



OPEN ACCESS

Child development assessment tools in low-income and middle-income countries: how can we use them more appropriately?

Saraswathy Sabanathan,^{1,2} Bridget Wills,^{1,2} Melissa Gladstone³

¹Oxford University Clinical Research Unit, Wellcome Trust Major Overseas Programme, Hospital for Tropical Diseases, Ho Chi Minh City, Viet Nam

²Nuffield Department of Medicine, Centre for Tropical Medicine, University of Oxford, Oxford, UK

³Department of Women and Children's Health, Institute of Translational Medicine, University of Liverpool, Alder Hey Children's NHS Foundation Trust, Liverpool, UK

Correspondence to

Dr Saraswathy Sabanathan, Oxford University Clinical Research Unit, Wellcome Trust Major Overseas Programme, Hospital for Tropical Diseases, 764 Vo Van Kiet, District 5, Ho Chi Minh City, Viet Nam; swhitehorn@oucru.org

Received 30 December 2014

Revised 24 February 2015

Accepted 25 February 2015

Published Online First

30 March 2015

ABSTRACT

Global emphasis has shifted beyond reducing child survival rates to improving health and developmental trajectories in childhood. Optimum early childhood experience is believed to allow children to benefit fully from educational opportunities resulting in improved human capital. Investment in early childhood initiatives in low-income and middle-income countries (LMICs) is increasing. These initiatives use early childhood developmental assessment tools (CDATs) as outcome measures. CDATs are also key measures in the evaluation of programmatic health initiatives in LMICs, influencing public health policy. Interpretation of CDAT outcomes requires understanding of their structure and psychometric properties. This article reviews the structure and main methods of CDAT development with specific considerations when applied in LMICs.

INTRODUCTION

It is estimated that up to 200 million children fail to reach their developmental potential in low-income and middle-income countries (LMICs).¹ This statistic underpins recent calls for urgent global action to improve early child development in the first 5 years of life.² Poverty, malnutrition, recurrent and/or chronic infectious diseases plus inadequate cognitive stimulation, can all have life-long consequences on development.^{3–5} Economic modelling studies demonstrate how early interventions and investment in disadvantaged children are effective in improving child development and cost-effective in improving the 'human capital' potential of individuals.⁶ Investment in early child development initiatives in LMICs is burgeoning and with this, prioritisation of global health research into interventions that have maximum impact on child development. Specific child development assessment tools (CDATs) are recommended by major research funding bodies as appropriate tools to evaluate child development interventions, despite limited evidence of the nature of the CDATs' relationship with quality of life measures and later potential human capital.^{7, 8} Since no CDAT is suitable for all populations, a variety of tools have been developed worldwide.⁹ Typically, knowledge of the structure, psychometric properties, applicability and limitations of CDAT is restricted to child development specialists. In practice, health workers from various backgrounds often administer and interpret these tests. This article describes the structure, main methods of CDAT development and their application, focusing on children less than 5 years of age in LMICs.

WHAT IS A CHILD DEVELOPMENTAL ASSESSMENT TOOL, WHAT DO THEY MEASURE AND HOW ARE THEY ADMINISTERED?

Early childhood is a period of rapid physical growth during which an individual acquires a complex set of skills and functional competencies that should facilitate achievement of their potential in life. Weight and height charts were developed on the premise that children from similar ethnic backgrounds follow similar growth trajectories. In a comparable fashion, acquisition of skills during early childhood is also expected to follow a set trajectory. For growth and development we know that it is best to intervene promptly in order to maintain a healthy trajectory.^{10–12} US legislative changes incorporated in the Individuals with Disabilities Education Act 2004 specify the use of standardised CDATs to evaluate the following developmental domains; cognitive, language, motor, adaptive and socioemotional, in order to guide diagnostic investigations and interventions (table 1).^{13, 14}

Each domain of development is assessed by the child's ability to carry out a series of increasingly complex activities that are thought to reflect the expected level of development at a particular age. The developmental domains evaluated by CDAT do not function as discrete entities but influence each other—for example, to successfully complete a jigsaw puzzle requires a combination of cognitive and fine motor skills, and sufficient understanding of language to process instructions.

Methods to assess child development may include: (A) direct assessment of standardised activities by a trained assessor in a clinical environment, (B) verbal reporting/completion of a questionnaire about the child's abilities by parents or teachers and (C) unstructured observation by a trained assessor in an environment familiar to the child (eg, home/school). Direct testing in an unfamiliar environment by an unfamiliar adult may restrict a child's engagement and participation in the assessment. Parental reporting, on the other hand, may be affected by recall bias.⁸ Unstructured observation can be difficult to reproduce and interpret. An ideal CDAT would include aspects of all three methods, but with limited financial resources, often one method is chosen.⁸

TYPES OF ASSESSMENT TOOLS—SCREENING VERSUS FORMAL ASSESSMENT?

Screening tools are administered quickly, using a limited sample of items representing a domain and rely on predetermined cut-off points. Screening tools are designed to identify children who may



Open Access
Scan to access more
free content



CrossMark

To cite: Sabanathan S, Wills B, Gladstone M. *Arch Dis Child* 2015;**100**:482–488.

Table 1 Developmental domains

Developmental domains and subdomains	Description of domain
Cognitive ^{39 40}	Strategies and processes children develop to interpret and respond to their environment and experiences including; memory (ability to encode, retain and recall information over time) attention (the ability to choose what to focus on for a sustained period), that influence memory language skills which as the brain develops children acquire and refine language skills. Newborn explores the world by mouthing objects; and later explores the world by imitating actions, manipulating objects and planning two-step strategies to get what he wants. From 2 years, children increase their use of language and start make-believe play. In children aged 3–5 years there is rapid development in information processing (the speed and fluency of response following stimuli), cognitive flexibility (the ability to make and change strategies as required, and to simultaneously process multiple stimuli) and goal setting (the ability to plan strategies in a coherent and efficient order).
Language ⁴¹	
Receptive	Understanding of the spoken word and sentence structure
Expressive	Spoken vocabulary
Motor ⁴²	
Fine motor	Ability to manipulate small objects
Gross motor	Ability to walk, run and coordinate complex physical activities
Social and emotional ^{42 43}	The ability to identify and understand one's own feelings and to accurately read and comprehend emotional states in others. Ability to regulate one's own behaviour, to develop empathy for others, and to establish and maintain relationships
Adaptive behaviour ⁴⁴	Collection of conceptual, social and practical skills that have been learned by people in order to function in their everyday lives

have impairment and require a comprehensive assessment. However, screening tools have poor utility in assessing subtle delays that may have a significant impact on subsequent development. Screening tools are beneficial when used within the context of a development surveillance programme, where there are appropriate norms and known applicability in specific subgroups.¹⁵

Most formal tools require specialised specific training of assessors. They allow administration of items above or below the age bracket of the child thus allowing the child's strengths and weaknesses to be categorised more accurately. Formal assessments can be administered repeatedly to evaluate changes over time and/or following an intervention.

IS THE TOOL MEASURING WHAT IT IS SUPPOSED TO—VALIDITY AND RELIABILITY?

Developmental domains are a theoretical concept, typically referred to as 'constructs' in the psychometric literature. These cannot be directly measured but are inferred through the child's performance on a number of observed variables (test items). Reliability is the variability of scores obtained by an individual if repeatedly given the same test. There are widely cited levels of acceptable reliability for testing,¹⁶ however the result should be interpreted in context. For example, retest reliability can be influenced by developmental maturation over the time interval, or training by the caregiver following mistakes seen in the first assessment.¹⁷ Validity is the accuracy of the score representing the construct of interest.⁸ In addition, screening tools are evaluated for sensitivity, specificity, positive predictive value and negative predictive value. Reliability and validity of CDAT can be evaluated through several methods, summarised in table 2.

HOW TO DETERMINE A CHILD'S PERFORMANCE ON THE CDAT—THE REFERENCE POPULATION

All assessment tools have a threshold that a child must achieve for the assessor to be confident that there are no current developmental concerns. There are two main ways of setting this threshold, norm referencing and criterion referencing. A norm-referenced test is usually administered in a standardised manner and the individual child's scores in each domain are compared

with scores from a large representative sample of children of the same age and sex (normative data). Standard scores allow comparison between scales evaluating the same domain and monitoring of individuals at different ages. By contrast, criterion-referenced tests assess whether a child has acquired a particular skill by a certain age, according to a specific curriculum. Administration may not be standardised, since individual children sometimes require additional instructions or physical aids to complete a particular task but are then able to complete the specified tests. Criterion-referenced tests are often used in screening procedures, for example assessment of reading skills at school entry, or evaluation of an intervention such as a programme of physical therapy. In some cases criterion-based tests have been standardised on large representative populations.

There are concerns that the rapidly changing nature of society means that normative data become outdated very quickly. To counter this, inclusion of a control group is now considered crucial for research studies, even when the normative population is contemporaneous and comes from the same linguistic and cultural background as the study population.

WHEN TO ASSESS DEVELOPMENT, WHAT DOES IT MEAN FOR THE FUTURE—PREDICTIVE VALIDITY?

CDAT will only give a snapshot of the child at that time point. Traits that have not yet evolved in early childhood clearly cannot be assessed until such time as they might reasonably be expected to be present, and yet impairment within these characteristics may impact subsequent development significantly.¹⁸

After an assessment, it is important to know the predictive validity of the CDAT score, whether normal or abnormal. Correlations between early childhood performance and school abilities in normal children are variable.¹⁹ Children with moderate developmental delay, despite the same initial difficulties, have great variability in developmental trajectories, with some children catching up while others exhibit evolving difficulties with age.²⁰ Aspects of early development may be predictive of later academic ability. Preschool English language acquisition of vocabulary, grammar and descriptive phrasing predicted school reading ability in a prospective longitudinal US study.^{20a} Research in normal infants suggests that acquisition of efficient

Table 2 Basic psychometric properties used to evaluate child development assessment tools (CDATs)

Relevance/Importance	Comment	
Reliability		
Internal consistency	Evaluates the similarity of test items assessed in one domain. One measure is split-half reliability, which compares the scores on two halves of a test in a single domain.	High internal consistency suggests that some items are too similar, so no additional information is gained from assessing them. Low internal consistency suggests the items may not be assessing the same domain.
Interobserver	Evaluates variability between different assessors on the same subject	There may be systematic errors, specific to a particular group of assessors, and this parameter may not be generalisable when the tool is used by a different group of assessors.
Intraobserver	Evaluates variability within a single assessor on a single subject	Commonly evaluated by the same assessor scoring video recordings of their own assessments. This is not essential unless there is low interobserver reliability
Validity		
Test-retest	Evaluates variability within the subject (influenced by random factors such as familiarity with items and mood)	Difficult to interpret in early childhood when changes in development occur over a short time. Usually the repeat assessment should be carried out within 2 weeks of the first test.
Content	Experts in the field make consensus agreement on whether the individual item and the range of items adequately sample and represent the domain of interest.	Subjective measure that cannot be used in isolation to evaluate validity.
Criterion	Ideally assessed by comparison to an established 'gold standard' test assessing the same construct	Usually 'gold standard' tests are not available so the comparison is typically against another recognised test regularly used in the same population and thought to measure the same domain.
Discriminant/convergent	Evaluates expected positive and negative correlations between scores in different domains or between different tests of the same or differing underlying construct.	Scores from two independent tests (eg, one using report method the other a direct test) of one domain should correlate where neither test is considered a 'gold-standard'. To ensure the test is not overlapping with constructs not of interest, the scores evaluating different constructs should poorly correlate, for example, test scores on 'fine motor' should correlate poorly with 'social emotional'.
Construct	Statistical evaluation to see whether values of observed data fit a theoretical model of the constructs (confirmatory) or to explore a possible model of the 'underlying traits' being measured.	Large numbers of assessments are required to evaluate this.

inhibitory control and planning aspects of cognitive executive function at 4 years of age are associated with improved acquisition of school mathematical and literacy abilities at 6–7 years of age.^{21 22}

There is limited research examining the predictive nature between CDAT and longer-term outcomes in LMICs. Three longitudinal nutritional intervention studies collected CDAT scores in infants from Indonesia and Guatemala.²³ Scores on 2 monthly assessments correlated poorly in infants less than 24 months, but score correlations increased on assessments carried out between 24 months and 30 months of age. The scores at age 20 months or younger had no predictive value for verbal reasoning and arithmetic scores at school age. Overall, predictive power in comparative US research was similar.²⁴ There is still a lack of evidence as to whether the interpretation of group differences in scores due to an intervention or exposure in infancy should be interpreted as predicting anything in terms of later intelligence.²³

APPROACHES AND CHALLENGES OF CDAT DEVELOPMENT IN LMICs

Currently, only a few CDATs are available in LMICs and most are used in research settings. These CDATs typically follow one of the following formats: (A) a standard western CDAT with no adaptations; (B) a western CDAT translated (linguistic equivalence) and/or adapted for the local cultural environment (cultural equivalence); (C) an amalgamation of a number of translated and/or adapted items from several different western CDATs; or (D) a locally developed, culturally specific CDAT consisting of original items designed to be relevant to the population of interest.²⁵ The format of CDAT adaptation depends on

the aim of the application. Rie *et al* evaluated the impact of HIV on neurodevelopment of children in Kinchasa, Congo. The researchers deliberately used a direct translation of a western CDAT without cultural adaptation to compare between groups in the same setting and then compare between past and future studies in other settings.²⁶ This method does not ensure that the tool is measuring the same underlying ability in all settings. Holding *et al*²⁷ aimed to facilitate comparisons between different cultural groups in differing countries and needed to measure the same underlying ability in all settings. This was done by cultural modification of a western tool. The Kilifi Developmental Checklist (KDC) was also based on items from a range of CDATs.²⁸ The items were chosen based on ease of observing item success, how well the item could differentiate within the population of interest and if the item could be readily described in the local languages. Additional considerations pertaining to LMICs include the prohibitive cost of the license fee to translate or adapt a western CDAT, the level of skill of the assessor and location of assessment. Fernandes *et al*²⁹ developed a screening tool for use in high-income countries and LMICs. The team faced the challenge of developing a culture-free tool that was low in cost, could be administered after limited level of assessor training and completed in 30 min. To achieve this, an advisory panel listed project-specific criteria that the CDAT needed to fulfil.

Once items are selected, the first step is pilot studies on a small number of children. These aim to discover the range of responses elicited from the tested population and identify problems with unfamiliar stimulus materials. Prado *et al*¹⁷ adapted a cognitive test for Indonesian children, who had difficulty identifying a picture of a bunny, so replaced it with a picture of a

Table 3 Examples of CDATs (children <5 years) used in LMIC countries sourced from published literature

Tool and country	Type of tool and item selection (translation±adaptation) Method of assessment Age	Reliability Internal consistency Interobserver Test-retest	Validity Criterion Discriminant Construct	Comments
Kilifi Developmental Checklist (KDC) Kenya ²⁸	Mixture of items from government local screening test, Western tools and LMIC tools developed elsewhere Direct and Report 11–109 months	Excellent* Excellent† Excellent‡	Items range from poor to good correlation with parental report Done	Aimed to be used by low-skilled workers. Identified that secondary school level education was minimal for an assessor.
Kilifi Developmental Inventory (KDI) Kenya ³⁰	Assessment based on KDC with additional items from western tools Direct observation 6–35 months	Excellent* Excellent‡ Excellent‡	Good correlation with maternal reports† Sensitive to neurodevelopmental disorders and underweight children.	Designed to monitor changes over time. Involved parental focus groups so procedures acceptable to community.
Developmental Milestone Checklist Kenya ⁴⁵	Screening Items from Western tools Report 3–24 months	Good* Good‡	Very good correlation with KDIt. ³⁰ Stunted children had lower scores. Done	Structured interview. Future plans to develop cut-offs using clinical samples. Can be administered by those with little training.
Guide to monitoring child development. Turkey ⁴⁶	Screening Core ideas were adapted from three seminal western models of child development. Report 0–24 months	Excellent* Good§	Excellent agreement with comprehensive evaluation§ Sensitivity 0.88; 95% CI 0.69 to 0.96 Specificity 0.93; 95% CI 0.83 to 0.97	Brief, open-ended precoded interview with parent. No specific cognitive domain questions. May not be appropriate for monitoring 'at-risk' populations.
Rapid neurodevelopmental assessment RNA 0–2 years Bangladesh ⁴⁷	Screening Mixture of neurological examination and neurodevelopment from western tools Direct 0–24 months	Excellent§	Significant differences in scores in children with neurodevelopment impairment Detected difference in scores between urban rural children in older age groups	Used in a small high-risk population and identified infants requiring intervention for retinopathy of prematurity. RNA identified additional problems such as impairments in vision and hearing.
Rapid neurodevelopmental assessment. RNA 2–5 years Bangladesh ⁴⁸	Screening Mixture of neurological exam and development assessment from western tools Direct 24–60 months	Excellent§	Sensitivity 80–90% Specificity 60–78% Stunted children had lower scores.	Comparison with 'gold-standard' tests already developed for Bangladesh.
Malawi Developmental Assessment Tool. Malawi ³²	Screening Western developmental tool and new items locally devised Direct and report 0–72 months	Excellent ⁴⁹ 60–80% of items at least fair§ 99–100% scoring at least fair§	Sensitivity 97% Specificity 82% Sensitive to malnutrition except social developmental domain Good predictive validity ^{50 51}	Good strategy to devise socioemotional items.

Continued

Table 3 Continued

Tool and country	Type of tool and item selection (translation±adaptation) Method of assessment Age	Reliability Internal consistency Interobserver Test-retest	Validity Criterion Discriminant Construct	Comments
INTERGROWTH-21st Neurodevelopment Assessment ²⁹	Screening Items from selection of Western and LMIC tools			Project specific criteria to create a tool for use in children from middle-class and upper class families across low-income, middle-income and high-income settings, carried out by non-specialists. Also evaluates vision, hearing and sleep-wake cycle.
Systematic approach to assess nutritional influences. Indonesia ¹⁷	Direct and report 24 months Assessment Items from selection of Western tools Direct 22–55 months	Good Excellent Items at least fair* At least fair agreement† Items at least fair†	Expected relationship with maternal depression and maternal education	Clear hypothesis driven tool creation. May not be suitable as a general developmental assessment.

Interpretation of statistical tests are listed below.^{16 52} These need to be interpreted in context of application and population.^{17 53}

Levels of κ Levels δ , Rt, intraclass \pm correlation and α coefficients*.	Level of proposed agreement‡ (%)	Levels of clinical or practical significance
<0.4	<70%	Poor
0.4–0.59	70–79	Fair
0.60–0.74	80–89	Good
0.75–1.00	90–100	Excellent

CDAT, child development assessment tool; LMICs, low-income and middle-income countries.

chicken. Despite modification to the KDC following pilot studies, audit of final study scores found some items could not be successfully completed by at least one child less than 36 months of age.³⁰ These data were used to refine and then expand the KDC to create the Kilifi Developmental Inventory.³⁰

In some situations, no appropriate tool exists and new test items need to be devised. This requires understanding of the domain being measured and the cultural references of the relevant population, but can still be problematic. The most common method of developing new items is to engage participant representatives of the population in focus groups. Decisions regarding who will participate and how the information will be transcribed and analysed are important factors in the process. Local experts evaluate the suitability of particular items for a target group, and evaluate whether items adequately cover the domain under investigation. Gladstone *et al*³¹ adapted a tool for Malawi but needed to devise new items for the social/emotional domain. Themes for new items were devised from focus groups, and despite good face and content validity, pilot testing results identified that the new items did not perform well when compared in terms of their psychometric properties alongside original items. An iterative process of adaptation and testing led to two more draft versions of the Malawi Developmental Assessment Tool before the team successfully developed the final version.³²

Table 3 presents examples of the basic structure and psychometric properties of several CDATs developed for LMICs.

WHAT HAVE WE LEARNT FROM CDATS IN LMICS: FACTORS THAT INFLUENCE EARLY CHILD DEVELOPMENT IN LMICS

Debate on the influence of nurture versus nature continues, although an increasing body of opinion now considers the early perinatal environment to be as important in determining cognitive ability as an individual's genetic background. Intrauterine environment, nutritional deficiency in infancy, infectious diseases and poverty, all shape the developmental trajectory of a child in a LMIC. What we know about these threats to development come from research using CDAT. Walker *et al*³³ reviewed the risk and protective factors of early child development that can be modified by interventions in children under the age of 5 years in LMICs, including iron deficiency, malaria and inadequate cognitive stimulation. These three factors illustrate challenges of CDAT interpretation. Three prospective iron supplementation trials unexpectedly found no change in cognitive development measured by an adapted western CDAT.³⁴ It was plausible that the CDAT was not sensitive enough to the effects of supplementation of a nutritional deficiency, resulting in a shift to a more hypothesis-driven, focused testing of specific areas of development.³⁵ This shift to ensure the CDAT can evaluate hypothesised outcomes based on biological mechanisms was incorporated into developing an appropriate CDAT for severe malaria.³⁶ A team in Kilifi adapted a western tool for school-age children and then successfully evaluated its construct validity and sensitivity for discriminating cognitive deficits following severe malaria (high-risk), mild malaria (medium risk) and no previous admission with malaria (low-risk).²⁷ There were significant differences in scores between unschooled high-risk children and unschooled low-risk children. There was no significant difference in scores between high-risk and low-risk malaria children who were at school, highlighting the plausible benefit of formalised education in protecting from neurocognitive sequelae in LMIC settings. This finding was put into context by the authors, who reported that the group attending school was not necessarily representative of severe children,

since school attendance is limited by economic restraints and parental perception of whether their child would benefit from formal schooling. A child perceived to have cognitive deficits by their parents might not be sent to school. In a setting with a multitude of biological threats to early child development, the addition of poverty will limit the beneficial opportunities of a child's environmental influence on their development. However, there is variability in parental attitudes and behaviours within one socioeconomic class and also variability within the same socioeconomic class between cultures. A South African Home Screening Questionnaire, found mental development was primarily accounted for by the mother's ability to structure the child's environment for learning.^{36a} This was through the provision of play materials and maternal involvement in the child's activities, which were believed to promote development. The authors concluded that the socioeconomic variables evaluated did not solely determine maternal behaviour. Supporting this premise are studies suggesting maternal stimulation programmes within early childhood community-based stimulation interventions that can have long-lasting benefits on cognitive, socioemotional well-being despite adverse socioeconomic environments.³⁷ CDAT research will continue to inform the agenda for early intervention, before lifelong trajectories leading to large health and economic inequalities become fixed.

CONCLUSIONS AND FUTURE DIRECTIONS

Childhood developmental outcomes are firmly on the current global health agenda particularly now that it is time to set the new millennium developmental goals. However the relationship between future employability, quality of life measures and performance on a CDAT before the age of 5 years, is not known. Longitudinal research is needed to evaluate these relationships. In LMICs where early intervention support is limited, robust CDAT construction within culturally acceptable early child development programmes will help explore the relationship with later outcomes and promote the factors which protect against the neurocognitive detrimental effects of adversities faced by children in LMICs.³⁸

QUICK CHECKLIST FOR APPRAISING A CDAT

1. Does the CDAT adequately measure all aspects of the domain(s) theoretically affected by a risk factor or intervention?
2. Has the CDAT been shown to be reliable and valid in the population of interest?
3. Is the CDAT sensitive enough in the setting required to identify the changes expected for the risk factor or intervention?
4. Are the number of evaluations and duration of follow-up suitable for evaluating the outcome of interest?
5. In research studies, is there a suitable control group and have potential influences been considered?

Contributors SS conceived the article and wrote the first draft of the manuscript, which was then critically revised by BW and MG. All authors have seen and approved the final manuscript and agree to be accountable for all aspects of the work.

Funding The Wellcome Trust.

Competing interests None.

Provenance and peer review Commissioned; externally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: <http://creativecommons.org/licenses/by/4.0/>

REFERENCES

- 1 Grantham-McGregor S, Cheung YB, Cueto S, *et al*. Developmental potential in the first 5 years for children in developing countries. *Lancet* 2007;369:60–70.
- 2 Maternal and Child Nutrition Study Group. Maternal and child nutrition: building momentum for impact. *Lancet* 2013;382:372–5.
- 3 Lake A. Early childhood development-global action is overdue. *Lancet* 2011;378:1277–8.
- 4 Mwaniki MK, Atieno M, Lawn JE, *et al*. Long-term neurodevelopmental outcomes after intrauterine and neonatal insults: a systematic review. *Lancet* 2012;379:445–52.
- 5 Scherzer AL, Chhagan M, Kauchali S, *et al*. Global perspective on early diagnosis and intervention for children with developmental delays and disabilities. *Dev Med Child Neurol* 2012;54:1079–84.
- 6 Heckman JJ. Skill formation and the economics of investing in disadvantaged children. *Science* 2006;312:1900–2.
- 7 Grand Challenges Canada. Saving Brains: Unlocking the Potential for Development. Secondary Saving Brains: Unlocking the Potential for Development. <http://www.grandchallenges.ca/saving-brains/>
- 8 Fernald L, Kariger P, Engle P, *et al*. *Examining early child development in low-income countries: a toolkit for the assessment of children in the first five years of life*. Washington DC: World Bank, 2009.
- 9 Thompson B, Vacha-Haase T. Psychometrics is datametrics: The test is not reliable. *Educ Psychol Meas* 2000;60:174–95.
- 10 Al Nofal A, Schwenk WF. Growth failure in children: a symptom or a disease? *Nutr Clin Pract* 2013;28:651–8.
- 11 Garza C, de Onis M. Rationale for developing a new international growth reference. *Food Nutr Bull* 2004;25(1 Suppl):S5–14.
- 12 Guralnick MJ. Effectiveness of early intervention for vulnerable children: a developmental perspective. *Am J Ment Retard* 1998;102:319–45.
- 13 Bellman M, Byrne O, Sege R. Developmental assessment of children. *BMJ* 2013;346:e8687.
- 14 Turnbull H, Huerta N, Stowe M. *The Individuals with Disabilities Education Act as Amended in 2004*. Upper Saddle River, New Jersey: Pearson Education, Inc., 2004.
- 15 Dworkin PH. Developmental screening: (still) expecting the impossible? *Pediatrics* 1992;89(6 Pt 2):1253–5.
- 16 Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol Assess* 1994;6:284–90.
- 17 Prado EL, Hartini S, Rahmawati A, *et al*. Test selection, adaptation, and evaluation: a systematic approach to assess nutritional influences on child development in developing countries. *Br J Educ Psychol* 2010;80(Pt 1):31–53.
- 18 Marcovitch S, Zelazo PD. A hierarchical competing systems model of the emergence and early development of executive function. *Dev Sci* 2009;12:1–18.
- 19 Roze E, Meijer L, Van Braeckel KN, *et al*. Developmental trajectories from birth to school age in healthy term-born children. *Pediatrics* 2010;126:e1134–42.
- 20 Roberts G, Anderson PJ, Doyle LW. The stability of the diagnosis of developmental disability between ages 2 and 8 in a geographic cohort of very preterm children born in 1997. *Arch Dis Child* 2010;95:786–90.
- 20a Coll CG. Pathways to reading: the role of oral language in the transition to reading. *Dev Psychol* 2005;41:428–42.
- 21 Clark CA, Pritchard VE, Woodward LJ. Preschool executive functioning abilities predict early mathematics achievement. *Dev Psychol* 2010;46:1176–91.
- 22 Bull R, Espy KA, Wiebe SA. Short-term memory, working memory, and executive functioning in preschoolers: longitudinal predictors of mathematical achievement at age 7 years. *Dev Neuropsychol* 2008;33:205–28.
- 23 Pollitt E, Triana N. Stability, predictive validity, and sensitivity of mental and motor development scales and pre-school cognitive tests among low-income children in developing countries. *Food Nutr Bull* 1999;20:45–52.
- 24 McCall RB. The development of intellectual functioning in infancy and the prediction of later IQ. In: Osofsky JD, ed. *Handbook of Infant development*. New York: Wiley, 1979:704–41.
- 25 Pena ED. Lost in translation: methodological considerations in cross-cultural research. *Child Dev* 2007;78:1255–64.
- 26 Van Rie A, Mupuala A, Dow A. Impact of the HIV/AIDS epidemic on the neurodevelopment of preschool-aged children in Kinshasa, Democratic Republic of the Congo. *Pediatrics* 2008;122:e123–8.
- 27 Holding PA, Taylor HG, Kazungu SD, *et al*. Assessing cognitive outcomes in a rural African population: development of a neuropsychological battery in Kilifi District, Kenya. *J Int Neuropsychol Soc* 2004;10:246–60.
- 28 Abubakar A, Van De Vijver FJ, Mithwani S, *et al*. Assessing developmental outcomes in children from Kilifi, Kenya, following prophylaxis for seizures in cerebral malaria. *J Health Psychol* 2007;12:417–30.
- 29 Fernandes M, Stein A, Newton CR, *et al*. The INTERGROWTH-21st project neurodevelopment package: a novel method for the multi-dimensional assessment of neurodevelopment in pre-school age children. *PLoS One* 2014;9:e113360.
- 30 Abubakar A, Holding P, van Baar A, *et al*. Monitoring psychomotor development in a resource-limited setting: an evaluation of the Kilifi Developmental Inventory. *Ann Trop Paediatr* 2008;28:217–26.
- 31 Gladstone MJ, Lancaster GA, Jones AP, *et al*. Can Western developmental screening tools be modified for use in a rural Malawian setting? *Arch Dis Child* 2008;93:23–9.
- 32 Gladstone M, Lancaster GA, Umar E, *et al*. The Malawi Developmental Assessment Tool (MDAT): the creation, validation, and reliability of a tool to assess child development in rural African settings. *PLoS Med* 2010;7:e1000273.
- 33 Walker SP, Wachs TD, Grantham-McGregor S, *et al*. Inequality in early childhood: risk and protective factors for early child development. *Lancet* 2011;378:1325–38.
- 34 Pollitt E. The developmental and probabilistic nature of the functional consequences of iron-deficiency anemia in children. *J Nutr* 2001;131(2S-2):669S–75S.
- 35 Siegel EH, Kordas K, Stoltzfus RJ, *et al*. Inconsistent effects of iron-folic acid and/or zinc supplementation on the cognitive development of infants. *J Health Popul Nutr* 2011;29:593–604.
- 36 Holding PA, Kitsao-Wekulo PK. Describing the burden of malaria on child development: what should we be measuring and how should we be measuring it? *Am J Trop Med Hyg* 2004;71(2 Suppl):71–9.
- 36a Richter LM, Grieve KW. Home environment and cognitive development of black infants in impoverished South African families. *Inf Ment Health J* 1991;12:2:88–102.
- 37 Alderman H, Behrman JR, Grantham-McGregor S, *et al*. Economic perspectives on integrating early child stimulation with nutritional interventions. *Ann N Y Acad Sci* 2014;1308:129–38.
- 38 Engle PL, Fernald LC, Alderman H, *et al*. Strategies for reducing inequalities and improving developmental outcomes for young children in low-income and middle-income countries. *Lancet* 2011;378:1339–53.
- 39 Richardson K. *Models of cognitive development*. East Sussex, UK: Psychology Press, 1998.
- 40 Anderson VM. Assessing executive functions in children: Biological, psychological, and developmental considerations. *Neuropsychol Rehabil* 1998;8:319–49.
- 41 Wegner LM, Poon JK, Macias MM. Disorders of cognition, attention, language and learning. In: Elzouki AY, Harfi HA, Nazer HM, *et al*. eds. *Textbook of clinical pediatrics*. 2nd edn. London: Springer, 2001:613–3.
- 42 Grossman L. Normal child development. In: Elzouki AY, Harfi HA, Nazer HM, *et al*. eds. *Textbook of clinical pediatrics*. London: Springer, 2001:571–81.
- 43 National Scientific Council on the Developing Child. Children's Emotional Development Is Built into the Architecture of Their Brains: Working Paper No. 2, 2004.
- 44 AIDD. Definition of Intellectual Disability. Secondary Definition of Intellectual Disability, 2014. <http://aidd.org/intellectual-disability/definition-U-qYvjK3efQ>
- 45 Abubakar A, Holding P, Van de Vijver F, *et al*. Developmental monitoring using caregiver reports in a resource-limited setting: the case of Kilifi, Kenya. *Acta Paediatr Suppl* 2010;99:291–7.
- 46 Ertem IO, Dogan DG, Gok CG, *et al*. A guide for monitoring child development in low- and middle-income countries. *Pediatrics* 2008;121:e581–9.
- 47 Khan NZ, Muslima H, Begum D, *et al*. Validation of rapid neurodevelopmental assessment instrument for under-two-year-old children in Bangladesh. *Pediatrics* 2010;125:e755–62.
- 48 Khan NZ, Muslima H, Shilpi AB, *et al*. Validation of rapid neurodevelopmental assessment for 2- to 5-year-old children in Bangladesh. *Pediatrics* 2013;131:e486–94.
- 49 Gladstone MJ. *Developmental assessment in Malawi: the creation of a culturally appropriate assessment tool*. Liverpool: Thesis M.D., [2010], 2010.
- 50 The correspondence validity of Malawian preschool developmental assessment and school-age neurocognitive performance in the longitudinal assessment of severe malaria. Global Health Metrics & Evaluation: Data, Debates, Directions; Seattle, Washington, USA. Lancet, 2013.
- 51 Gandhi M, Teivaanmaki T, Maleta K, *et al*. Child development at 5 years of age predicted mathematics ability and schooling outcomes in Malawian adolescents. *Acta Paediatr* 2013;102:58–65.
- 52 Cicchetti DV. Methodological commentary the precision of reliability and validity estimates re-visited: distinguishing between clinical and statistical significance of sample size requirements. *J Clin Exp Neuropsychol* 2001;23:695–700.
- 53 Rankin G, Stokes M. Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. *Clin Rehabil* 1998;12:187–99.