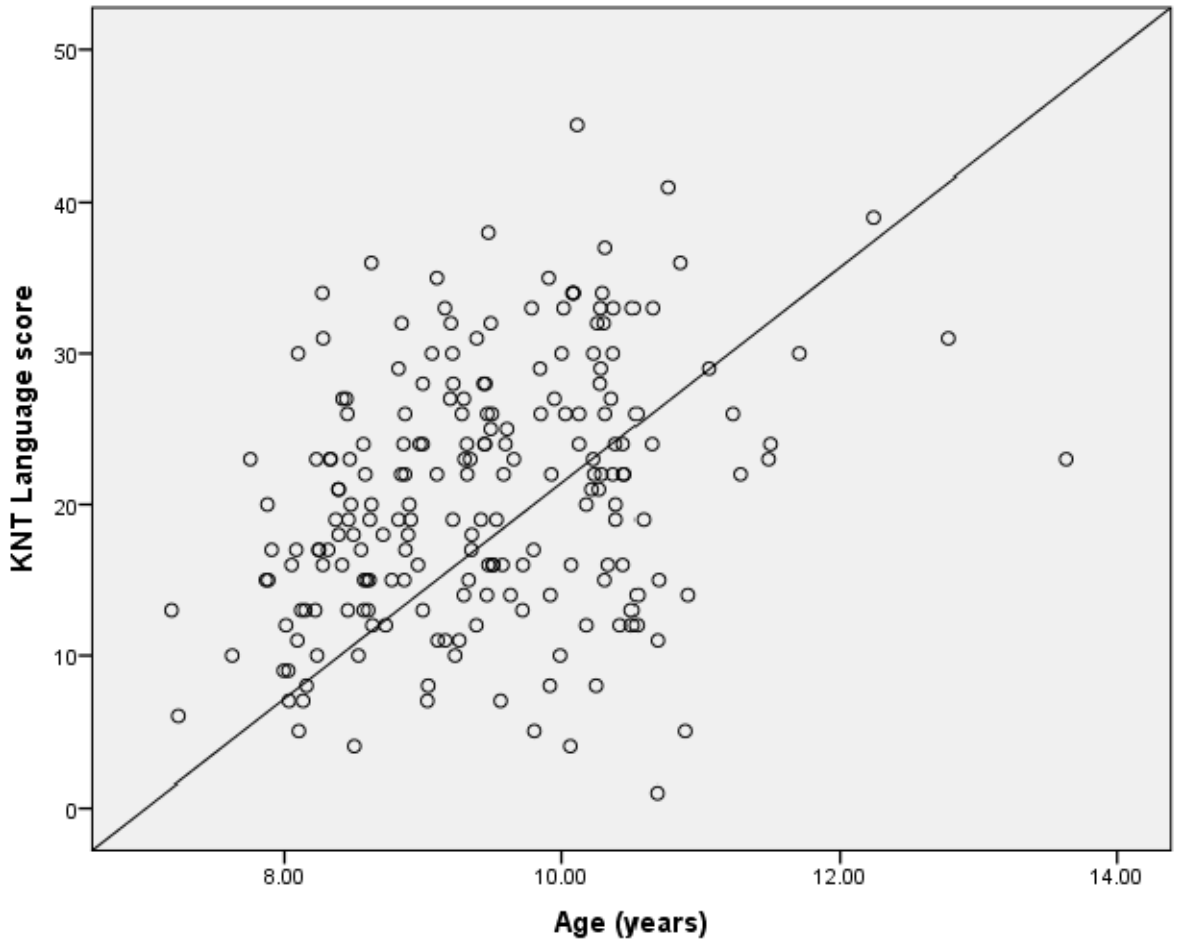

Table 1. Description of tool development and validation study samples

Variables	Tool development N (%)	Validation N (%)	Total N (%)
N	100	208	308
Gender			
Female	52 (52.0)	108 (51.9)	160 (51.9)
Male	48 (48.0)	100 (48.1)	148 (48.1)
Age (years)			
≤ 8.4	27 (27.0)	39 (18.8)	66 (21.4)
8.5 - 9.4	32 (32.0)	72 (34.6)	104 (33.8)
≥ 9.5	41 (41.0)	97 (46.6)	138 (44.8)
Linear growth retardation			
Present	25 (25.0)	49 (23.6)	74 (24.0)
Absent	75 (75.0)	159 (76.4)	234 (76.0)
School experience (years)			
None	6 (6.0)	29 (13.9)	35 (11.4)
1-2 years (recent attenders)	40 (40.0)	61 (29.3)	101 (32.8)
> 2 years (longer-term attenders)	54 (54.0)	118 (56.7)	172 (55.8)
Household status			
Level 1 (Least wealthy)	45 (45.0)	68 (32.7)	123 (39.9)
Level 2 (Moderately wealthy)	30 (30.0)	74 (35.6)	94 (30.5)
Level 3 (Most wealthy)	25 (25.0)	66 (31.7)	91 (29.5)
		-	
Residence			
Rural	100 (100)	145 (69.7)	245 (79.5)
Urban	0	63 (30.3)	63 (20.5)

Table 2. Total number and proportion correct of selected items, N = 208

Item number	Target word	Total correct (n)	Proportion correct (%)
47	Horn	2	1.0
59	Sail	3	1.4
31	Adze	3	1.4
58	Anchor	4	1.9
55	Torch	4	1.9
34	Udder	5	2.4
56	Mat coil	6	2.9
28	Pipe	6	2.9
5	Adam's apple	10	4.8
48	Xylophone	13	6.3
32	Traditional pot holder (*kata*)	16	7.7
26	Traditional sieve (*kifumbu*)	19	9.1
37	Pilot	20	9.6
50	Turtle	26	12.5
52	Gutter	32	15.4
43	Ring	38	18.3
53	Jingles	43	20.7
45	Drummer	43	20.7
39	Lobster	44	21.2
40	Owl	49	23.6
38	Turkey	55	26.4
49	Fish scales	56	26.9
44	Guinea fowl	57	27.4
41	Earrings	62	29.8
60	Praying mantis	68	32.7
35	Hooves	88	42.3
51	Snail shell	88	42.3
4	Eyebrows	91	43.8
36	Hump	93	44.7
46	Mat	105	50.5
10	Traditional skirt (*hando*)	105	50.5
42	Horse	109	52.4
3	Elbow	110	52.9
54	Charm	130	62.5
27	Desk	146	70.2
30	Trap	153	73.6
29	Leaf	173	83.2
17	Traditional ladle (k*ipawa*)	175	84.1
24	Maize	175	84.1
33	Tail	195	93.8
13	Door	195	93.8
20	Cat	200	96.2
18	Cup	202	97.1
8	Tap	204	98.1
14	Ball	207	99.5

Table 3. Significant predictors of KNT scores

	b	SE *b*	*β*	95% CI for *b*	
				Lower	Upper
Step 1 (R^2 = .216)					
Constant	-.542	.096			
School experience	.262	.032	.443***	.200	.324
Area of residence	-.824	.132	-.335***	-1.083	-.565
Step 2 (ΔR^2 = .010)					
Constant	-.367	.153			
School experience (years)	.247	.035	.417***	.178	.315
Area of residence	-.837	.131	-.340***	-1.095	-.578
Household wealth	.000	.014	-.001	-.028	.027
Nutritional status	.098	.051	.102^{+}	-.002	.198
Step 3 (ΔR^2 = .008)					
Constant	-.251	.172			
School experience (years)	.243	.035	.410***	.175	.311
Area of residence	-.883	.131	-.338***	-1.091	-.575
Household wealth	-.001	.014	-.005	-.028	.026
Nutritional status	-.100	.051	.104^{+}	.000	.199
Gender	-.183	.101	-.092^{+}	-.382	.015

$^{+}p < 0.1$, $*p < 0.05$; $**p < 0.01$, $p < 0.001$

References

Alcock, K. J., Abubakar, A., Newton, C. R. J. C., & Holding, P. (2016). The effects of prenatal HIV exposure on language functioning in Kenyan children: establishing an evaluative framework. *BMC Research Notes, 9*, 463.

Alcock, K. J., Holding, P. A., Mung'ala-Odera, V., & Newton, C. R. J. C. (2008). Constructing tests of cognitive abilities for schooled and unschooled children. *Journal of Cross-Cultural Psychology, 39*(5), 529-551.

Alcock, K. J., Rimba, K., Holding, P., Kitsao-Wekulo, P. K., Abubakar, A., & Newton, C. R. J. C. (2015). Developmental inventories using illiterate parents as informations: Communicative Developmental Inventory (CDI) adaptation for two Kenyan languages. *Journal of Child Language, 42*(4), 763-785. doi: 10.1017/S0305000914000403

Anderson, P., Anderson, V., & Lajoie, G. (1996). The Tower of London Test: Validation and Standardization for Pediatric Populations. *The Clinical Neuropsychologist, 10*(1), 54-65.

Anderson, P., Anderson, V., Northam, E., & Taylor, H. G. (2000). Standardization of the Contingency Naming Test (CNT) for School-Aged Children: A Measure of Reactive Flexibility. *Clinical Neuropsychological Assessment, 1*, 247-273.

Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 77-116). Newark, DE: International Reading Association.

Apiwattanalunggarn, K. L., & Luster, T. (2005). Individual differences in the school performance of 2nd-grade children born to low-income adolescent mothers: Findings from an 8-Year longitudinal study. *Journal of Research in Childhood Education, 19*(4), 314-332.

Ardila, A., & Rosselli, M. (1994). Development of language, memory, and visuospatial abilities in 5- to 12-year-old children using a neuropsychological battery. *Developmental Neuropsychology, 10*(2), 97-120.

Awiti, J. A. O. (2011). *Communication rules of the Maasai and the Akamba: A comparative analysis.* (Doctor of Literature and Philosophy PhD), University of South Africa.

Barker-Collo, S. L. (2001). The 60-item Boston Naming Test: Cultural bias and possible adaptations for New Zealand. *Aphasiology, 15*(1), 85-92.

Basilio, C. S., Puccini, R. F., Silva, E. M. K. d., & Pedromónico, M. r. M. (2005). Living conditions and receptive vocabulary of children aged two to five years. *Rev Saude Publica, 39*(5), 725-730.

Bates, E., Dale, P. S., & Thal, D. (1995). Individual differences and their implications for theories of language development. In P. Fletcher & B. MacWhinney (Eds.), *Handbook of Child Language*. Oxford: Basil Blackwell.

Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing words to life: Robust vocabulary instruction*. New York: Guilford.

Bhargava, A., Jukes, M., Ngorosho, D., Khilma, C., & Bundy, D. A. P. (2005). Modeling the effects of health status and the educational infrastructure on the cognitive development of Tanzanian schoolchildren. *American Journal of Human Biology, 17*(3), 280-292.

Bornstein, M. H., Hahn, C.-S., & Haynes, M. O. (2004). Specific and general language performance across early childhood: Stability and gender considerations. *First Language, 24*, 267-304.

Bortz, M. (1995). *South Africa Language Assessments Manual*.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments in nature and design* Cambridge, MA: Harvard University Press.

Bronfenbrenner, U. (1995). Developmental ecology through space and time: A future perspective. In P. Moen, J. G H Elder, & K. Luscher (Eds.), *Examining lives in context: Perspectives on the ecology of human development* (pp. 619-647). Washington DC: American Psychological Association.

Brown, J. L., & Pollitt, E. (1996). Malnutrition, poverty and intellectual development. *Scientific American, 274,* 38-43.

Burman, D. D., Bitan, T., & Booth, J. R. (2008). Sex differences in neural processing of language among children. *Neuropsychologia, 46*(5), 1349-1362.

Cantalini, M. (1987). *The effects of age and gender on school readiness and school success.* (PhD), Ontario Institute for Studies in Education, Toronto, Canada.

Carter, J. A., Lees, J. A., Gona, J. K., Murira, G., Rimba, K., Neville, B. G. R., & Newton, C. R. J. C. (2006). Severe falciparum malaria and acquired childhood language disorder. *Developmental Medicine, 48*, 51-57.

Carter, J. A., Ross, A. J., Neville, B. G. R., Obiero, E., Katana, K., Mung'ala-Odera, V., . . . Newton, C. R. J. C. (2005). Developmental impairments following severe falciparum malaria in children. *Tropical Medicine and International Health, 10*(1), 3-10.

Chall, J. S., Jacobs, V. A., & Baldwin, L. E. (1990). *The reading crisis: Why poor children fall behind*. Cambridge, MA: Harvard University Press.

Cheung, R. W., Cheung, M.-C., & Chan, A. S. (2004). Confrontation naming in Chinese patients with left, right or bilateral brain damage. *Journal of the International Neuropsychological Society, 10*, 46-53.

Childers, J. B., Vaughan, J., & Burquest, D. A. (2007). Joint attention and word learning in Ngas-speaking toddlers in Nigeria. *Journal of Child Language, 34*(2), 199-225.

Christian, K., Morrison, F. J., Frazier, J. A., & Massetti, G. (2000). Specificity in the nature and timing of cognitive growth in kindergarten and first grade. *Journal of Cognition and Development, 1*(4), 429-448.

Clacherty, G., & Kushlik, A. (2004). *Meeting the challenge of research with very young children: a practical outline of methodologies used in the formative research and pre-testing of the Takalani Sesame HIV and AIDS television and radio programmes.* Paper presented at the Fourth International Education Entertainment Conference, Cape Town.

Commission on Revenue Allocation. (2013). *Kenya County Fact Sheets*. Nairobi: Commission on Revenue Allocation.

Connolly, K. J., & Grantham-McGregor, S. M. (1993). Key issues in generating a psychological-testing protocol. *American Journal of Clinical Nutrition, 57*(suppl), 317S- 318S.

Connolly, K. J., & Pharoah, P. O. D. (1993). Subclinical effects of iodine deficiency: Problems of assessment. In J. B. Stanbury (Ed.), *The damaged brain of iodine deficiency: Cognitive, behavioral, neuomotor, educative aspects* (pp. 27-36). New York, NY: Cognizant Communications.

Connor, C. M., Morrison, F. J., & Slominski, L. (2006). Preschool instructino and children's emergent literacy skill growth. *Journal of Educational Psychology, 98*, 665-689. doi: 10.1037/0022-0663.98.4.665

Court, J. H., & Raven, J. (1995). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 7: Research and references: Summaries of normative, reliability, and validity studies and references to all sections*. Oxford, England/San Antonio, TX: Oxford Psychologists Press/The Psychological Corporation.

Cunningham, A. E., & Stanovich, K. E. (1998). What reading does for the mind. *American Educator*.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1989). *California Verbal Learning Test - Children's Version*. San Antonio: Harcourt Brace & Company.

Demuth, K. (1990). Subject, topic and Sesotho passive. *Journal of Child Language, 17*, 67-84.

Dickinson, D. K., & Tabors, P. O. (Eds.). (2001). *Beginning literacy with language: young children learning at home and school*. Baltimore: Paul H. Brookes.

Dollagan, C. A. (2007). *The Handbook for Evidence-based Practice in Communication Disorders*. Baltimore, MD: Paul H. Brookes.

Dupéré, V., Leventhal, T., Crosnoe, R., & Dion, É. (2010). Understanding the positive role of neighborhood socioeconomic advantage in achievement: The contribution of the home, child care and school environments. *Developmental Psychology, 46*(5), 1227-1244. doi: 10.1037/a0020211

Eriksson, M., Marschik, P. B., Tulviste, T., Almgren, M., Pereira, M. P., Wehberg, S., . . . Gallego, C. (2012). Differences between girls and boys in emerging language skills: Evidence from 10 language communities. *British Journal of Developmental Psychology, 30*, 326-343. doi: 10.1111/j.2044-835X.2011.02042.x

Evans, G. W., Maxwell, L. E., & Hart, B. (1999). Parental language and verbal responsiveness to children in crowded homes. *Developmental Psychology, 35*, 1020-1023.

Fernald, L. C. H., Weber, A., Galasso, E., & Ratsifandrihamanana, L. (2011). Socioeconomic gradients and child development in a very low income population: evidence from Madagascar. *Developmental Science, 14*(4), 832-847.

Fletcher, J. M. (1985). Memory for verbal and nonverbal stimuli in learning disability subgroups: Analysis by selective reminding. *Journal of Experimental Child Psychology, 40*, 244-259.

George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update* (4th ed.). Boston: Allyn & Bacon.

Grantham-McGregor, S., Cheung, Y. B., Cueto, S., Glewwe, P., Richter, L., Strupp, B., & Group, T. I. C. D. S. (2007). Child development in developing countries 1: Developmental potential in the first 5 years for children in developing countries. *The Lancet, 369*, 60-70.

Greenfield, P. M. (2009). Linking social change and developmental change: shifting pathways of human development. *Developmental Psychology, 45*(2), 401-418.

Hall, R., Greenberg, D., Gore, J. L., & Pae, H. K. (2014). The Relationship Between Expressive Vocabulary Knowledge and Reading Skills for Adult Struggling Readers. *Journal of Research in Reading, 37*(Suppl 1), 87-100. doi: 10.1111/j.1467-9817.2012.01537.x

Halperin, J. M., Healey, J. M., Zeitchik, E., Ludman, W. L., & Weinstein, L. (1989). Developmental aspects of linguistic and mnestic abilities in normal children. *Journal of Clinical and Experimental Neuropsychology, 11*(4), 518-528. doi: 10.1080/01688638908400910

Hamadani, J. D., Baker-Henningham, H., Tofail, F., Mehrin, F., Huda, S. N., & Grantham-McGregor, S. M. (2010). Validity and reliability of mothers' reports of language development in 1-year-old children in a large-scale survey in Bangladesh. *Food and Nutrition Bulletin, 31*(2 (supplement)), S198 - S206.

Hart, B., & Risley, T. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore: Paul H. Brookes Publishing.

Hawkins, K. A., Sledge, W. H., Orleans, J. F., Quinlan, D. M., & Huffman, R. E. (1993). Normative implications of the relationship between reading vocabulary and Boston naming test performance. *Archives of Clinical Neuropsychology, 8*(6), 525-537.

Hayes, D. P., & Ahrens, M. (1988). Vocabulary simplification for children: A special case of 'motherese'. *Journal of Child Language, 15*, 395-410.

Hoff-Ginsberg, E. (1991). Mother-child conversation in different social classes and communicative settings. *Child Development, 62*(4), 782-796.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development, 74*(5), 1368-1378.

Hoff, E., Laursen, B., & Tardif, T. (2002). Socioeconomic status and parenting. In M. H. Bornstein (Ed.), *Handbook of parenting* (2nd ed., pp. 231-252). Mahwah, NJ: Erlbaum.

Hoffman, L. M., Loeb, D. F., Brandel, J., & Gillam, R. B. (2011). Concurrent and Construct Validity of Oral Language Measures with School Age Children with Specific Language Impairment. *Journal of Speech, Language and Hearing Research, 54*(6), 1597-1608.

Holding, P. A., Abubakar, A., & Kitsao-Wekulo, P. (2009). Where there are no tests: A systematic approach to test adaptation. In M. L. Landow (Ed.), *Cognitive impairment: Causes, diagnosis and treatments* (pp. 189-200). New York, NY: Nova Science.

Holding, P. A., Taylor, H. G., Kazungu, S. D., Mkala, T., Gona, J., Mwamuye, B., . . . Stevenson, J. (2004). Assessing cognitive outcomes in a rural African population: Development of a neuropsychological battery in Kilifi District, Kenya. *Journal of the International Neuropsychological Society, 10*, 246-260.

Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology, 45*, 337-374.

Hyde, J. S., & Linn, M. C. (1988). Gender differences in verbal ability: A meta-analysis. *Psychological Bulletin, 104*(1), 53-69.

Jordan, F. M., & Ashton, R. (1996). Language performance of severely closed head injured children. *Brain Injury, 10*(2), 91-97.

Kahuthu, R., Muchoki, T., & Nyaga, C. (2005). *Kilifi District Strategic Plan 2005-2010 for Implementation of the National Population Policy for Sustainable Development*. Nairobi: National Coordination Agency for Population and Development.

Kanjilal, B., Mazumdar, P. G., Mukherjee, M., & Rahman, M. H. (2010). Nutritional status of children in India: household socio-economic condition as the contextual determinant. *International Journal for Equity in Health, 9*, 19-31.

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *The Boston Naming Test*. Philadelphia: Lea and Febiger.

Kim, H., & Na, D. L. (2008). A normative study of the Boston Naming Test in 3- to 14-year-old Korean children. *The Clinical Neuropsychologist, 22*(1), 84-97. doi: 10.1080/13854040601064526

Kitsao-Wekulo, P. K., Holding, P. A., Taylor, H. G., Abubakar, A., & Connolly, K. (2013). Neuropsychological testing in a rural African school-age population: evaluating contributions to variability in test performance. *Assessment, 20*(6), 776-784. doi: 10.1177/1073191112457408

Kornilova, S. A., Lebedeva, T. V., Zhukova, M. A., Prikhoda, N. A., Korotaeva, I. V., Koposov, R. A., . . . Grigorenko, E. L. (2017). Language development in rural and urban Russian-speaking children with and without developmental language disorder. *Learning and Individual Differences, 46*, 45-53. doi: 10.1016/j.lindif.2015.07.001

Leaper, C. (2002). Parenting girls and boys. In M. H. Bornstein (Ed.), *Handbook of parenting: Vol. 1. Children and parenting* (2nd ed., Vol. 1, pp. 127–152). Mahwah, NJ: Erlbaum.

Leventhal, T., & Brooks-Gunn, J. (2000). The neighborhoods they live in: The effects of neighborhood residence upon child and adolescent outcomes. *Psychological Bulletin, 126*, 309-337.

Levine, S. C., Huttenlocher, J., Taylor, A., & Langrock, A. (1999). Early sex differences in spatial skills. *Developmental Psychology, 35*, 940-949. doi: 10.1037/0012-1649.35.4.940

Luo, S., & Zhang, X. (2011). Multiple-choice item and its backwash effect on language teaching in China. *Theory and Practice in Language Studies, 1*(4), 423-425.

Luoni, C., Balottin, U., Rosana, L., Savelli, E., Salini, S., & Termine, C. (2015). Confrontation naming and reading abilities at primary school: A longitudinal study. *Behavioural Neurology, 2015*, 10 pages. doi: 10.1155/2015/981548

Maccoby, E. (1980). *Social Development*. New York, NY: Harcourt Brace Jovanovich.

Malvern, D. D., & Richards, B. J. (1997). A new measure of lexical diversity. In A. Ryan & A. Wray (Eds.), *Evolving models of language*. Bristol, England: Multilingual Matters.

Manly, T., Robertson, I. H., Anderson, V., & Nimmo-Smith, I. (1999). *TEA-Ch: The Test of Everyday Attention for Children*. Bury St Edmunds England: Thames Valley Test Company Limited.

Massaro, D. W., & Perlman, M. (2017). Quantifying iconicity's contribution during language acquisition: Implications for vocabulary learning. *Frontiers in Communication, 2*(4). doi: 10.3389/fcomm.2017.00004

Mbise, A. S., & Kysela, G. M. (1990). Developing appropriate screening and assessment instruments: the case of Tanzania. In M. J. Thorburn & K. Marfo (Eds.), *Practical Approaches to Childhood Disability in Developing Countries: Insights from Experience and Research*. Tampa, FL: Global Age.

Miller, G. A., & Gildea, P. M. (1987). How children learn words *Scientific American* (pp. 94-99).

Mills, M. T. (2015). The Effects of Visual Stimuli on the Spoken Narrative Performance of School-Age African American Children. *Language, Speech, and Hearing Services in Schools, 46*, 337-351.

Miotto, E. C., Sato, J., Lucia, M. C. S., Camargo, C. H. P., & Scaff, M. (2010). Development of an adapted version of the Boston Naming Test for Portuguese speakers. *Revista Brasileira de Psiquiatria, 32*(3), 279-282.

Nampijja, M., Apule, B., Lule, S., Akurut, H., Muhangi, L., Elliot, A. M., & Alcock, K. J. (2010). Adaptation of Western measures of cognition for assessing 5-year-old semi-urban Ugandan children. *British Journal of Educational Psychology, 80*(1), 15-30.

Nation, K., & Snowling, M. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading, 27*(4), 342-356. doi: 10.1111/j.1467-9817.2004.00238.x

Newcomer, P. L., & Hammill, D. D. (1997). *Test of Language Development-Primary* (3rd Edition ed.). Austin, TX: Pro-Ed Inc.

Noble, K. G., McCandliss, B. D., & Farah, M. J. (2007). Socioeconomic gradients predict indvidual differences in neurocognitive abilities. *Developmental Science, 10*(4), 464-480.

Nyasani, J. M. (1997). *The African psyche*. Nairobi: University of Nairobi Theological Printing Press.

Pakendorf, C., & Alant, E. (1997). Culturally valid assessment tools: Northern Sotho translation of the Peabody Picture Vocabulary Test - Revised. *South African Journal of Communication Disorders, 44*, 3-12.

Petrides, M., & Milner, B. (1982). Deficits on subject-ordered tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologia, 20*(3), 249-262.

Raven, J. C., Court, J. H., & Raven, J. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scales: Section 2 - The Coloured Progressive Matrices*. San Antonio, Texas: Psychological Corporation.

Ricketts, J., Nation, K., & Bishop, D. V. M. (2007). Vocabulary is important for some, but not all reading skills. *Scientific Studies of Reading, 11*(3), 235-257.

Sampson, R. J., Morenoff, J. D., & Gannon-Rowley, T. (2002). Assessing 'Neighborhood Effects': Social Processes and New Directions in Research. *Annual Review of Sociology, 28*, 443-478.

Schady, N., Behrman, J., Araujo, M. C., Azuero, R., Bernal, R., Bravo, D., . . . Vakis, R. (2015). Wealth gradients in early childhood cognitive development in five Latin American countries. *Journal of Human Resources, 50*(2), 446-463.

Sénéchal, M., Ouellette, G., & Rodney, D. (2006). The misunderstood giant: On the predictive role of vocabulary to reading In S. B. Neuman & D. Dickinson (Eds.), *Handbook of early literacy* (pp. 173-182). New York: Guilford Press.

Sigman, M., McDonald, M. A., Neumann, C., & Bwibo, N. (1991). Prediction of cognitive competence in Kenyan children from toddler nutrition, family characteristcis and abilities. *Journal of Child Psychology and Psychiatry, 32*, 307-320.

Sigman, M., Neumann, C., Carter, E., Cattle, D. J., D'Souza, S., & Bwibo, N. (1989). Home interactions and the development of Embu toddlers in Kenya. *Child Development, 59*, 1251-1261.

Sigman, M., Neumann, C., Jansen, A. A. J., & Bwibo, N. (1989). Cognitive abilities of Kenyan children in relation to nutrition, family characteristics, and education. *Child Development, 60*(6), 1463-1474.

Sinha, D., & Misra, R. C. (1982). Some socio-environmental disadvantages and skill for perception of pictorial symbols. *Journal of Personality Study and Group Behaviour, 2*, 111-121.

Skibbe, L. E., Connor, C. M., Morrison, F. J., & Jewkes, A. M. (2011). Schooling effects on preschoolers' self-regulation, early literacy, and language growth. *Early Childhood Research Quarterly, 26*, 42-49.

Storms, G., Saerens, J., & Deyn, P. P. D. (2004). Normative data for the Boston Naming Test in native Dutch-speaking Belgian chldren and the relation with intelligence. *Brain and Language, 91*(3), 274-281.

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). New York, NY: Oxford University Press.

Suzman, S. M. (1987). Passives and prototypes in Zulu children's speech. *African Studies, 46*(2), 241-254.

Tatsuta, N., Nakai, K., Suzuki, K., Kurokawa, N., Hosokawa, T., & Satoh, H. (2013). Normative study of the Boston Naming Test in 7-year-old Japanese children. *Japanese Journal of Hygiene, 68*(3), 189-196.

UNESCO Institute for Statistics. (2011). Education (all levels) profile - Kenya. *UIS Statistics in brief*. http://stats.uis.unesco.org/unesco/TableViewer/document.aspx?ReportId=289&IF_Language=eng&BR_Country=4040&BR_Region=40540

van de Vijver, F., & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: an overview. *European Review of Applied Psychology, 54*, 119-135.

Vogt, P., Douglas, J. M., & Aussems, S. (2015). Early vocabulary development in rural and urban Mozambique. *Child Development Research, 2015*, 15. doi: 10.1155/2015/189195

Wachs, T. D. (1995). Relation of Mild-to-Moderate Malnutrition to Human Development: Correlational Studies. *The Journal of Nutrition, 125*(8), 2245-2254.

Wachs, T. D., Bishry, Z., Moussa, W., Yunis, F., McCabe, G., Harrison, G., . . . Shaheen, F. (1995). Nutritional intake and context as predictors of cognition and adaptive behaviour of Egyptian school-age children. *International Journal of Behavioural Development, 18*(3), 425-450.

Weizman, Z. O., & Snow, C. E. (2001). Lexical input as related to children's vocabulary acquisition: Effects of sophisticated exposure and support for meaning. *Developmental Psychology, 37*(2), 265-279.

Wenger, M. (1989). Work, play and social relationships among children in a Giriama community. In D. Belle (Ed.), *Children's Social Networks and Social Supports* (pp. 91-115). New York: Wiley.

Whaley, S. E., Sigman, M., Beckwith, L., Cohen, S. E., & Espinosa, M. P. (2002). Infant-caregiver interaction in Kenya and the United States: the importance of multiple caregivers and adequate comparison samples. *Journal of Cross-Cultural Psychology, 33*(3), 236-247.

World Health Organization. (2007). Growth Reference Data for 5-19 Years. Geneva: WHO.

Appendix 1. Battery of tests administered to children in the main study

a) The Tower of London (P. Anderson, Anderson, & Lajoie, 1996) is a non-verbal test of executive function that measures problem-solving and planning ability. The test makes use of three coloured balls which can be placed onto three wooden pegs of varying heights. The child is required to match a pattern presented on a stimulus card in a prescribed number of moves, while adhering to specified rules.

b) Dots (Fletcher, 1985) is a test of verbal memory where the child is required to point at a special dot on a sheet. The examiner progressively points at a series of one up to eight special dots from a series of designs.

c) The Self-Ordered Pointing Test (SOPT; Petrides & Milner, 1982) assesses verbal/visual selective reminding in terms of the capacity to initiate a sequence of responses, retain the responses and monitor the consequences of behaviour. The child is shown three series of six, eight, ten and twelve pictures each presented three times. The positions of the pictures differ on each page and the child is required to point at one picture on each page without pointing at the same one twice.

d) The Verbal List Learning (Delis, Kramer, Kaplan, & Ober, 1989) is a test of learning and working memory. The test consists of five serial verbal presentations of a 15-item word list composed of items semantically related to four common categories. Following each presentation, the child is asked to recall as many items as they can in any order. A second list of different items is read out once.

e) The Coloured Progressive Matrices (CPM; Raven et al., 1998) is a non-verbal test of reasoning. The 36-item test comprises a matrix of abstract patterns in a multiple-choice format which the child is required to complete by selecting the appropriate missing symbol from a set of six alternatives.

f) The Contingency Naming Test (CNT; P. Anderson, Anderson, Northam, & Taylor, 2000) is a test of executive function designed to assess response inhibition, attentional shift and cognitive flexibility. The child is taught a set of rules to name nine drawings consisting of a large outer coloured shape and a smaller inner coloured shape displayed in a single series.

g) The Score test (Manly, Robertson, Anderson, & Nimmo-Smith, 1999) is a test of auditory sustained and selective attention in which the child is required to place beads on one of two plates only after a special sound is heard on a cassette tape.

h) The People Search (Connolly & Grantham-McGregor, 1993; Connolly & Pharoah, 1993) is a test of visual sustained and selective attention. A stimulus sheet comprising complete and incomplete stick figures is presented. The child is required to cross out all the complete figures as quickly as possible.