



Published in final edited form as:

Child Neuropsychol. 2017 October ; 23(7): 761–802. doi:10.1080/09297049.2016.1216536.

Selecting Measures for Neurodevelopmental Assessment in Children in Low and Middle-Income Countries

Margaret Semrud-Clikeman¹, Regilda Anne A. Romero², Elizabeth L. Prado³, Elsa G. Shapiro¹, Paul Bangirana⁴, and Chandy C. John^{1,5}

¹Department of Pediatrics, University of Minnesota, Minneapolis, MN ²Licensed Psychologist, Keystone Behavioral Pediatrics ³Department of Nutrition, University of California Davis

⁴Department of Psychiatry, Makerere University, Kampala, Uganda ⁵Department of Pediatrics, Indiana University, Indianapolis, IN

Abstract

Diseases affecting millions of children in low and middle-income countries (LMIC), such as malnutrition, micronutrient deficiency, malaria, and HIV, can lead to adverse neurodevelopmental outcomes. Thus, a key health outcome in children is neurodevelopmental status. We reviewed the neurodevelopmental screening and testing measures most commonly utilized in LMIC, and created a matrix to help researchers and clinicians determine which measures may be most useful for various LMIC inquiries. The matrix was based on an internet literature review of 114 publications from January 1998 to May 2016 reporting the psychometric properties of instruments tested in LMIC children. Measures were classified as screening tests or more detailed tests that included both comprehensive batteries of general development and tests of specific domains. We have also had two experts as well as the authors review this paper for completeness. We provide an overview of tests used to date, including the benefits and drawbacks of each test, to provide researchers and developmental clinicians with a way to decide which tests may be best suited to their developmental assessment goals. Remarkable progress has been made in neurodevelopmental testing in children in LMIC over the past two decades but there remains a need for additional research in this area to develop new tests, better evaluate and adapt current tests, and assess test validity and reliability across cultures.

Keywords

LMIC; neurodevelopmental testing; reliability; validity

Overview

Introduction

The World Bank defines the world's economies based on gross national income (GNI) per capita for the previous year. As of July 2014 low income countries are defined as \$1,045 or

less GNI in 2013 and middle-income as those with a GNI of more than \$1045 and less than \$12,746 (World Bank, 2015). Iron deficiency, malaria, HIV, helminth infections and other disorders that affect neurodevelopment continue to afflict millions of children each year in low and middle income countries (LMIC) (Fernald, Kariger, Engle, & Raikes, 2009). Great progress has been made over the past twenty years in improving care of children with these diseases and child survival rates are improving (World Bank, 2011). Therefore, emphasis is now being placed on the long-term consequences of disease in survivors. As more children survive previously deadly diseases such as cerebral malaria and HIV, the high prevalence of disability and poor neurodevelopmental outcomes in low and middle-income countries may increase (Grantham-McGregor, Cheung, Cueto, Glewwe, Richter, Strupp, & International Child Development Steering, 2007; Murray & Lopez, 1997; Walker et al., 2007). A major result of the increase in survival rates from these diseases is the number of intervention programs in several countries that have now begun (Engle et al., 2007). The ability to measure outcome from these programs through well-validated and reliable measures continues to lag behind these efforts.

In LMIC, illnesses and malnutrition are common in the first 3 years of life, a period in which the brain is developing at a rapid rate. Diseases and negative experiences such as violence, abuse, malnutrition and neglect can seriously impact the development of neural networks and normal brain development (Spinazzola et al., 2005). How these negative experiences affect neurodevelopment, behavior, and quality of life of families and children is not yet fully delineated, particularly in LMIC. It has been estimated that over 200 million children under the age of 5 will be unable to reach their full potential due to disease and poverty (Grantham-McGregor, Cheung, Cueto, Glewwe, Richter, Strupp, & Group, 2007). Due to these awareness raising articles, there has been a push to provide improved food sources, remediate iodine and iron deficiencies, and improve remediation programs in such diverse countries as Guinea, Cape Verde, Burma, Vietnam, Nepal, Colombia, Bolivia, Turkey, and Bangladesh

Establishing a consensus for which tests of neurocognitive development should be used in specific LMIC cultures has been limited by the use of many different instruments that measure cognition in dissimilar ways, making comparability across studies or cultures difficult. Recent reviews have provided helpful information on test selection, adaptation, and assessment of preschool-aged children in LMIC (Gladstone et al., 2010; Prado et al., 2010), but these reviews focused on a specific disease process or culture, so their applicability to other cultures or other disease processes is limited.

The most comprehensive review of many measures used for preschoolers in LMIC and high-income countries provided additional information about neurodevelopmental tools (Fernald et al., 2009). Tests were selected for the review that had been found to discriminate between groups of children and focused on children aged 5 and under. The measures reviewed were those that are administered individually to the child as well as measuring a variety of domains. In addition, the authors evaluated whether the tests were psychometrically adequate, valid, and reliable, were enjoyable for the children, could be adapted to various cultures, and were easy to use.

This review is very helpful in that it reports how tests are best adapted in differing cultures and what is necessary for the test to be viable in LMIC. It also provides a blueprint for developing new tests in LMIC. The toolkit also lays out how to go about train individuals in different settings and how to make sure standardization is followed. In the final part of this toolkit the authors made recommendations for selection of assessments based on the following properties of the tests: psychometrically adequate (not defined), enough easy items, enjoyable, easy to adapt, easy to use in low-resource settings, not too hard to obtain or expensive, and able to be used with a wide age range. For this evaluation over 300 studies were evaluated using most of the search engines. Many of these studies, the tests, and the locations in which the tests were utilized are delineated in an excellent appendix. Many of the measures that were reviewed in this early paper are currently out of date and are rarely used in current research (e.g., British Ability Scales, Griffiths, McCarthy Scales, etc). For example the Griffiths was found to have a different factor structure with poorer South African children compared to the findings in the United Kingdom making its validity suspect for use particularly with that population (Luiz, Foxcroft, & Povey, 2006).

A more recent review published in 2014 evaluated 14 tools for screening developmental disabilities in children 3 and under which were developed in LMIC (Fischer, Morris, & Martines, 2014). This review focused mostly on children 0 to 3 and was particularly interested in measures that could be used by non-specialist primary healthcare providers. Of the 14 measures only 4 had adequate reliability and validity. Another review of tests by Robertson et al. (2012) found similar results with reliability and validity not being reported or developed for many of the measures.

The goal of the present review is to address update our knowledge in the following ways: (1) by providing comprehensive information on the strengths and limitations of specific tests that have been used to assess child neurodevelopment in LMICs, (2) reviewing the validity and reliability of these tests which had not been reported previously, (3) documenting the cognitive domains assessed by the major tests used, and (4) objectively assessing the clinical applicability of these measures in different settings. We define clinical applicability as the ability of the measure to be used in a manner consistent with its intended purpose, for it to be readily obtained in the LMIC, and for it to be possible for personnel to be adequately trained in the administration and interpretation of the results. Few measures have been evaluated for sensitivity and specificity for appropriate clinical diagnoses. While the positive predictive value (PPV), which is the probability that the score obtained truly indicates that the disorder/disease is present, unfortunately the PPV reported of many measures is not reported or studied. In cases where it is reported we have noted that in the tables. We also hope that this review will encourage further research into the psychometric properties of these measures. Many of these measures are general use without a concerted effort to make sure they are appropriate for the populations in LMIC. In addition, we did not seek to evaluate every possible instrument that is present; but rather to evaluate the most commonly used instruments as present in current literature and the most viable instruments that are readily available.

Key issues in assessment of neurodevelopmental testing in children in LMIC

Comprehensive testing vs. screening tests—It is important to distinguish between comprehensive measures of neurodevelopment and screening measures. Screening measures are brief tests to determine whether the child is at risk for developmental problems in one or more areas (for example, motor, language, cognition, or visual-reception). These measures may sample multiple areas and either provides a score or a qualitative determination of whether the child is at risk for difficulty and requires further assessment. Screening measures have been used in LMIC, but often without normative data or validation in the country of testing. It has been suggested that measures that do not have normative data for the country in which they are being used, should compare children with difficulty with a control group from the same country (Fernald et al., 2009). It was also concluded that norms from a different country may not be appropriate for conclusions as to the particular child's skill development. Screening measures should not be considered a comprehensive evaluation of a child's neurodevelopment.

In contrast to screening measures, comprehensive neurodevelopmental assessments generally utilize instruments that sample a broad range of skills and provide a standardized score from the child's performance based on large normative samples usually in developed countries. These tests usually provide scores that are characterized as developmental quotients or intelligence quotients as their goal is to assess overall cognition. A higher level of training is required to administer the comprehensive measures compared to the screening instruments. They are generally more specialized, so as to be able to provide recommendations for interventions and to measure progress following an intervention. Apart from measures of overall cognitive ability, comprehensive testing across multiple cognitive domains (e.g., executive function, gross and fine motor function, attention) may require more than one test measure (John et al., 2008).

Test adaptation and reliability and validity—Measuring the neurodevelopment consequences of such diseases requires reliable and valid tests of neurodevelopment. Without adequately studied neurodevelopmental assessment tools in children in LMIC, it is difficult to determine the full extent of factors contributing to neurocognitive impairment. A relatively small number of studies have assessed neurocognitive development across age groups in LMIC children or determined the validity and reliability of measures for a particular culture. It has been strongly suggested that adapting a test to a LMIC context requires the use of local psychologists if available as well as local informants from different disciplines including parents and nurses and teachers as well as pediatricians (Kammerer, Isquith, & Lundy, 2013). In addition, review of the translations by a multidisciplinary group to ensure the accuracy of the words is important. Kammerer et al. (2013) further suggest that the items should be videotaped to determine how the task is understood and managed by the child.

Validity is difficult to evaluate, particularly in the LMIC, because few measures have been used in each particular language and context. Moreover, tests that have been used in multiple studies may have been solely translated into the local language without validation with local norms or adapted to the local culture. While some measures may show acceptable

psychometric reliability, there is not an established benchmark or gold standard with which to gold standard to validate the target test. In the absence of a gold standard the use of other variables such as academic achievement and factor analysis can provide some evidence of convergent validity (Bangirana et al., 2009; Maida, van de Vliver, Srinivasan, Transler, & Sukumar, 2010). Other studies have used observation as well as community reports of functioning with which to validate a test. This practice is not widespread, as it requires attention to quantifying such observations, but may be a useful approach. Gladstone utilized reporters from Malawi in developing the MDAT as well as evaluating what children are able to do using focus groups (Gladstone et al., 2009). This technique is currently being used in developing measures in India and Ethiopia as reported on the Saving Brains Grand Challenges Canada website: <http://www.grandchallenges.ca/saving-brains/>.

One way to validate a measure is to determine which developmental milestones are seen in healthy children at the various ages in the target country (Fernald et al., 2009). Specific measures such as executive functioning and attention may yield the most information about neurodevelopment as they evaluate the child's ability to respond to day-to-day requirements, how the child responds to new situations, and what strategies a child employs to achieve a goal. These specific measures may be less dependent on the nature of the child's educational experiences than comprehensive measures of cognitive ability and may reflect abilities needed in all cultures.

Tasks that are unfamiliar to a child in an LMIC but familiar to children in Western cultures may measure the LMIC child's ability to adapt to new situations rather than their ability to complete the task rendering the scores invalid for the domain being measured. For example, in Malawi most children have not had experience playing with puzzles. A test with puzzles is a novel situation and may not tap the child's visual-spatial skills in the same way it does in the country where it was standardized. A critical review of nonverbal neuropsychological measures used in different cultures strongly suggests that significant differences arise among cultures and that these nonverbal measures may not be evaluating the same thing across cultures (Roselli & Ardila, 2003). Thus, the results cannot be interpreted in the same manner. These tasks require cultural sensitivity as well as knowledge as to what is appropriate in a particular culture. Methods of adapting and developing valid tests for LMIC requires understanding of the ways different cultures shape experience and hence the brain (Chiao, 2010). Solely translating a measure into another language without regard to the culture (some label this a universal translation), the child's experience, or how social experiences have shaped the child's world view may result in the test measuring skills in ways the original test did not intend (Chiao, 2010).

Validity can also be assessed as convergent validity through determining whether a test shows expected associations with the child's performance in the real world. For example, a study in Zambia found scores for the Draw a Person Test and a locally developed test (Panga Munthu Test) correlated less with local educational outcomes in girls than in boys (Serpell & Jere-Folotiya, 2011). In contrast, ratings of the child's ability by adults familiar with the children in the context of the local village correlated equally well with local educational outcomes in both boys and girls. Thus, the skills assumed to be important in Zambia in school and later in life were not the same as those predicted to be important in cognitive

tests developed and standardized in a Western country, and differed according to gender. Valid interpretation of the results of testing is dependent on this knowledge.

Locally developed tests vs. tests from high-income countries adapted to local context—There are two main strategies in adapting/adopting tests for use in a local context in a LMIC. The first is to adapt an existing test that has shown good validity and reliability in HIC and the second is to develop culturally specific tasks (Kammerer et al., 2013). Both of these practices come with strengths and with drawbacks. Cultures differ as to the emphasis put on some developmental tasks and what might be expected from a child in the U.S. can differ from one in Africa, Asia, or South America. Using Western norms may not fairly represent the child from LMIC accurately as these norms may be not be applicable due these children (Kirova & Hening, 2013). Experts have cautioned against using norms established in Western countries to apply to non-Western cultural backgrounds as there is a tendency to over-identify these children as delayed (Chow, Gokiart, Parsa, & Rajani, 2009; Heydon & Iannacci, 2008). To avoid this difficulty, it has also been suggested to do pre and post interventions and use the size of change to determine neurocognitive functioning (Boivin & Giordani, 2009).

Major issues are present in developing tests from scratch in LMIC including time, resources, costs, and most of all trained faculty in psychometrics. We are just beginning to appreciate how the brain interacts with the social world, how cognition shapes social behavior, and how cultural expectations can have effects on the developing brain and hence how tasks and skills are learned and adapted (Schutt, Seidman, & Keshavan, 2015). These aspects particularly affect the validity of the measures (Agiloti & Liuzza, 2011).

Because of difficulty of determining cross-cultural validity of a test developed in Western countries, development of neurocognitive test batteries in the local areas of study has been accelerating, with culturally appropriate questions and testing tools, for the LMIC children. Examples of these tests include the Malawi Developmental Assessment Test (MDAT) (Gladstone et al., 2010) and the Kilifi Developmental Inventory (KDI) (Abubakar, Holding, Van Baar, Newton, & van de Vijver, 2008). While these tests are ideally suited for the areas in which they were developed, they face the same questions as tests developed in Western countries when they are used in other countries. A test developed in one area or country of Africa may not be suitable for another area or country in Africa, let alone areas in Latin America or Asia.

Another limitation of appropriate local test batteries is the considerable resources they take to develop. One suggestion to improve the validity of measures is to use situational sampling. Situational sampling generates stimuli that reveals cultural variations in behavior (Chiao et al., 2010). This method provides the ability to determine that stimuli used across cultural contexts share a similar meaning; that is what one is measuring is what one thinks is being measured.

Finally, an advantage of valid tests that are applicable to multiple countries is that results can be compared across countries, particularly if the results in each country have age-normed data from healthy children in that country. This allows comparison with children with a

particular disease state so that cognitive consequences can be identified. Tests developed specifically for one area of an LMIC, though they have performed well in the populations of study, may be useful only in that area and results may not generalize. However, use of cut-offs and values for differentiation of abnormal and normal behavior can only occur with culturally specific normative data and clinical validation of a tool within the culture (Van Widenfelt, Treffers, De Beurs, Siebelink, & Koudijs, 2005).

For this review, we considered all tests, including those developed in high income countries that have had a rigorous evaluation of psychometric properties and for which results were published in at least one LMIC. We also evaluated tests developed for a specific LMIC population, as it is possible that these tests could have broader applicability. This review differs from previously published reviews because it emphasizes not only the use of measures, but also peer-reviewed data on the usefulness of the measures as well as their psychometric properties. We were surprised at how many of commonly used measures in the LMIC do not have peer-reviewed published data on specificity, sensitivity, validity, or reliability.

Methods

Developmental screening tests and neuropsychological tests for infants and preschool-aged children developed and normed in high income or low and middle-income countries (LMIC) were reviewed. The terms “child,” “children,” “infant,” “pediatric,” “early childhood,” or “preschool” were used together with the terms “development,” “cognitive,” “IQ,” “intelligence,” “neurocognitive,” “neuropsychological,” “neurodevelopment,” “attention,” “executive functioning,” “memory,” “behavior” and “adaptive functioning” to search for tests in *EBSCOhost*, *Google*, *Google Scholar*, *PubMed*, *PsychArticles*, *PsychInfo*, MEDLINE (PubMed), Psych Info, Educational Resources Information Center (ERIC), World Health Organization, the World Bank, the International Bureau for UNESCO (United Nations Educational, Scientific and Cultural Organization), LILACS (Latin American and Caribbean Health Services, UNICEF and *ProQuest*. The following inclusion criteria were used: (1) study was peer reviewed and published from January 1998 to February 2016, (2) publications that reported tests administered to children 8 years old or rating scales completed by parents/caregivers or teachers of children aged 8 years old, (3) information about what the test assessed, and (4) the test was used in LMIC. A total of 114 articles were identified, of which 60 were not utilized because the study or paper did not meet all of these requirements.

Information from test manuals, test publishers, correspondence with authors, and published validation/research studies were also obtained. Tests with published results and normative values were included in our evaluation. Tests developed and normed in LMIC with a strong theoretical/empirical basis but with limited validation/research studies were also included. While the most recent version or edition of the test was included in Table 1, in some cases the only research available was on older editions. In these cases, that is noted in the table.

We used the following criteria in the evaluation of the measures: adequate psychometric standards (age range, norms and standardization, reliability/validity, and administration/

scoring), several domains measured, the training needed for administration, the ease of administration, whether the child was directly evaluated or observed, and the availability of the measure. For some promising tests, this is a call for additional studies rather than a statement that this is an inferior test. For some well-regarded tests used in high-income countries (HIC), there was little to no information about use in LMIC. These measures include the Delis-Kaplan Tests of Executive Function, Differential Abilities Scale I and II, and Woodcock Johnson Achievement and Cognitive Batteries III and IV to name a few. These measures could not be assessed for applicability in LMIC and so were not included in the Tables 2 and 3.

Results and Discussion

Summary of Neurodevelopmental Tests used in LMIC

Table 1 presents the major measures that have been used in LMIC, and provides specific information about the instrument, languages/countries in which it has been used, training requirements, and whether the test is a screening test or a comprehensive measure for one or more areas of neurodevelopment. Table 2 (screening tests) and Table 3 (measures of development and neuropsychological tests) provide information about ages for which the test is appropriate, administration time, scoring methods, and neurocognitive domains assessed. Some tests included in Table 1 are not included in Tables 2 and 3 because the tests had limited use or information about use in LMIC.

Review of Screening Tests (Table 2)

The most used screening tools were the Strengths and Difficulties Questionnaire (Kashala, Elgen, Sommerfelt, & Tylleskar, 2005) and the Ten Questions Questionnaire (TQQ) (Birbeck et al., 2010; Durkin, Hasan, & Hasan, 1995; Mung'ala-Odera et al., 2004). These measures require little training for administration and interpretation, and are readily available in the LMIC. These measures have different purposes as illustrated in Tables 1 and 2. The TQQ was designed to be a screening tool for neurological difficulties in children aged 2 to 9 years. The SDQ screens for behavioral difficulties in children aged 3 and higher. They provide quick information as to the child's performance. They also have a wealth of research concerning their applicability in many countries.

While most of the screening measures show usefulness, they have been utilized in very few countries limiting their applicability for additional cultures. Assessment of the psychometric utility of these measures has just begun. These measures appear to be most appropriate for the countries in which they have been developed unless piloted in new areas. Moreover, some of the measures (Guide for Monitoring Child Development, Parental Report Scales, Shoklo Neuological Test, Test de Desarrollo Psicomotor) measure solely one or two domains, which are generally motor and language and do not provide sufficient information as to the child's development.

The Escala Argentina de Inteligencia Sensorimotriz, and Escala de Evaluacion del Desarrollo require observation of the child and determination of whether the child can perform certain tasks. These direct observation scales pose a particular challenge as the

examiners need to be carefully trained to be sensitive to developmental difficulties. These measures also required a significant level of training needed and the lack of generalizability across cultures beyond that of South America and/or Thailand.

There have been at least 5 screening measures developed for use in India; The Baroda Developmental Screening Test (Phatak & Khurana, 1991), Developmental Assessment tool for Anganwadis (Nair et al., 1991), the Disability Screening Schedule (DSS)(Chopra, Verma, & Seetharaman, 1999), Lucknow Developmental Screen for Indian Children (Bhave, Bhargava, & Kumar, 2010) and the Screening Test Battery for Assessment of Psychosocial Development (Vazir, Naidu, Vidyasagar, Lansdown, & Reddy, 1994). These measures have been solely used in India limiting applicability to countries outside of that region. Despite this limitation the Baroda and Disability Screening schedule show relatively good psychometric properties and are easy to administer.

The authors of a review of these measures (Fischer et al., 2014) concluded that no measure met criteria for use proposed by a panel of experts. Three screening measures (DSS, GMCD, and Ten Questions) had adequate overall psychometric properties and were more readily available and user friendly than the other measures. The DSS and GMCD had only been used in one context (India and Turkey) and need further validation. The Malawi Developmental Assessment Tool (MDAT) was reviewed also. The major drawback for the MDAT was the relatively long time for administration (35 minutes) and it was suggested that a screening measure developed from the MDAT would be useful.

There are several screening measures that have been used solely in only one or two countries and so applicability to other countries in the area is unknown. The Abbreviated Developmental Scale (van Meerbeke, Talero-Gutierrez, & Gonzalez-Reyes, 2007) was developed through interviews with parents and teachers as to expected developmental milestone and has become the most frequent used tool in Colombia (Robertson, Hatton, Emerson, & Yasamy, 2012). It has solely been used in Colombia. The Access Portfolio (Wirz, Edwards, Flower, & Yousafzai, 2005) is not the typical screening measure. It has 2 sections: identification and intervention suggestions. While psychometric properties were not reported, it was reported to be effective in identifying disabilities in Sri Lanka and Uganda (Robertson et al., 2012). The Guide for Monitoring Child Development (GMCD) has been used solely in Turkey and found in two reviews to be helpful (Fischer et al., 2014; Robertson et al., 2012). The Infant Neurological International Battery test used in Iran (Soleimani & Dadkhah, 2006) and the Monitoring Child Development (Lansdown et al., 1996) in Thailand, India, and China show promise but research demonstrating their usefulness is lacking.

General, Comprehensive and Specific Neurodevelopment Tests (Table 1)

Table 1 provides a selected list of most commonly utilized comprehensive and specific neurodevelopmental tests in LMIC. Tests have been divided according to the area of functioning that they measure. Not all of the measures listed in Table 1 are reviewed as some are out of date and others were not used in more than one study to determine applicability to

LMIC. Please refer to Tables 1 and 3 for additional information about the individual measures.

General Tests of Child Development

General tests of child development require more training and support for interpretation and administration than screening tests, but less than the comprehensive measures of cognitive functioning and can be used by medical personnel, nurses, and clinical specialists. The MDAT and Kilifi scales are scales that provide the strongest psychometric properties and which were developed for use in Africa. Both of these measures have been found to show adequate reliability (adequate reliability > 0.70 as defined by Sattler) (Sattler, 2001) and validity scores of > 0.21 . It is important to refer to Table 3 as these measures have different purposes. For example the MDAT is a general measure as described in the next paragraph while the Kilifi scale measures fine and gross motor abilities.

Cognitive Tests

Eleven measures designed to evaluate comprehensive child neurodevelopment through direct interaction with the child were found in the literature review. The BSID-II and III, and the KABC-II have been utilized in more populations than the other measures. While the WPPSI III and Stanford Binet have been utilized in many countries that are not LMIC, the psychometric applicability to LMIC has not been well established (Grantham-McGregor, Cheung, Cueto, Glewwe, Richter, & Strupp, 2007; Hamadani et al., 2011). The BSID II and II are designed for use with children aged 1 month to 42 months. The Bayley Scales of Infant and Toddler Development-III Screening Test (Bayley, 2005) was developed from the BSID III. It requires approximately 15 to 25 minutes to complete and is for ages 1 to 42 months. It measures cognitive, language, and motor ability. This measure has not been used outside of HIMC.

BSID

While the BSID is now in its third revision, many studies have utilized the second and even first version with findings of significant cultural differences with some difficulties also reported for the third revision. Similar difficulties were found in Nigeria with the BSID II. In this study scores dropped significantly at age 1 due to increased language demands (Ogunnaiké & Houser, 2002). When items were selected from the BSID II that were judged to be more culturally appropriate, no such decline was found. These items formed the Yoruba Mental Scale (see Tables 1 and 3). Cultural influences in performance have also been found on the BSID II with Lakota children in South Dakota, U.S.A. (Hagie, Gallipo, & Svien, 2003). When the individual items were inspected, linguistic and cultural differences within the Lakota culture appeared to preclude performance on some items. While an improvement on the BSID II is the BSID III, which now separates out the language skills from the overall cognitive scores, language and gross motor skills continue to be found to be different in at least one LMIC. Studies in Cameroon found significant differences on the BSID III that were influenced by cultural/language variations particularly in the areas of language and gross motor (Kolling et al., 2014; Vierhaus et al., 2011).

These empirical studies indicate that the BSID-II and -III may be more culturally affected than was previously believed. They also indicate that the use of Western norms may be inappropriate for these children with the BSID and BSID II, particularly if used to predict future performance or to evaluate the presence of an overall delay. These findings highlight the importance of testing healthy children in the area, and norming the test with scores from these children when using the test in a group of children with a specific disease. These instruments may prove useful for tracking the same child over time using a within-child approach, but a limited number of studies have done such follow-up.

Cogstate

CogState is a computerized measure with few language demands. It has been used in 7 countries and does not require translation as material is presented in picture/visual format except for list learning. CogState uses training items to reduce task novelty. It consists of 9 measures that evaluate a specific area of cognition. Not all of the scales need to be utilized for all patients and the test can be adjusted for individuals. Measures include visual-motor, spatial problem solving, speed of processing, visual attention, visual learning and memory, attention, working memory, social cognition and response inhibition. Many of these measures are designed for adults but can be used for children particularly if concussion, traumatic brain injury, or neurological damage is suspected.

Boivin et al. (2010) found that CogState scores significantly correlated with the global scales for children on the KABC-II, and KABC-II scores correlated with level of education and educational achievement (class rank) in Uganda (Boivin et al., 2010), providing additional support for the Cogstate. Studies that directly correlate CogState with educational achievement would lend further credence to interpretation of the validity of this measure.

While CogState has been used somewhat with children in the LMIC, it has been used frequently with children in the US and Western countries. It has been found to be useful with children with ADHD (Molica, Maruff, & Vance, 2004). In addition, CogState has been adopted for use for high-risk pre-B acute lymphoblastic leukemia in one center and is recommended for use as a common protocol for assessment of cognitive functioning after treatment (Noll et al., 2013). It has also been used in determining recovery from concussion particularly following a sports injury (Schatz & Zillmer, 2003). The use of CogState for these populations has not been demonstrated in LMIC but would be useful as a next step in research.

CANTAB

The Cambridge Neuropsychological Testing Automated Battery (CANTAB) was originally designed for adult populations particularly to evaluate patterns of cognitive decline. The CANTAB focuses on measuring functioning of the frontal and temporal lobes. There are three domains: Visual memory, visual attention, and working memory/planning. Within each domain there are 4 subtests. In the visual memory domain are measures of delayed memory and paired associate learning tasks. In the visual attention section are measures of set-shifting tasks and continuous performance. The working memory/planning domain uses measures of spatial memory span, searching and a Tower of London planning task.

More recent research has expanded the use of the CANTAB to pediatric and adolescent populations. While children with developmental disorders have been tested using the CANTAB (Luciana, Sullivan, & Nelson, 2001), research as to how healthy children perform on the measure has been strongly suggested (Luciana & Nelson, 2002). Luciana and Nelson (2002) evaluated children aged 4 to 11 years that spoke English as well as a sizeable sample of those for whom English was a second language (German, Russian, Spanish, Middle-Eastern or Indian, African or Ethiopian, and Asian were the primary languages). Analysis of performance based on language proficiency found that the primary English-speaking children performed no differently from the children for whom English is a second language. This study also found that the CANTAB is useful in children with diverse language skills. It was concluded that the CANTAB may be very useful for children with weak verbal skills due to its low demand on language proficiency. A recent study found that the reaction time subtest was the most reliable and stable test of the subtests with stability found for the visual memory, reaction time, and visual information processing subtests (Syvaaja et al., 2015). It was found that the set shifting measure and the spatial recognition memory subtests were not reliable or stable with typically developing children.

While the CANTAB may be useful for children assessed in LMIC, a literature search did not find any published, peer-reviewed articles that demonstrate its usefulness. A recent article indicated some difficulty with reliability and stability of this measure with typically developing children in a HIMC (Syvaaja et al., 2015). For this reason the use of CANTAB in LMIC suggests that this instrument may not be appropriate.

KABC-II

The KABC-II and its predecessor the KABC have been used extensively in sub-Saharan Africa, including in Uganda, Kenya, Senegal, and Malawi (Bangirana et al., 2009; Boivin, 2002; Holding et al., 2004). In Uganda the factor structure of the KABC-II has been found to be similar to the U.S. (Bangarina et al., 2009). Other studies have also used the KABC and KABC-II, and have found that children with cerebral malaria performed poorly on measures of visual-spatial processing and memory (Boivin, 2002; Boivin et al., 2007).

Similarly, published reliability and validity data are within generally accepted guidelines (Sattler, 2001) and have been found for the use of the KABC-II with populations in India (Malda et al., 2010) and the KABC in Zaire (Giordani, Boivin, Opel, Nseyila, & R.E., 1996) (see Table 3). Malda et al. (2010) stressed measures such as the KABC-II need to be validated through analysis of how the measure predicts school success in LMIC. Giordani et al. (1996) found lower scores for the children in Zaire compared to U.S., which likely reflects cultural differences as reflected on the test. Specific portions of the KABC-II have been shown to be inappropriate for children in specific LMIC. For example, in Uganda, Riddles, Verbal Knowledge, Expressive Vocabulary and some items in Gestalt Closure were not culturally appropriate (Bangirana et al., 2009). Studies that provide evidence of reliability and stability within the same child are being completed in Malawi and Uganda. Of note, most studies of the KABC and KABC-II have been used with children 5 years of age, so the validity and reliability of this test in children age 3–5 years remains to be demonstrated.

The Mullen Scales of Early Learning (MSEL) (Mullen, 1995) has been studied in Benin (Durkin et al., 1995) and Uganda (Bangirana, Menk, John, Boivin, & Hodges, 2013). The reliability of the MSEL in the HIV-exposed or infected children and in the children with severe malaria (Bangirana et al 2014) was reported as good (see Table 3), and minimal modifications were required for its use in these populations. An advantage of the MSEL is that the test can be used in children from birth to 68 months, so they can be used in longitudinal studies to follow children from birth to 5 years, whereas the BSID can only be tested in children from birth to 42 months, and the KABC II starts at age 3 years, and has not been extensively testing in LMIC children between 3 and 5 years of age. A drawback for the Mullen is that the norms have not been recently updated and some sections are highly language-based.

Similar to the MSEL is the Rapid Neurodevelopmental Assessment Instrument (RNAI). The RNAI has been developed for use in Bangladesh with acceptable reliability and validity (Kahn et al., 2010) (see Table 2). The drawback on this measure is that it has been used in only one country. It does show promise and is a measure that should be evaluated for more general use.

Fernandes et al. (Fernandes et al., 2014) used an advisory panel that evaluated 47 tools to develop a new multi-dimensional assessment, the Intergrowth-21st Neurodevelopmental Assessment (INTER-NDA). A panel of international child development experts designed it. This assessment included measures of auditory evoked potential, cognition, language skills, behavior, motor skills, and attention. In addition sleep-wake patterns were assessed. This program has been used in Brazil, India, Italy, Kenya and the United Kingdom in children aged < 14 week of gestation to 2 years of age. This measure, however, has a very narrow age range from 22 to 26 months of age. It has been designed for children from middle and upper class families and the author's caution about its use for children from poorer backgrounds. It also includes test that could be quite expensive. Its main design was to develop a population based screening instrument for early childhood disability.

Attention

A few measures of attention have been piloted in Sub-Saharan Africa. The Test of Variables of Attention (TOVA) (Greenberg, Kindschi, & Corman, 2000) can be used in individuals from 5 years of age to adulthood. This measure has been utilized in Uganda, Senegal, and Malawi to date. While studies in Uganda, Malawi, and Senegal have found the TOVA as reliably identifying attentional deficits in children who have recovered from cerebral malaria, psychometric properties are not reported in these peer-reviewed articles (Boivin, 2002; Boivin et al., 2007; John et al., 2008). The Conners' Continuous Performance Test II (Conners, 1994) has been used in Cameroon (Ruffieux et al., 2010). While reliability was not reported in this one study, the test showed factor loadings on the executive function factor. An important difference between the Conner's and the TOVA is that the Conner's uses numbers and letters while the TOVA uses a square either at the top of the screen or at the bottom. The use of numbers and letters may be culturally biased. No studies have compared these measures across LMIC to determine whether this suggestion is, indeed, valid.

Picture Search from the Tests of Everyday Attention for children (Manly et al., 2001) has been used in Indonesia (Sakti et al., 1999), China (Nokes et al., 1999) and Zambia (Nampijja et al., 2010). Reliability from the Zambian study found good reliability (.76) and internal consistency (.84). Further research is needed to establish this measure as useful in determining attentional ability.

Kitsao-Wekulo et al. (Kitsao-Wekulo, Holding, Taylor, Abubakar, & Connolly, 2012) evaluated 3 measures of attention in a group of children aged 8 to 11 in Kenya. The Contingency Naming Test (CNT: attention and attention shifting) (Anderson, Anderson, Northam, & Taylor, 2000), Score (Auditory sustained and selective attention) (Manly, Robertson, Anderson, & Nimmo-Smith, 1999) and People Search (visual sustained and selective attention) (Holding et al., 2004) were used. Reliability was not generally within expected limits (see Table 3). Validity was not psychometrically established. Score and CNT correlated significantly with school exposure while People search did not correlate with any background variables.

For children 18 months to 5 years of age, the Early Childhood Vigilance Test (ECVT) (Goldman, Shapiro, & Nelson, 2004; Shapiro, 2010) was used to assess children in Uganda who have recovered from severe malaria (Bangirana et al., 2014). The coding system requires extensive training, is labor and time-intensive to complete, and the measure has only been used in one LMIC. Attention appears to be significantly affected in illnesses like severe malaria (John et al., 2008) so it is important to directly assess this area, which most general cognitive tests do not. For this reason, further study of the ECVT, which is the only instrument generally available that can do this assessment in children as young as 18 months, in other countries will be important.

Executive function

Similar to attention, tests are limited that have been used to assess executive function in children in LMIC. The NEPSY II has been mostly tested in countries in the upper middle income (South Africa, Romania and Iran) and the NEPSY in Zambia (Mungala-Odera & Newton, 2007) and Uganda (Kitsao-Wekulo et al., 2012). The word generation and inhibition subtests from the NEPSYII have been tested in South Africa (Dalen, Jellestad, & Kamaloodien, 2007). The NEPSY (not the revision NEPSY-II) has also been used in Iran and in Romania (Abedi, Malekpour, Oraizi, Faramarzi, & Paghale, 2012) (Visu-Petra, Benga, & Miclea, 2007). The Romanian study and Zambian studies (NEPSY) as well as the South African study (NEPSY II) found that strong language skills were needed for success even for English speaking subjects; this language component should be carefully evaluated as this test becomes adapted to others areas of the world.

A Ugandan study evaluated subtests from the NEPSY of working memory (sentence repetition, verbal fluency), impulse control (knock tap) and used the Wisconsin Card Sort as a measure of cognitive flexibility (Nampijja et al., 2010). Test retest reliability for these measures was found to be acceptable for sentence repetition and verbal fluency (see Table 2) but not for Knock Tap (.38), and Wisconsin Card Sort (.23). Good internal consistency was also found for sentence repetition, verbal fluency, and the knock tap game. It is important to note that the knock tap game did not have good test-retest reliability (.28). Validity was not

established for the Wisconsin Card Sorting Test because internal consistency could not be established due to the nature of the task. It is important to note that poor reliability makes tests suspect for use even if shown to be internally consistent. The use of the Knock-Tap Game and Wisconsin Card Sorting Test should be used with extreme caution. Additionally tests of executive function may be most susceptible to practice effects complicating the ability to measure validity as seen above for the Wisconsin Card Sorting Tests. This problem has arisen in studies of disease outcome in Western countries particularly in cancer research where serial evaluations are needed. In this case CogState has been used because it has shown fewer practice effects (Noll et al., 2013)

One study reported on the use of individual executive functioning subtests and provided reliability and validity information (Kitsao-Wekulo et al., 2012). The Tower Test (Shallice, 1982), Verbal List Learning (VLL: working memory) (Delis, Kramer, Kaplan, & Ober, 1994), and Coloured Progressive Matrices (CPM: reasoning) (Raven, Court, & Raven, 1998) were evaluated in Kenya as part of a larger study. Test retest reliability for the Tower was found to .654, for the VLL .43, and for the CPM .77. These reliabilities are a bit lower than would be expected in HIC. Although validity was not established against a gold standard, these tests correlated significantly with school exposure ($p < .0001$).

The Behavior Rating Inventory of Executive Functions (BRIEF) (Gioia, Isquith, Guy, & Kenworthy, 2000) is a rating scale that can be completed by the parent or major caregiver as well as the teacher and at older ages by the client. It is being adapted for use in Malawi and Uganda. In Ugandan children infected with HIV, higher scores on the BRIEF were associated with higher viral load (Ruiseñor-Escudero et al., 2015) while HIV infected children meeting the DSM criteria for ADHD had higher BRIEF scores (Burkey et al 2015). A validation study in Uganda of the BRIEF using factor analysis resulted in two main factors comprising of the Behavioral Regulation and the Meta Cognition scales (Familiar et al 2015).

One of the issues with the BRIEF in Western countries is the lack of positive correlation with direct measures of executive functioning. Studies have not found a positive relationship between the BRIEF and the Wisconsin Card Sorting Test, Trails B, and verbal fluency (Vriezen & Pigott, 2002). Others have found the BRIEF to be sensitive in identifying children with ADHD performance measures of executive functioning (Toplak, Bucciarelli, Jain, & Tannock, 2008). An interesting finding from this study was that the individual subscales of the BRIEF (i.e., working memory) did not correlate significantly with direct measures of executive functioning. These findings suggest that the BRIEF may reflect that the child's inability to perform these skills in day-to-day life while individual measures of executive functioning may show that the child can perform the task.

Memory

Studies have found that children who recover from cerebral malaria show particular problems with working memory (Atlantis subtest) as measured by the KABC-II (Bangirana et al., 2011; Boivin et al., 2011; John et al., 2008). Further study is needed to demonstrate how these subtests relate to every day difficulties with memory for these children. The

NEPSY II short term and delayed memory subtests are currently being used in a study in Bangladesh sponsored by the Grand Challenges Canada, but is not currently published.

Kitsao-Wekulo et al. (2012) utilized memory measures with children aged 8 to 11 in Kenya. These measures included self-ordered pointing test (Petrides & Milner, 1981)(SOPT: verbal and visual selective reminding), and Dots (nonverbal memory) (Fletcher, 1985). While validity studies against a gold standard were not performed, the memory subtests showed significant positive correlations with school exposure but not nutritional status, household status, or area of residence.

The Color Object Association Test (COAT) (Jordan, Johnson, Hughes, & Shapiro, 2008) has been utilized as an assessment of memory for younger children (aged 18 to 60 months). There are currently two published studies to date using this test in an LMIC, both from Uganda (Bangarina et al., 2014; Boivin, Bangirana, et al., 2013). The measure was reported to have good reliability although the precise reliability co-efficient was not reported. It was also found to detect differences in memory between children with cerebral malaria and community children. This test requires further testing in other populations before it can be widely recommended, and administration of the test is complex, but it holds promise because it can assess memory down to a very young age.

Adaptive behavior

Adaptive behavior is an area that is just beginning to be examined. The International Classification of Functioning, Disability, and Health (ICF) (World Health Organization, 2007) divides adaptive behaviors into those activities a child is able to perform but doesn't perform independently and those that are completed independently (Oakland, 2007). The Vineland Adaptive Behavior Scales (VABS) (Sparrow, Cicchetti, & Balla, 2005) has been used successfully in South Africa with families with HIV (Allen et al., 2014).

The Independent Behavior Assessment Scale (IBAS, Level 2) (Munir, Zaman, & McConachie, 1999) has been used in Bangladesh, Jamaica and Pakistan. Psychometric properties were found to be acceptable. This measure has not been widely adapted for use outside of these three countries.

As with many of the tests reviewed, further studies are required before the VABS or the IABS can be generally recommended, but studies of adaptive behavior would be very useful, as there are few tests available that can assess adaptive behavior, and this is clearly an important functional area for a child with illness. It is highly likely that adaptive behavior and expectations of daily living skills and socialization are culturally dependent. Community input into the appropriate types of questions to be utilized for adaptive behavior assessment and local normative data will be important for any future studies.

A measure which is just beginning to be adapted for use in LMIC and Middle income countries is the Children's Assessment of Participation and Enjoyment (CAPE) and a companion measure, Preferences for Activities of Children (PAC). There is no published data on these measures outside of Canada so they are not included in Table 3. These measures evaluate a child's self report and are designed for ages 6 to 21. Questions about

what types of activities the child enjoys are presented (King et al., 2004). A recent study in Canada established the usefulness of these measures for children with physical disabilities (King et al., 2006).

An adaptation of the original Vineland Adaptive Behavior Scales has been developed in Indonesia (Tombakan-Runtukau & Nitko, 1992). While this adaptation is promising, there is a lack of validation of the measures to observed performance of the child. The initial findings are of a similar factor structure to the original VABS.

Motor

There are few studies that have evaluated fine motor and gross motor skills in the LMIC. The Bruininks-Oseretsky Test of Motor Proficiency (second edition) (BOT-2) (Bruininks & Bruininks, 2005) has been used to assess motor skills in HIV infected children and in children with Konzo disease in sub-Saharan Africa (Buomoko et al., 2015). HIV infected children and those with Konzo had lower motor scores than controls. BOT-2 scores correlated with a neurologic index score in children with Konzo ($r = -0.43$, $p < 0.0001$), no other psychometric results of the BOT-2 were provided in these studies (Boivin, Okitundu, et al., 2013; Buomoko et al., 2015).

The BOT-2 was also used in the north west province of South Africa with first graders to determine the relation of this measure of academic performance as well as to the Visual Motor Integration test (VMI) (Beery, 1997) (Pienaar, Barhorst, & Twisk, 2014). Findings were that a strong relationship was found between the VMI and academic performance as well as between the BOT-2 and overall academic performance. A relationship was also found between low SES schools and academic performance. These findings begin to provide validation for this measure through strong correlations with academic performance.

The Movement Assessment Battery for Children-2 (MABC-2) (Henderson, Sugden, & Barnett, 2007) is another test that has been used in Asia, South America and South Africa for motor assessment especially in children with developmental coordination disorder. This battery differs significantly from the original battery. It provides qualitative ratings for significant movement difficulty, at risk, and no difficulty. The therapist is able to rate the child's performance and supplement the numerical findings from the direct assessment. One study concluded that the MABC-2 is useful with children with various neurodevelopmental disorders across HIC and LMIC (Brown & Lator, 2009). The limitation of this measure was the lack of research on validity and reliability—a more recent literature search did not find any additional information on this score. It was suggested that this test be used with caution due to the lack of reliability information or construct validity.

Both the Bayley MSEL scales have been used to assess for fine motor and gross motor functions in children. In children with malaria, deficits in fine motor skills were more pronounced in children with cerebral malaria than those with severe malarial anemia while another study found that helminth infection was associated with poor gross motor function (Bangarina et al., 2016; Mireku et al., 2015).

Fine motor skills have been assessed using the Purdue Pegboard. One study that utilized the Purdue Pegboard (Ruffieux et al., 2010) found that the pegboard was useful in fine motor evaluation. Nampijja et al. (2010) evaluated bead threading, the coin box (adapted from the Kilifi Developmental Inventory) and balancing on one leg (taken from the Movement Assessment Battery for Children (Henderson & Sugden, 1992) to be useful measures in the LMIC. Reported reliability for the Coin box (.45), bead threading (.34), and balancing (.64) is less than desirable while internal consistency measures for these tasks is within acceptable limits (.76 to .80). The authors suggest that balancing on one leg was the only stable measure. It is important to note here that none of these measures shows strong stability.

Emotion and Behavior

The Child Behavior Checklist (CBCL) (Achenbach, 2010) is one of the most highly studied measures in LMIC, but review of the literature shows that test adaptation may be required, and that reliability may be sub-optimal in some settings. It has been translated into over 80 languages and there are over 7,000 published articles using the CBCL (Achenbach, 2010). The CBCL is also often adapted to an interview format, which is a different format from the original questionnaire administration. Reliability for the measure in one LMIC (Uganda) was somewhat lower than would be expected for tests in high-income countries, ranging from .64 for the internalizing scales, .78 for externalizing scales, and .83 for total problems (Bangarina et al., 2009). Generally minimal test-retest reliability for tests administered in high-income countries is required to be above 0.8 (Sattler, 2001). Of the individual scales, only aggressive behavior meets this standard (.82), and withdrawn/depressed (.73), somatic complaints (.74) and social problems (.73) approach the level considered acceptable for reliability in studies in high-income countries. Significantly poorer reliabilities were found for anxious/depressed, thought problems, and rule-breaking behavior. Validity of this measure is questionable when reliability indices are poor. None of the CBCL Diagnostic Statistical Manual scales were within acceptability levels for reliability with oppositional behavior being the highest at .70 and affective problems low at .45. These findings suggest that the CBCL, at least in sub-Saharan Africa, needs to be more strongly validated and if used, should be used with caution. Higher reliabilities have been found in Mexico (.90 for internalizing, .94 for externalizing, and .97 for total problems) (Albores-Gallo et al., 2007) and China (internalizing, .76, externalizing .81, Total problems .83) (Leung et al., 2006). Thus, the use of summary scales, internalizing and externalizing, should be used as the reliabilities for these scales is acceptable.

Conclusions

Our review of the extant literature provides support for the use of a limited number of measures in LMIC. For comprehensive or detailed neurodevelopmental tests, the need for in-depth training is a major obstacle, as all accurate, valid, and reliable tests to date of comprehensive or detailed neurodevelopment have required significant training. Other tests assessing various neurodevelopmental domains show promise in limited studies. These tests include the Cogstate, MDAT, and Rapid Neurodevelopmental Assessment Instrument, for general neurodevelopmental testing, the TOVA and ECVT for testing of attention, the NEPSY-II (selected subtests) for executive function, and the VABS and IABS for adaptive

behavior (with limitations). When these tests are studied further in different cultural settings, and validated and shown to be reliable in these settings, they may become valuable tools for assessment of these domains across a variety of cultures. This review points out the dearth of information that is currently published as to the sensitivity and specificity of these measures (clinical utility). While some of these measures are fairly commonly utilized and scores used to determine interventions, it is concerning that so little data is present as to the reliability and validity of these measures. There is even less data provided as to Positive Predictive Value (PPV). In some cases the number of children who are found to have difficulties on the measures which have been studied are at a high rate of false positives. This finding is particularly disturbing as it means that children are identified as having difficulty where none exists. Further evaluation of these tests to provide information as to not only the reliability of the measures but also the validity (particularly ecological validity) is crucial.

In summary, the recent interest in neurodevelopmental assessment of children in LMIC has resulted in a number of important findings about how different diseases and cultural factors may affect a child's neurodevelopment. Further research is required to determine whether specific screening and neurodevelopment tests can be used across multiple cultures and countries, but significant progress has been made in recent years in introducing testing in multiple neurodevelopmental domains. The tables in this paper provide a framework for those planning neurodevelopmental testing to assess which tests may be most appropriate for the specific disease or context in which they wish to do testing, or for the specific area of development they wish to test, and provide guidance as to areas that may require adaptation. Additional studies using appropriately adapted measures, and/or new locally-developed measures of neurodevelopment hold the promise of increasing understanding of the burden of neurodevelopmental impairment in children in LMIC and will allow accomplishment of the long-term goal of interventions to prevent or treat this impairment.

Databases searched were *EBSCOhost*, *Google*, *Google Scholar*, *PubMed*, *PsychArticles*, *PsychInfo*, MEDLINE (PubMed), Psych Info, Educational Resources Information Center (ERIC), World Health Organization, the World Bank, the International Bureau for UNESCO (United Nations Educational, Scientific and Cultural Organization), LILACS (Latin American and Caribbean Health Services), and UNICEF. Keywords used were “developing countries”, or ‘developing nations’ or ‘third world’ or ‘LMIC’ and ‘child,’ ‘children,’ ‘infant,’ ‘pediatric,’ ‘early childhood,’ or ‘preschool’ were used together or separately with the terms ‘development,’ ‘cognitive,’ ‘IQ,’ ‘intelligence,’ ‘neurocognitive,’ ‘neuropsychological,’ ‘neurodevelopment,’ ‘attention,’ “executive functioning,” “memory,” “behavior” and “adaptive functioning, ‘child development,’ ‘cognition’ ‘neurocognitive tests’, ‘reliability’, ‘validity’, ‘psychometrics’, ‘neuropsychological’, ‘developmental measure’, ‘motor’, ‘language’, ‘assessment’

References

- Abedi A, Malekpour M, Oraizi H, Faramarzi S, Paghale SJ. Standardization of the neuropsychological test of NEPSY on 3–4 years old children. *Iranian Journal of Psychiatry and Clinical Psychology*. 2012; 18:52–60.

- Abubakar A, Holding P, Van Baar A, Newton CRJC, van de Vijver FJR. Monitoring psychomotor development in a resource-limited setting: An evaluation of the Kilifi Developmental Inventory. *Annals of Tropical Paediatrics*. 2008; 28:217–226. [PubMed: 18727851]
- Achenbach, TM. The multicultural supplement to the manual for the ASEBA Preschool forms and Profiles ADM Modules for ages 11/2–5 with multicultural options. Burlington, VT: ASEBA; 2010.
- Agiloti SM, Liuzza MT. Social and cultural neuroscience: At the intersection between brain and social sciences. *Sistemi Intelligenti*. 2011; 23:359–366.
- Albores-Gallo L, Lara-Munoz C, Esperon-Vargas C, Cardenas Zetina JA, Perez Soriano AM, Villanueva CG. Validity and reliability of the CBCL/6-18. Includes DSM scales. *Actas Espanolas de Psiquiatria*. 2007; 35:393–399. [PubMed: 18004676]
- Allen AB, Finestone M, Eloff I, Sipsma H, Makin J, Triplett K, et al. The role of parenting in affecting the behavior and adaptive functioning of young children of HIV-infected mothers in South Africa. *AIDS and Behavior*. 2014; 18:605–616. [PubMed: 23892768]
- Anderson P, Anderson V, Northam E, Taylor HG. Standardization for the Contingency Naming Test (CNT) for school-aged children: A measure of reactive flexibility. *Clinical Neuropsychological Assessment*. 2000; 1:247–273.
- Bangarina P, Nakasujja N, Giordani B, Opoka RO, John CC, Boivin MJ. Reliability of the Luganda version of the Child Behavior Checklist in measuring behavioural problems after cerebral malaria. *Child and Adolescent psychiatry and Mental Health*. 2009; 3:38–41. [PubMed: 19995426]
- Bangarina P, Opoka RO, Boivin MJ, Idro R, Hodges JS, John CC. Neurocognitive domains affected by cerebral malaria and severe malarial anemia in children. *Learning and Individual Differences*. 2016; 46:38–44. [PubMed: 27212870]
- Bangarina P, Opoka RO, Boivin MJ, Idro R, Hodges JS, Shapiro E, John CC. Severe malarial anemia is associated with long-term neurocognitive impairment. *Clinical Infectious Diseases*. 2014; doi: 10.1093/cid/ciu293
- Bangirana P, Menk J, John CC, Boivin MJ, Hodges JS. The association between cognition and academic performance in Ugandan children surviving malaria with neurological involvement. *PLoS Medicine*. 2013; 8:e55653.
- Bangirana P, Musisi S, Boivin MJ, Ehnvall A, John CC, Bergemann TL, Allebeck P. Malaria with neurological involvement in Ugandan children: Effect on cognitive ability, academic achievement and behaviour. *Malaria Journal*. 2011; 10:334–341. [PubMed: 22047193]
- Bangirana P, Seggane-Musisi C, Allebeck P, Giordani B, John CC, Opoka OR, Boivin MJ. A preliminary examination of the construct validity of the KABC-II in Ugandan children with a history of cerebral malaria. *African Health Sciences*. 2009; 9:186–192. [PubMed: 20589149]
- Bayley, N. Bayley Scales of Infant and Toddler Development III (screening test). San Antonio, TX: Pearson; 2005.
- Beery, KE. The Visual-Motor Integration Test. 4th. Austin, TX: Pro-Ed; 1997.
- Bhave A, Bhargava R, Kumar R. Development and validation of a new Lucknow Development Screen for Indian children aged 6 months to 2 years. *Journal of Child Neurology*. 2010; 25:57–60. [PubMed: 19448137]
- Birbeck GL, Molyneux ME, Kaplan PW, Seydel KB, Chimalizeni Y, Kawaza K, Taylor TE. Blantyre Malaria Project Epilepsy Study (BMPES) of neurological outcomes in retinopathy-positive paediatric cerebral malaria survivors: A prospective cohort study. *Lancet*. 2010; 9:1173–1181. [PubMed: 21056005]
- Boivin MJ. Effects of early cerebral malaria on cognitive ability in Senegalese children. *Journal of Developmental and Behavioral Pediatrics*. 2002; 23:353–364. [PubMed: 12394524]
- Boivin MJ, Bangarina P, Byarugaba J, Opoka RO, Idro R, Jurek AM, John CC. Cognitive impairment after cerebral malaria in children: A prospective study. *Pediatrics*. 2007; 119:e360–366. [PubMed: 17224457]
- Boivin MJ, Bangirana P, Nakasujja N, Page C, Shohet C, Givon D, Klein PS. A year-long caregiver training program to improve neurocognition in preschool Ugandan HIV-exposed children. *Journal of Developmental and Behavioral Pediatrics*. 2013; 34:269–278. [PubMed: 23535340]

- Boivin MJ, Busman RA, Sujal M, Parikh A, Bangirana P, Page CF, Giordani B. A pilot study of the neuropsychological benefits of computerized cognitive rehabilitation in Ugandan children with HIV. *Neuropsychology*. 2010; 24:667–673. [PubMed: 20804255]
- Boivin, MJ., Giordani, B. Neuropsychological assessment of African children: Evidence for a university basis to cognitive ability. In: Chiao, JY., editor. *Cultural Neuroscience: Cultural influence on brain function*. New York: Elsevier; 2009. p. 113-135.
- Boivin MJ, Gladstone MJ, Vokhiwa M, Birbeck GL, Magen JO, Page C, Taylor TE. Developmental outcomes in Malawian children with retinopathy-confirmed cerebral malaria. *Tropical Medicine and International Health*. 2011; 16:263–271. [PubMed: 21143354]
- Boivin MJ, Okitundu D, Makila-Mabe Bumoko G, Sombo MT, Mumba D, Tylleskar T, Tshala-Katumbay D. Neuropsychological effects of konzo: A neuromotor disease associated with poorly processed cassava. *Pediatrics*. 2013; 131:1231–1239.
- Brown T, Lalor A. The movement assessment battery for children-second edition (MABC-2): A review and critique. *Physical and Occupational Therapy in Pediatrics*. 2009; 29:86–103. [PubMed: 19197761]
- Bruininks, RH., Bruininks, BD. *Bruininks-Oseretsky Test of Motor Proficiency*. second. San Antonio, TX: Pearson; 2005.
- Buomoko GM, Sadiki NH, Rwatambuga A, Kayembe KP, Okitundu DL, Mumba ND, Tshala-Katumbay D. Lower serum levels of selenium, copper, and zinc are related to neuromotor impairments in children with konzo. *Journal of Neurological Science*. 2015; 349:149–153.
- Chiao JY. At the frontier of cultural neuroscience: Introduction to the special issues. *SCAN*. 2010; 5:109–110. [PubMed: 20592041]
- Chiao JY, Hariri AR, Harada T, Mano Y, Sadato N, Parrish TB, Iidaka T. Theory and methods in cultural neuroscience. *Cultural Neuroscience*. 2010; 5:356–361.
- Chopra G, Verma IC, Seetharaman P. Development and assessment of a screening test for detection childhood disabilities. *Indian Journal of Pediatrics*. 1999; 66:331–335. [PubMed: 10798079]
- Chow, W., Gokiert, R., Parsa, B., Rajani, N. *Cross-cultural lessons: Early childhood developmental screening and approaches to research and practice*. Edmonton, Alberta: Community University Partnership; 2009.
- Dalen K, Jellestad F, Kamaloodien K. The translation of the NEPSY II to Afrikaans, Some ethical reflections. *Cognitie, Creier, Compartament*. 2007; 11:609–620.
- Delis, DC., Kramer, JH., Kaplan, E., Ober, BA. *The California Verbal Learning Test: Children's Version*. San Antonio, TX: Psychological Corporation; 1994.
- Durkin MS, Hasan ZM, Hasan KM. The ten questions screen for childhood disabilities: Its uses and limitations in Pakistan. *Epidemiology*. 1995; 5:283–289.
- Engle P, Black MM, Cabral del Mello M, Gertler PJ, Kapiriri L, Martorell R, Group, International Child Development Steering. Strategies to avoid the loss of developmental potential in more than 200 million children in the developing world. *Lancet*. 2007; 369:229–242. [PubMed: 17240290]
- Fernald, LCH., Kariger, P., Engle, P., Raikes, A. *Examining early child development in low-income countries: A toolkit for the assessment of children in the first five years of life*. Washington, D.C.: The International Bank for Reconstruction and Development; 2009.
- Fernandes M, Stein A, Newton CR, Cheikh-Ismaïl L, Kihara M, Wulff K, Villar J. The INTERGROWTH-21st project neurodevelopment package: A novel method for the multi-dimensional assessment of neurodevelopment in preschool age children. *PLoS ONE*. 2014; 9:1–34.
- Fischer VJ, Morris J, Martines J. Developmental screening tools: Feasibility of use at Primary Healthcare level in low- and middle-income settings. *Journal of Health and Popular Nutrition*. 2014; 32:314–326.
- Fletcher JM. Memory for verbal and nonverbal stimuli in learning disability subgroups: Analysis by selective reminding. *Journal of Experimental Child Psychology*. 1985; 40:244–259. [PubMed: 4045379]
- Gioia, GA., Isquith, PK., Guy, K., Kenworthy, L. *Behavior Rating Inventory of Executive Function*. Lutz, FL: Psychological Assessment Resources, Inc; 2000.

- Giordani B, Boivin MJ, Opel B, Nseyila DND, Lauer RE. Use of the KABC with children in Zaire, Africa: An evaluation of the sequential-simultaneous processing distinction within an intercultural context. *International Journal of Disability, Development and Education*. 1996; 43:5–24.
- Gladstone M, Lancaster G, Umar E, Nyirenda M, Kayira E, Van den Broek N, Smyth RL. Perspectives of normal child development in rural Malawi—a qualitative analysis to create a more culturally appropriate developmental assessment tool. *Child care, health, and development*. 2009; 36:346–353.
- Gladstone M, Lancaster G, Umar E, Nyirenda M, Kayira E, Van den Broek N, Smyth RL. The Malawi Developmental Assessment Tool (MDAT): The creation, validation, and reliability of a tool to assess child development in rural African Settings. *PLoS Medicine*. 2010; 7:1–13.
- Goldman D, Shapiro E, Nelson CA. Measurement of vigilance in 2-year old children. *Developmental Neuropsychology*. 2004; 25:227–250. [PubMed: 15147998]
- Grantham-McGregor S, Cheung AM, Cueto S, Glewwe P, Richter L, Strupp B, Group, International Child Development Steering. Developmental potential in the first 5 years for children in developing countries. *Lancet*. 2007; 369:60–69. [PubMed: 17208643]
- Grantham-McGregor S, Cheung YB, Cueto S, Glewwe P, Richter L, Strupp B, International Child Development Steering, Group. Child Development in Developing Countries 1: Developmental potential in the first 5 years for children in developing countries. *The Lancet*. 2007; 369:60–70.
- Grantham-McGregor S, Cheung YB, Cueto S, Glewwe P, Richter L, Strupp B. Developmental potential in the first 5 years for children in developing countries. *The Lancet*. 2007; 369:60–70.
- Greenberg, LM., Kindschi, CL., Corman, CM. *Test of Variables of Attention: Clinical guide*. Los Alamitos, CA: Universal Attention Disorders; 2000.
- Hagie MU, Gallipo PL, Svien L. Traditional culture versus traditional assessment for American Indian students: An investigation of potential test item bias. *Assessment for Effective Intervention*. 2003; 29:15–25.
- Hamadani JD, Tofail F, Nermell B, Gardner R, Shiraji S, Bottai M, Vahter M. Critical windows of exposure for arsenic-associated impairment of cognitive function in pre-school girls and boys: A population-based cohort study. *International Journal of Epidemiology*. 2011; 40:1593–1604.
- Henderson, SE., Sugden, DA. *The Movement Assessment Battery for Children (Movement-ABC)*. New York: Psychological Corporation; 1992.
- Henderson, SE., Sugden, DA., Barnett, AL. *Movement assessment battery for children-2*. London, UK: The Psychological Corporation; 2007.
- Heydon, RM., Iannacci, L. *Early childhood curricula and the de-pathologizing of childhood*. Toronto: University of Toronto Press; 2008.
- Holding PA, Taylor HG, Kazungu SD, Mkala T, Gona J, Mwamuye B, Stevenson J. Assessing cognitive outcomes in a rural African population: Development of a neuropsychological battery in Kilifi District, Kenya. *Journal of the International Neuropsychological Society*. 2004; 10:246–260. [PubMed: 15012845]
- John CC, Bangirana P, Byarugaba J, Opoka RO, Idro R, Jurek AM, Boivin MJ. Cerebral malaria in children is associated with long-term cognitive impairment. *Pediatrics*. 2008; 122:e92–99. [PubMed: 18541616]
- Jordan CM, Johnson AL, Hughes SJ, Shapiro EG. The Color Object Association test (COAT): The development of a new measure of declarative memory for 18- to 36-month-old toddlers. *Child Neuropsychology*. 2008; 14:21–41. [PubMed: 18097800]
- Kammerer, B., Isquith, PK., Lundy, S. Approaches to assessment of very young children in Africa in the context of HIV. In: Boivin, MJ., Giordani, B., editors. *Neuropsychology of children in Africa: Perspectives on risk and resilience*. New York: Springer; 2013. p. 17-36.
- Kashala E, Elgen I, Sommerfelt K, Tylleskar T. Teacher ratings of mental health among school children in Kinshasha, Democratic Republic of Congo. *European Child and Adolescent Psychiatry*. 2005; 14:208–215. [PubMed: 15981132]
- King GA, Law M, King S, Hurley P, Hanna S, Kertoy M, Rosenbaum P. Measuring children's participation in recreation and leisure activities: Construct validation of the CAPE and PAC. *Child care, health, and development*. 2006; 33:28–39.

- King, GA., Law, M., King, S., Hurley, P., Hanna, S., Kertoy, M., Young, N. Children's Assessment of Participation and Enjoyment (CAPE) and Preferences for Activities of Children. San Antonio, TX: Harcourt Assessment, Inc.; 2004.
- Kirova A, Hening K. Culturally responsive assessment practices: Examples from an intercultural multilingual early learning program for newcomer children. *Power and Education*. 2013; 5:106–119.
- Kitsao-Wekulo PK, Holding PA, Taylor HG, Abubakar A, Connolly K. Neuropsychological testing in a rural African school-age population: Evaluating contributions to variability in test performance. *Assessment*. 2012; 20:776–784. [PubMed: 22936783]
- Kolling T, Lamm B, Vierhaus M, Knopf M, Lohaus A, Fassbender I, Keller H. Differential development of motor abilities in western middle-class and Cameroonian Nso infants. *Journal of Cross-Cultural Psychology*. 2014; 45:1502–1508.
- Lansdown RG, Goldstein H, Shah PM, Orley JH, Di G, Kaul KK, Reddy V. Culturally appropriate measures for monitoring child development at family and community level: A WHO collaborative study. *Bulletin of the World Health Organization*. 1996; 74:283–290. [PubMed: 8789927]
- Leung PWL, Kwong SL, Tang CP, Ho TP, Hung SF, Lee CC, Liu WS. Test-retest reliability and criterion validity of the Chinese version of the CBCL, TRF, and YSR. *Journal of Child Psychology & Psychiatry*. 2006; 47:970–973. [PubMed: 16930392]
- Luciana M, Nelson CA. Assessment of neuropsychological function through the use of the Cambridge Neuropsychological Testing Automated Battery: Performance in 4- to 12-year-old children. *Developmental Psychology*. 2002; 22:595–624.
- Luciana M, Sullivan J, Nelson CA. Individual differences in the phenylalanine to tyrosine ratios moderate cognitive outcomes in adolescents treated early and continuously for PKU. *Child Development*. 2001; 72:1637–1652. [PubMed: 11768137]
- Luiz DM, Foxcroft CD, Povey JL. The Griffiths Scales of Mental Development: A factorial validity study. *South African Journal of Psychology*. 2006; 36:192–214.
- Maida M, van de Vijver FJ, Srinivasan K, Transler C, Sukumar P. Traveling with cognitive tests: Testing the validity of a KABC II adaptation in India. *Assessment*. 2010; 17:107–115. [PubMed: 19745212]
- Malda M, Fons JR, van de Vijver FJR, Krishnamachari S, Transler C, Sukuman P. Traveling with cognitive tests: Testing the validity of a KABC II adaptation in India. *Assessment*. 2010; 17:107–115. [PubMed: 19745212]
- Manly T, Nimmo-Smith I, Watson P, Anderson V, Turner A, Robertson IH. The differential assessment of children's attention: The Test of Everyday Attention for Children (TEA-ch), normative sample and ADHD performance. *Journal of Child Psychology*. 2001; 42:1056–1081.
- Manly, T., Robertson, IH., Anderson, V., Nimmo-Smith, I. The Test of Everyday Attention for Children (TEA-ch). Bury St. Edmunds, England: Thames Valley Test; 1999.
- Mireku MO, Boivin MJ, Davidson LL, Ouedraogo JB, Koura GK, Alao MJ, Bodeau-Livinec F. Impact of helminth infestation during pregnancy on cognitive and motor functions of one-year-old children. *PLOS Neglected Tropical Diseases*. 2015; 9:E0003463.
- Molica CM, Maruff P, Vance A. Development of a statistical approach to classifying treatment response in individual children with ADHD. *Human Psychopharmacology*. 2004; 19:445–456. [PubMed: 15378673]
- Mullen, EM. Mullen Scales of Early Learning. Torrance, CA: WPS; 1995.
- Mung'ala-Odera V, Njuguna MR, Mturi N, Alcock K, Carter JA, Newton CRJC. Validity and reliability of the 'Ten Questions' questionnaire for detecting moderate to severe neurological impairment in children aged 6–9 years in rural Kenya. *Neuroepidemiology*. 2004; 23:67–72. [PubMed: 14739570]
- Mungala-Odera V, Newton CRJC. Identifying children with neurological impairment and disability in resource-poor countries. *Child care, health, and development*. 2007; 33:249–256.
- Munir SZ, Zaman S, McConachie H. Development of an Independent Behaviour Assessment Scale for Bangladesh. *Journal of Applied Research in Intellectual Disabilities*. 1999; 12:241–252.

- Murray CJ, Lopez AD. Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. *Lancet*. 1997; 349(9063):1436–1442. doi:S0140-6736(96)07495-8 [pii] 10.1016/S0140-6736(96)07495-8. [PubMed: 9164317]
- Nair MKC, George B, Philip E, Lekshmi MA, Haran JC, Sathy N. Trivadrurum developmental screening chart. *Indian Pediatric*. 1991; 28:869–872.
- Nampijja M, Apule B, Lule S, Akurut H, Muhangi L, Elliott AM, Alcock KJ. Adaptation of Western measures of cognition for assessing 5-year-old semi-urban Ugandan children. *British Journal of Educational Psychology*. 2010; 80:15–30. [PubMed: 19594989]
- Nokes C, McGarvey ST, Shiue L, Guanling W, Hawai W, Bundy AP, et al. Evidence for an improvement in cognitive function following treatment of *Schistosoma japonicum* infection in Chinese primary school children. *American Journal of Tropical Medicine and Hygiene*. 1999; 60:556–565. [PubMed: 10348228]
- Noll RB, Patel SK, Embry L, Hardy KK, Pelletier W, Annett RD, Barakat LP. Children’s oncology group’s 2013 blueprint for research: Behavioral Science. *Pediatric Blood Cancer*. 2013; 60:1048–1054. [PubMed: 23255478]
- Oakland, T. How universal are test development and use?. In: Grigorenko, EL., editor. *Multicultural psychoeducational assessment*. New York: Springer; 2007. p. 1-40.
- Ogunnaike OA, Houser RF Jr. Yoruba Toddlers’ engagement in errands and cognitive performance on the Yoruba Mental Subscale. *International Journal of Behavioral Development*. 2002; 26:145–153.
- Petrides M, Milner B. Deficits on subject-ordered tasks after frontal-and temporal-lobe lesions in man. *Neuropsychologia*. 1981; 20:249–262.
- Phatak AT, Khurana B. Baroda development screening test for infants. *Indian Pediatric*. 1991; 28:31–37.
- Pienaar AE, Barhorst R, Twisk JW. Relationships between academic performance, SES school type, and perceptual motor skills in first grade South African learners: NW-CHILD study. *Child Care Health and Development*. 2014; 40:370–378.
- Prado EL, Hartini S, Rahmawati A, Ismayani E, Hidayati A, Hikmah N, et al. Test selection, adaptation, and evaluation: A systematic approach to assess nutritional influences on child development in developing countries. *British Journal of Educational Psychology*. 2010; 80:31–53. [PubMed: 19772714]
- Raven, JC., Court, JH., Raven, J. *Manual for Raven’s Progressive Matrices and Vocabulary Scales: Section 2. The Coloured Progressive Matrices*. San Antonio, TX: Psychological Corporation; 1998.
- Robertson J, Hatton C, Emerson E, Yasamy MT. The identification of children with, or at significant risk of, intellectual disabilities in low- and middle-income countries: A review. *Journal of Applied Research in Intellectual Disabilities*. 2012; 25:99–118. [PubMed: 22473963]
- Roselli M, Ardila A. The impact of culture and education on nonverbal neuropsychological measurements: A critical review. *Brain and Cognition*. 2003; 52:326–333. [PubMed: 12907177]
- Ruffieux N, Njamshi AK, Mayer E, Sztajzel R, Eta SC, Doh RF, Hauert CA. Neuropsychology in Cameroon: First normative data for cognitive tests among school-aged children. *Child Neuropsychology*. 2010; 16:1–19. [PubMed: 19370432]
- Sakti H, Nokes C, Subagio WH, Hendratino S, Hall A, Bundy DA, et al. Evidence of an association between hookworm infection and cognitive function in Indonesian school-children. *Tropical Medicine and International Health*. 1999; 4:322–334. [PubMed: 10402967]
- Sattler, JM. *Assessment of children: Cognitive approaches*. San Diego: Author; 2001.
- Schatz P, Zillmer EA. Computer-based assessment of sports-related concussion. *Applied Neuropsychology*. 2003; 10:42–47. [PubMed: 12734074]
- Schutt, RK., Seidman, LJ., Keshavan, MS. *Social neuroscience: Brain, mind, and society*. Cambridge, MA: Harvard University Press; 2015.
- Serpell R, Jere-Folotiya J. Basic education for children with special needs in Zambia: Progress and challenges in the translation of policy into practice. *Psychology and Developing Societies*. 2011; 23:211–245.
- Shallice T. Specific impairments of planning. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*. 1982; 298:199–209. [PubMed: 6125971]

- Shapiro EG. Early Childhood Vigilance Test. 2010
- Soleimani F, Dadkhah A. Validity and reliability of Infant Neurological International Battery for detection of gross motor developmental delay in Iran. *Child: Care, Health, and Development*. 2006; 33:262–265.
- Sparrow, SS., Cicchetti, DV., Balla, DA. Vineland Adaptive Behavior Scales. 2. Minneapolis, MN: Pearson; 2005.
- Spinazzola J, Ford JD, Zucker M, van der Kolk BA, Silva S, Smith SS, Blaustein M. Survey evaluates complex trauma exposure, outcome, and intervention among children and adolescents. *Psychiatric Annals*. 2005; 35:433–439.
- Syvaoja HJ, Tammelin TH, Ahonen T, Rasanen P, Tovanen A, Kankaanpaa A, Kantomaa MT. Internal consistency and stability of the CANTAB neuropsychological test battery in children. *Psychological Assessment*. 2015; 27:698–709. [PubMed: 25528164]
- Tombokan-Runtukau J, Nitko AJ. Translation, cultural adjustment, and validation of a measure of adaptive behavior. *Research in Developmental Disabilities*. 1992; 13:481–501. [PubMed: 1410714]
- Toplak ME, Bucciarelli SM, Jain U, Tannock R. Executive functions: Performance-based measures of the Behavior Rating Inventory of Executive Function (BRIEF) in adolescents with Attention Deficit/Hyperactivity Disorder. *Child Neuropsychology*. 2008; 1:53–72.
- van Meerbeke AV, Talero-Gutierrez C, Gonzalez-Reyes R. Prevalence of delayed neurodevelopment in children from Bogota, Colombia. *Neuroepidemiology*. 2007; 29:74–77. [PubMed: 17925597]
- Van Widenfelt BM, Treffers PDA, De Beurs E, Siebelink BM, Koudijs E. Translation and cross-cultural adaptation of assessment instruments used in psychological research with children and families. *Clinical Child and Family Psychology Review*. 2005; 8(2):135. [PubMed: 15981582]
- Vazir S, Naidu AN, Vidyasagar P, Lansdown RG, Reddy V. Screening test battery for assessment of psychosocial development. *Indian Pediatric*. 1994; 31:1465–1475.
- Vierhaus M, Lohaus A, Kolling T, Teubert M, Keller H, Fassbender I, Schwarzer G. The development of 3- to 9-month-old infants in two cultural contexts: Bayley longitudinal results for Cameroonian and German infants. *European Journal of Developmental Psychology*. 2011; 8:349–366.
- Visu-Petra L, Benga O, Miclea M. Dimensions of attention and executive functioning in 5-to-12-years-old children: Neuropsychological assessment with the NEPSY battery. *Cognitie Creier Comportament*. 2007; 11:585–608.
- Vriezen ER, Pigott S. The relationship between parental report on the BRIEF and performance-based measures of executive function in children with moderate to severe traumatic brain injury. *Child Neuropsychology*. 2002; 8:296–303. [PubMed: 12759826]
- Walker SP, Wachs TD, Gardner JM, Lozoff B, Wasserman GA, Pollitt E, Group, International Child Development Steering. Child development: Risk factors for adverse outcomes in developing countries. *Lancet*. 2007; 369:145–157. [PubMed: 17223478]
- Wirz S, Edwards K, Flower J, Yousafzai A. Field testing of the ACCESS materials: A portfolio of materials to assist health workers to identify children with disabilities and offer simple advice to mothers. *International Journal of Rehabilitation Research*. 2005; 28:293–302. [PubMed: 16319554]
- World Bank. Development Data - Malawi. 2011. Retrieved from <http://data.worldbank.org/country/malawi>
- World Bank. Updated income classifications. 2015.
- World Health Organization. International classification of functioning, disability and health (ICF): Child and youth version. Geneva, Switzerland: Author; 2007.

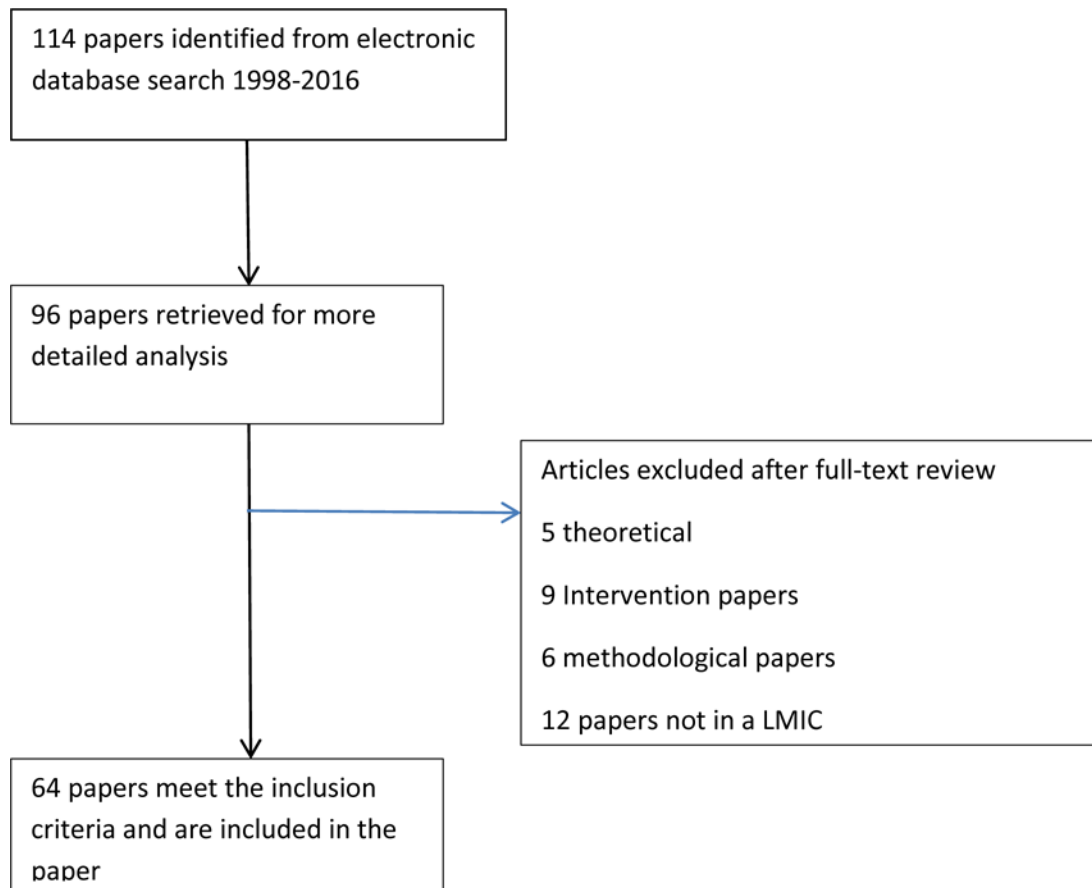


Figure 1.
Flow Diagram for Selection of Articles

Table 1

Summary of test-specific information, training required and key issues for developmental screening tests and comprehensive or specific neurodevelopmental tests assessed in LMIC

Measure	LMIC for which the test has been adapted	Test-specific information	Extensive training required?	Comments
Screening Measures				
Abbreviated Developmental Scale (van Meerbeke et al., 2007)	Colombia	Used as a primary tool in Colombia to evaluate developmental delays	No	No reliability and validity published
ACCESS Portfolio Wirz et al., 2005)	Uganda, Sri Lanka	Administered by public health officers	No	Reliability and validity not reported. Parent and medical officers reported the questionnaire was useful.
A not B test (Diamond, 1990)	Kenya	An (Tombokan-Runtukau & Nitko, 1992) object is hidden in front of the child in one of 2 locations; after short delay child is asked to find the item; location is changed between trials	No	Limited psychometric properties for LMIC countries—well-established findings in HIC
Ages and Stages Questionnaire - Third Edition (Squires, 2009)	Korea, Thailand, Ecuador	Screening instrument for developmental delays 4 to 60 months of age	No	No norms, validity and reliability are reasonable, parent questionnaire
Angkor Hospital for Children Developmental Milestone Assessment Tool (Ngoun, Stoev, Van't Ende, & Kumar, 2012)	Cambodia	Measure of fine and gross motor, and social emotional development	No	Adapted from Denver Developmental Test; added 43% of items based on expert opinion; reliability and validity being established
Australian Early Development Index (Sayers, 2004)	Indigenous children in Australia (rch.org.au/aedi)	Teacher completed	No	Used to collect data as to community functioning—does not provide individual scores
Baroda Development Screening Test (Phatak & Khurana, 1991)	India	Child Administered	Some	Adapted from the BSID; 54 items (22 motor; 32 mental) Designed to be used by health care workers going door to door
Denver Prescreening Developmental Questionnaire (Denver II PDQ) (Frankenburk et al., 1990)	Brazil, Arab countries not specified, Armenia, Malawi, Sri Lanka, Egypt, China, Philippines, Singapore, Turkey, United Arab Emirates, Zaire	Parent Questionnaire of developmental skills expected at certain ages	No	Children with more than 2 delays or 3 items that were not performed totally should be referred for a further assessment to determine needs
Developmental Assessment Tool for Anganwadis (Nair et al., 1991)	India	Child Administered	Minimal	Designed to test children 2 ½ years old for developmental delays—quick and inexpensive. Anganwadis are government sponsored childcare and mother care centers in India
Developmental Milestones Checklist (Prado et al., 2013)	Kenya, Burkina Faso, Cambodia	Evaluates motor, language and personal-social skills	No	Good internal and test-retest reliability; validity established by correlations with play

Measure	LMIC for which the test has been adapted	Test-specific information	Extensive training required?	Comments
				materials in home and activities
Disability Screening Test (Chopra et al., 1999)	India	Parental reports, observation of the child, direct assessment	Minimal	Obtains information on prenatal and birth history, physical and sensory functions, and direct developmental assessment
ICMR Psychosocial Development Screening Test (India) (Malik, Pradhan, & Prasuna, 2007)	India, Indonesia, Thailand	Screens for developmental delays	No	Administered by community health care workers Can be used for screening
Infant Neurological International Battery Test (Soleimani & Dadkhah, 2006)	Iran	Screens for developmental delays particularly in the areas of motor: does not test for cognitive delay	Designed for professionals with some level of medical training	Not designed for use by low-level workers for administration
Lucknow Development Screen for Indian Children (Bhave et al., 2010)	India	Administered to the main caregiver	Minimal	27 milestones in motor, language and social domains. Validated against the Developmental Assessment Scale
Monitoring Child Development (Lansdown et al., 1996)	China, India, and Thailand	Child administered	Minimal	Multicenter study in collaboration with WHO. Purpose was to identify a small number of key milestones to be placed in child's medical record
Screening Test Battery for Assessment of Psychosocial Development (Vazir et al., 1994)	India (particularly rural India)	Child administered	Yes	66 items in 5 areas: gross motor, vision and fine motor, hearing and language and concept development, self-help skills, and social skills. Found to be culturally appropriate by the users
Strengths and Difficulties Questionnaire (1997) http://www.sdqinfo.com/b1.html	Bangladesh China, Iran, Malawi, Brazil, Pakistan, Yemen, Democratic Republic of Congo	Brief behavioral screening questionnaire for ages 3 to 16 years	No	Free and in the public domain
Ten Questions Questionnaire (Durkin, Hasan, & Hasan, 1995)	Sub-Saharan Africa, Asia, Caribbean, India, Jamaica, Pakistan, Bangladesh	Used to detect severe neurological impairment in children in resource poor countries	No	Good sensitivity rates ranging from 87.4% for hearing, 70% for cognitive, and 100% for epilepsy with specificity rates at 96% or higher (Mung'ala-Odera et al., 2004)
Test de Desarrollo Psicomotor (TEPSI) (Haussler & Marchant, 1980)	Chile	Evaluates motor function, coordination and language	Minimal training needed	Tester observes the child doing selected tasks
Comprehensive Neurodevelopmental Measures				
Bayley Scales of Infant Development-III (Bayley, 2005)	Bosnia, Argentina, Bangladesh, Brazil, Chile, China, Costa Rica, Czech Republic, Guatemala, Indonesia, Jamaica, Kenya, Lithuania, Malaysia, Mexico, Nicaragua, Nigeria, Phillipines, Poland, Romania, Seychelles, South Africa,	Major measure of development including fine and gross motor, language, social skills, and reasoning.	Yes	Most research completed on BSID II; BSID III now used and does not have the language scale factored into the composite

Measure	LMIC for which the test has been adapted	Test-specific information	Extensive training required?	Comments
	Tanzania, Turkey, Thailand, Viet Nam, Zimbabwe			
Cambridge Neuropsychological Testing Automated Battery (CNTAB)	Used with children whose first language is German, Russian, Spanish, Middle-eastern, Indian, African, Ethiopian, Asian	Computerized measure of visual memory, reaction time, processing speed, working memory, planning	No	Some familiarity required for computer use.
Cogstate (Westerman, Darby, Maruff, & Collie, 2001)	Jamaica, Uganda, China, Hungary, India, South Africa and Lithuania.	Computerized measure of thinking, memory and reasoning with no language component	No	*“None of the CogState tasks are language dependent, so they are adaptable cross-culturally” (Boivin et al., 2010)
Escala Argentina de Inteligencia Sensorimotriz (Oiberman, Orellana, & Mansilla, 2006)	Argentina, Chile	Uses direct observation of children completing Piagetian tasks	Need experience in child development	No studies for the validity and reliability of this screening measure
Escala de Evaluacion del Desarrollo Psicomotor (de Andraca, P., de La Parra, & Rivera y Marcwla, 1998 Schonhaut, Rojas, & Kaempffer, 2005)	Argentina, Chile	Observation and report	Need experience in child development	Used to identify risk factors for developmental delay—found to be useful particularly with low-income children
Griffiths Mental Development Scales (Griffiths, 1984)	South Africa, Pakistan	Measure of overall ability; measures gross motor, activities of daily living, interaction with others, language, eye hand coordination and visuospatial skills as well as practical reasoning	Requires training and certification	Older version well validated—newer version published in 2006 is not as well studied
Grover-Counter Scale of Cognitive Development (South Africa) (Grover & Sebate, 2005)	South Africa	Based on Piagetian theory—designed to be used for children and adults as well as those with hearing impairments—mostly nonverbal	Yes	Only provisional norms are provided—obtained from sample of normal children, mentally handicapped children, and normal African-speaking children
IEA Preprimary Program Assessments (Multi-national) (jmontie@highscope.org)	Guinea, Cape Verde, Bangladesh, Indonesia, Thailand	Documents how a child is performing in a given setting through observation	Need extensive training,	Is observation system for children 4 to 7 years
Intergrowth 21st Neurodevelopment Package (Fernandes et al., 2014)	India, Italy, Brazil, Kenya	Cognition, language, motor, behavior, attention	Yes	New measure being used—validity to be established
Kaufman Assessment Battery for Children-II (Kaufman & Kaufman, 2004)	Sub-Saharan Africa, India, Uganda, Romania, Benin, Laos, China, Senegal, Zaire	Tests general cognitive ability, and skills in visual spatial reasoning, sequential thinking, planning, learning, and memory	Yes	It maintained its factor structure in Ugandan children.
Kilifi Developmental Inventory (Abubakar, Holding, Van Baar, Newton, & van de Vijver, 2008)	Kenya Nigeria	69 items that are explained and demonstrated prior to the child attempting the task	Yes	Found high community acceptability of the measure in Kenya
Leiter International Performance Scale-Revised (1997) (Roid & Miller, 1997)	Saudia Arabia, Taiwan	Nonverbal measure	Training is needed for administration	Language is not involved but is not a culture free measures

Measure	LMIC for which the test has been adapted	Test-specific information	Extensive training required?	Comments
MacArthur-Bates Communicative Development Inventory (1993) (Fenson et al., 2007)	Sign language adaptation Hungary	Completed by the parent—parent asked about whether child understands words from a list and then is asked to fill in a vocabulary production checklist	Some training required.	Norms are developed in Britain. Unclear how would be adapted in other countries.
Malawi Developmental Assessment Tool (Gladstone et al., 2010)	Malawi	Direct assessment of the child as well as observation	Some training required	Sensitivity found to be around 97% and specificity at 82%
Mullen Early Developmental Scales (Mullen, 1995)	South Africa, Sub-Saharan Africa	Evaluates language (expressive, receptive), fine and gross motor skills, and social ability	Yes	Community Mental Health workers can administer
NEPSY II (2007) (Korkman, Kirk, & Kirk, 2007)	Zambia, Indonesia, South Africa	Measure of executive functions, memory, as well as language and reasoning	Requires extensive training	Limited use in LMIC; unclear as to reliability and validity
Rapid Neurodevelopmental Assessment Tool (Kahn et al., 2010)	Bangladesh	8 age related forms from ages 0 to 24 months; Measures reflexes, motor, vision, hearing, speech, cognition, behavior, and seizures	Partially; need experience in child development	Reliability and validity indices found to be good to excellent; Correlated strongly with BSID III
Shoklo Neurological Test (Thailand) (Haataja et al., 2002)	Thailand	No norms, assesses motor, cognition, social-emotional behavior, and speech	Yes	Must be administered by health care worker Uses observation and parental report Does not provide age norms
Stanford Binet Intelligence Scale, 5th edition (2003) (Roid, 2003)	India, Japan, China, Taiwan, Thailand, Turkey	Child is directly evaluated on verbal reasoning, abstract visual reasoning, quantitative comprehension, and short-term memory	Yes	Previous version validated—new version has not been evaluated
Wechsler Preschool and Primary Scale of Intelligence-third edition and Wechsler Intelligence Scales for Children (WISC) (Wechsler, 2003)	Brazil, Chile, South Korea, Bangladesh, China, Colombia, Ecuador, India, Indonesia, Iran, Mexico, Peru, Thailand, Turkey, Taiwan, Venezuela, Pakistan, Yugoslavia,	Measures of IQ; verbal comprehension, perceptual reasoning, processing speed, working memory	Need extensive training	Reliability and validity well established in HIC but not in LMIC
Yoruba Mental Subscale (Ogunnaiké & Houser, 2002)	Yoruba, Nigeria (urban, semi-urban, and rural communities)	Based on the Bayley Scales from 1969	Requires extensive training	Out of date norms—needs to be rescaled for BSID III
Rating Scales				
Behavior Rating Inventory of Executive Function (Gioia, Isquith, Guy, & Kenworthy, 2000)	Malawi, Uganda	Caregiver completed—may need to be read to parent	No	Rating scale; Just beginning to be used in Sub-Saharan Africa—appropriateness is currently unknown
CBCL/1.5–5 (Achenbach Child Behavior Checklist) (Achenbach & McConaughy, 2003)	Ethiopia, Bulgaria, China, Taiwan, Hong Kong, Croatia, Czech Republic, Estonia, Iran, India, Hungary, Lithuania, Malaysia, Poland, Brazil, Romania, Serbia, Slovenia, Philippines, Tamil, Thailand, Turkey, Ukraine, Pakistan, Viet Nam	Parent questionnaire—can also use a teacher form. Provides information about internalizing, externalizing and social behaviors	No	*The CBCL is translated into over 80 languages/not all forms available in all languages. *There are over 7,000 published studies on the CBCL tests (for multiple ages) multiple forms)

Measure	LMIC for which the test has been adapted	Test-specific information	Extensive training required?	Comments
Early Childhood Care and Development Checklist (Armecín et al., 2007)	Philippines	Checklist of items presented to parent and to child to evaluate general development	Requires experience in child development	Skills are either present or not present; Also has a semi-structured parent interview
Molteno Scale Cited in (Laughton, 2010)	South Africa	Measures overall developmental skills	Yes	Screening tool; weakly correlated with Griffiths Mental Development Scales
Academic, Attention, Visual-Motor, and Motor Measures				
Beery Visual Motor Integration Test-6th edition (Beery & Buktenica, 1997)	Recommended for use but not well documented as to countries	Nonverbal measure	No	Found to correlate significantly with academic achievement in South Africa
Bracken Basic Concept Scale-Revised (Bracken, 2008)	China	Academic measure including knowledge of letters, numbers, shapes, and pre academic skills	No	Measure of readiness—directly administered to child
Bruininks-Osteresky Test of Motor Proficiency	South Africa, South America, Sub-saharan Africa	Measure of motor dexterity, fine motor skills, and gross motor (balance, running, gait, etc)	Yes	Measure of motor skills—directly administered to the child as well as a rating scale for caregivers
Color Object Association Test (Jordan, Johnson, Hughes, & Shapiro, 2008)	Sub-Saharan Africa	Mostly nonverbal measure—can use materials commonly used in country of study	Yes	*Very little information found on this instrument
Early Childhood Vigilance Test (Shapiro, 2010)	Sub-Sahara Africa	Child watches a video and examiner codes where the child is viewing	Administration does not require extensive training. Scoring of this measure does	No published studies at this time
Early Development Instrument (EDI) (Brinkman et al., 2007)	Jamaica	Administered as a questionnaire for teachers	No	Measures school readiness
Motor Assessment Battery for Children-2	Sub-Saharan Africa, Asia, Pacific Islanders	Administered directly to the child	Requires some training	Significant revision from initial battery
Parental Report Scales (kvalsvig@gmail.com)	Tanzania, Nepal	Assesses language and motor skills through parent questionnaire	Requires some training	Psychometric properties being established but as of yet unpublished.
Peabody Picture Vocabulary Test-4th edition (2010) (Dunn & Dunn, 2007)	China, Jamaica, South Africa, Translated and adapted in Peru, Viet Nam, India, Ethiopia, West Indies	Measure of receptive vocabulary and screening test for verbal ability	No	Some words do not translate well; Spanish version normed on small and high SES children
Test of Variables of Attention (Greenberg, 1989)	Malawi, Uganda	Computerized measure of attention. Provides scores for attention, impulsivity, reaction time, and variability	Little training needed for administration—training needed for interpretation	Child needs some initial exposure to computers
Woodcock-Johnson Third edition (Woodcock, McGrew, & Mather, 2001a, 2001b)	Costa Rica, Seychelles	Measure of cognitive and academic functioning	Need extensive training	Not validated in LMIC; Not clear how useful in many countries
Adaptive Behavior/Activities				
Children's Assessment of Participation and Enjoyment (Cape)/ Preferences for Activities of Children (PAC) (King, Law, King et al., 2004)	Beginning to be used in South Africa—no published data at this time	Self-report of a child on levels of participation and enjoyment in formal and informal activities	No	Construct validity established in Canada. Intensity enjoyment and preference were correlated with environmental, family and child variables

Measure	LMIC for which the test has been adapted	Test-specific information	Extensive training required?	Comments
Independent Behavior Assessment Scale (Munir, Zaman, & McConachie, 1999)	Bangladesh, Jamaica, Pakistan	Four scales: motor, socialization, communication and daily living skills	No	Test-retest reliability and inter-rater reliability ranged from .71 to .98 and judged to be excellent
Indian Adaptation of the Vineland Adaptive Behavior Scales (Mahli & Singi, 2002)	India	Same scales as Vineland	No	Authors indicate needs additional validation. Initial findings were of similar structure to the VABS
Indonesian Adaptation of the Vineland Adaptive Behavior Scales (Tombokan-Runtukahu & Nitko, 1992)	Indonesia	Same scales as Vineland—254 questions	No	Authors indicate needs more validation before it is used widely in Indonesia
Vineland Adaptive Behavior Scales second edition (Sparrow, Cicchetti, & Balla, 2005)	Uganda, Kenya; Indonesia, also translated into Spanish	Parent Interview	Need training	Generally administered by community health workers

Tombokan-Runtukahu, J., & Nitko, A. J. (1992). Translation, cultural adjustment, and validation of a measure of adaptive behavior. *Research in Developmental Disabilities, 13*, 481–501.

Table 2

Screening Measures

Test	Age	Adminis- tration	Procedure			Score Type						Domain					Reliability	Validity
			Child Administration	Rating Scale /Questionnaire	Interview	Raw Score	Standardized Score	Age Equivalence	Developmental Range	Percentile	Cognitive/ Reasoning	Language	Motor	Visual-Spatial	Learning/ Memory	Socio-Emotional/ Beh		
Abbreviated Developmental Scale	Less than 5 years	Less than 20 minutes															Not reported	Not reported
ACCESS Portfolio	0-36 months	Less than 20 minutes															Not reported	Congruent with expert assessment 76-82%
Ages and Stages Questionnaire	1 mos to 5 1/2 years	10-15 minutes															Test-retest $r = .91$ Inter-rater = $.92$	Sensitivity (.86) Specificity (.85)
Baroda Development Screening Test	0 to 30 months	Less than 20 minutes															Not reported	Sensitivity 95% ; Specificity 65%
Cambodian Developmental Assessment Test	3 to 5 years	30 minutes															None reported	None reported
Early Childhood Care and Development Checklist (Philippines)	0 to 5 years	45 min															None reported	None reported
Denver Prescreening Developmental Questionnaire	0-6 years	20 minutes															Inter-rater and test-retest $\geq .90$	Sensitivity = .66 Specificity = .88
Developmental Assessment Tool for Anganwadis	2.5 years	10 minutes															Not available	Not available
Disability Screening Test	0 to 6 years	5 minutes															Not available	Sensitivity=89% Specificity = 98%
Early Development Instrument (Canada)	4 to 5 years	7 to 20 minutes															Test retest = .82 to .94 Inter-rater = .56 to .80	Correlation with parent interview ranged from .15 to .48 Predictive validity ranged from .19 to .46
Escala Argentina de Inteligencia Sensoriomotriz (EAIS) (Argentina)	6 to 24 month																None reported	None reported
Escala de Evaluación del Desarrollo Psicomotor (EEDP) (Chile)	0 to 24 month	20 min															None reported	None reported
Guide for Monitoring Child Development (Turkey)	0 to 24 months	20 minutes															Internal consistency .83-.88	Concurrent validity 91.1 with Bayley II and neurological. Sensitivity 88% Specificity 93%
ICMR Psychosocial Developmental Screening Test (India)	0 to 6 years	20 min															None reported	None reported
Infant Neurological International Battery Test	4-18 months	10 minutes															Inter-rater .9 Test retest .86 to .99	Sensitivity .9 Specificity .83
Kilifi Developmental Checklist (Kenya)	1 to 5	10 to 15 minutes															Internal consistency = .96 Inter-observer agreement = .98 Test-retest = .96	Reported high face validity: sensitivity demonstrated in children with disorders
Locknow Developmental Screen for Indian Children	6 months to 2 years	10 minutes															Not reported	Sensitivity 95.9% Specificity 73%
Monitoring Child Development	0-6 years																Inter-rater .90-.96	Not reported

Test	Age	Administration	Procedure			Score Type					Domain					Reliability	Validity	
			Child Administration	Rating Scale /Questionnaire	Interview	Raw Score	Standardized Score	Age Equivalence	Developmental Range	Percentile	Cognitive/ Reasoning	Language	Motor	Visual-Spatial	Learning/ Memory			Socio-Emotional/ Belt
Molteno Scale	12 to 48 month	10 to 15 minutes															None Reported	
Parental Report Scales (Tanzania, Nepal)	6 to 59 month	15 to 20 minutes															None Reported	
Screening Test Battery for Assessment of Psychosocial Development	0 to 6 years	Less than 30 minutes															Test retest reliability ranged from 95 to 99%	None reported
Shoklo Developmental Test (Thailand)	3 to 12 month	20 to 25 minutes															None Reported	
Shoklo Neurological Test (Thailand)	9 to 36 month	15 to 20 minutes															None Reported	
Strengths and Differences	3 to 16 years	15 minutes																
Ten Questions Questionnaire	2-9 years	10 to 15 minutes																Sensitivity = 84% for cognitive and motor
Test de Desarrollo Psicomotor (TEPSI) (Chile)	2 to 5 years	30 to 40 minutes															None Reported	

/References are provided in the manuscript text

Table 3

Characteristics of major neurodevelopmental tests used in low- and middle-income countries.

Test	Age	Procedure				Score Type				Domains							Validity	
		Child Administration	Rating Scale/ Questionnaire	Observation	Raw Score	Standardized Score	Age Equivalence	Developmental Range	Percentile	Cognitive/ Reasoning	Language	Attention	EF	Motor	Visual-Spatial	Learning/ Memory		Reliability
Cognitive Development																		
Australian Early Development Index	4 to 5 years																None Reported	Construct Validity=.62, .69, .72 with teacher measures ¹
Angkor Hospital for Children Developmental Milestone Assessment Tool	3 to 5 years																None Reported	None Reported
IEA Preprimary Program Assessment (Multinational)	4 to 7 yrs																None Reported	None Reported
Malawi Developmental Assessment Tool	0 to 6 yrs																Inter-observer=.79 Test-retest=.99	Identified 97% of children as having a documented neurodisability with 18% of normal scoring poorly
Molteno Scale	12 to 48 mos																None Reported	None Reported
Bayley Scales of Infant Development- Third Edition	1 to 42 months																Internal consistency=.89 Test-retest=.85	Concurrent Validity=.79 with McCarthy scales .73 with WPPSI
CANTAB	6 to adult—mostly used with adults																None Reported	None Reported
CogState	6 to 106 years																Test-retest =.32 to .57	3 factor structure found; processing speed, processing accuracy, and mazes
Griffiths Mental Development Scales	0 to 8 years																Reported to be good but statistics not provided	Concern expressed about cultural bias based on factor analysis
Grower-Counter Scale of Cognitive Development-Revised (South Africa)	3 to 10 for																Not Reported	Not Reported
Intergrowth 21 st Assessment Package	14 weeks to 2 years																Inter-rater=.95% Test-retest=.95%	None Reported
Kaufman Assessment Battery for Children, KABC-II	3 to 18 years																Internal reliability=.70 to .96	Being developed in Malawi and Uganda
Kilifi Developmental Inventory	6 months to 35 months																Retest reliability = .87 for locomotor, .95 for eye-hand; .96 for psychomotor	Correlated positively with the Association for the Physically Disabled of Kenya observation schedule at .89
Letter-International Performance Scale-Revised, Letter-R	2 to 20 years																Not Reported	Not Reported
Mullen Scales of Early Learning	Birth to 5 years, 8 months																Test-retest=.78 to .99	Not Reported
Rapid Neurodevelopmental Assessment Instrument	25 mos to 60 mos																Inter-rater reliability=.63 to .99	Discriminated between children with impairment and those without 83%

Test	Age	Procedure			Score Type				Domains						Reliability	Validity	
		Child Administration	Rating Scale/ Questionnaire	Observation	Raw Score	Standardized Score	Age Equivalence	Developmental Range	Percentile	Cognitive/ Reasoning	Language	Attention	EF	Motor			Visual-Spatial
Stanford-Binet Intelligence Scales, 5th Edition, SB-5	2 to 85+ years															Not directly reported	Found to correlate with stunted growth and school attainment
Wechsler Preschool and Primary Scale of Intelligence Test - Third Edition, WPPSI-III	2 years, 6 months to 7 years, 3 months															Not directly reported	Has been found to be related to the child's nutritional status (Hamdam et al., 2011)
Yoruba Mental Subscale (Nigeria)	0 to 26 mos		?														
Attention																	
Early Childhood Vigilance Test, ECVT	12 to 46 months															None Reported	None reported
Conners Continuous Performance Test II	5 years to adult															None Reported	Loaded on vigilance factor .737
Test of Variables of Attention	5 years to adult															None Reported	Not Reported
Executive Functioning																	
Behavior Rating Inventory of Executive Function - Preschool Version	2 to 5 years, 11 months															Not Reported	Not Reported
NEPSY-II	3 to 16 years, 11 months															Reliability ranging from .88 to .28 (knock tap)	Internal consistency reported .88 to .94
Color Object Association Test, COAT	18 to 36 months old															Reported as good without specifying co-efficient	None Reported
Picture Search (from TEEACH)																Test-retest reliability = .76	Internal consistency = .76
Behavior																	
Child Behavior Checklist, CBCL	Preschool; 1 year to 5 months to 18 years															Reliability ranged from .64 (internalizing) to .78 (externalizing) to .83 (total)	None Reported
Adaptive Behavior																	
Independent Behavior Assessment Scale (Bangladesh, Jamaica, Pakistan)	2 to 9 years															None Reported	None Reported
Vineland Adaptive Behavior Scales, Second Edition	Birth to 90 years															None Reported	None Reported
Indonesia Adaptation of the VABS	6-18 years															Inter-rater reliability .4 for 82% of items	Face validity established by expert assessment
Gross and Fine Motor																	
Bruninks-Oseretsky Test of Motor Proficiency-2	4 to 21 years															Reliability not reported in LMIC	Good validity with school performance
Motor Assessment Battery for Children-2	3 to 16 years															Reported in Chinese sample (test retest from .62 to .92) and ICC .79 (intra-rater and .79 (inter-rater)	Content validity established by expert panel No construct validity reported Some support of differences culturally in Asian Pacific countries (Livesey, Coleman, & Pick, 2007)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Test	Age	Procedure			Score Type				Domains						Reliability	Validity		
		Child Administration	Rating Scale/ Questionnaire	Observation	Raw Score	Standardized Score	Age Equivalence	Developmental Range	Percentile	Cognitive/ Reasoning	Language	Attention	EF	Motor			Visual-Spatial	Learning/ Memory
Purdue Pegboard	2 to 16 years																Limited information as to reliability in LMC—good reliability in HIC	Correlated significantly with academic achievement