SUPPLEMENT ARTICLE

The MAL-ED Cohort Study: Methods and Lessons Learned When Assessing Early Child Development and Caregiving Mediators in Infants and Young Children in 8 Low- and Middle-Income Countries

Laura E. Murray-Kolb,[1] Zeba A. Rasmussen,[2] Rebecca J. Scharf,[3] Muneera A. Rasheed,[4] Erling Svensen,[5,6] Jessica C. Seidman,[2] Fahmida Tofail,[7] Beena Koshy,[8] Rita Shrestha,[9,10] Angelina Maphula,[11] Angel Orbe Vasquez,[12,13] Hilda P. da Costa,[14] Aisha K. Yousafzai,[4] Reinaldo B. Oria,[14] Reeba Roshan,[8] Eliwasa B. Bayyo,[6] Margaret Kosek,[12,13] Sanjaya Shrestha,[9,10] Barbara A. Schaefer,[1] Pascal Bessong,[11] Tahmeed Ahmed,[8] Dennis Lang,[2,15] and The MAL-ED Network Investigators[a]

[1]The Pennsylvania State University, University Park; [2]Fogarty International Center, National Institutes of Health, Bethesda, Maryland; [3]University of Virginia, Charlottesville; [4]Aga Khan University, Karachi, Pakistan; [5]University of Bergen, Norway; [6]Haydom Lutheran Hospital, Haydom, Tanzania; [7]icddr,b (formerly the International Centre for Diarrhoeal Disease Research, Bangladesh), Dhaka, Bangladesh; [8]Christian Medical College, Vellore, India; [9]Institute of Medicine, Kathmandu, Nepal; [10]Royal Thai Army–Armed Forces Research Institute of Medical Sciences, Nepal and Thailand; [11]University of Venda, Thohoyandou, South Africa; [12]Asociacion Benefica PRISMA, Iquitos, Peru; [13]Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland; [14]Federal University of Ceará, Fortaleza, Brazil; and [15]The Foundation for the National Institutes of Health, Bethesda, Maryland

More epidemiological data are needed on risk and protective factors for child development. In The Etiology, Risk Factors and Interactions of Enteric Infections and Malnutrition and the Consequences for Child Health and Development (MAL-ED) cohort study, we assessed child development in a harmonious manner across 8 sites in Bangladesh, Brazil, India, Nepal, Pakistan, Peru, South Africa, and Tanzania. From birth to 24 months, development and language acquisition were assessed via the Bayley Scales of Infant and Toddler Development and a modified MacArthur Communicative Development Inventory. Other measures were infant temperament, the child's environment, maternal psychological adjustment, and maternal reasoning abilities. We developed standard operating procedures and used multiple techniques to ensure appropriate adaptation and quality assurance across the sites. Test adaptation required significant time and human resources but is essential for data quality; funders should support this step in future studies. At the end of this study, we will have a portfolio of culturally adapted instruments for child development studies with examination of psychometric properties of each tool used.

Keywords. child; cognitive development; infant; low-and middle-income countries; methods.

More than 200 million children aged <5 years from low- and middle-income countries (LMICs) fail to reach their full cognitive potential, leading to poor school performance and subsequent reduction in economic productivity [1]. Both biological and psychosocial factors including nutrient deficiencies, infectious diseases, environmental exposures, lack of adequate stimulation and/or learning opportunities, and maternal depressive symptoms place children at risk for poor cognitive development [2]. Children are often exposed to multiple risks simultaneously and as these risks accumulate, the developmental consequences become more pronounced [3–5].

In the 2007 *Lancet* series on Child Development in Developing Countries [1, 2], the International Child Development Steering Group (ICDSG) identified factors with sufficient evidence to recommend implementing risk reduction or prevention strategies. These risk factors include

[a]MAL-ED Network Investigators are listed in the Supplementary Appendix.

Correspondence: Laura E. Murray-Kolb, PhD, MA, The Pennsylvania State University, 219 Chandlee Laboratory, University Park, PA 16802 (lem118@psu.edu).

Clinical Infectious Diseases® 2014;59(S4):S261–72

DOI: 10.1093/cid/ciu437

inadequate provision of cognitively stimulating materials, growth retardation and low birth weight, iodine and iron deficiencies, malaria, violence, and exposure to heavy metals; many of these factors affect at least 25% of children in developing countries [6,7]. Development of strategies targeting these risk factors is hindered by a limited understanding of the mechanisms by which these factors alter cognitive development: specifically, how timing, duration, and severity of these risk factors affect cognition and whether negative cognitive effects are long-lasting and/or reversible. The ICDSG also acknowledged that a paucity of data from LMICs exists with respect to other domains of child development [2].

Hence, while child development consists of multiple interdependent domains (ie, cognitive, language, motor, and socioemotional), too few studies in LMICs have assessed the motor and socioemotional domains to make any significant conclusions regarding the above risk factors and these outcomes. Although new global studies on child development are under way, most of the evidence available on early childhood development has been gathered from the 10% of children who reside in high-income countries. Less is known about the risk and protective factors and developmental trajectories of the majority of children living in LMICs [8].

Grand Challenges Canada recently defined human capital as the

> ". . . productive wealth embodied in the intellectual, social and physical capability of the individual, and is key to the development of individuals, communities and whole societies. It reflects the cognitive abilities, social functioning (for example behavior and psychological functioning), physical capacity and health status that allow an individual to be a productive member of society." [9]

Other definitions of human capital exist, with all of them acknowledging the need for investment in health and education to improve human capital. The "productive wealth of societies" mentioned above begins in early childhood. Human brain development begins shortly after conception, and continues throughout gestation. Although a full-term infant is born with the number of neurons he or she will have as an adult, these neurons require growth, myelination, neural connections, and pruning over the early years of a child's life. Nutrients and various growth factors regulate brain development during fetal and early postnatal life; malnutrition and illness can influence neuronal maturation, particularly in the first few years of life [10–12]. During early life, the brain demonstrates its greatest degree of plasticity, but may also be more vulnerable to nutrient insufficiency relative to later periods. These early periods of brain development are crucial for later cognitive development and learning [13, 14].

The Etiology, Risk Factors, and Interactions of Enteric Infections and Malnutrition and the Consequences for Child Health (MAL-ED) study is a multidisciplinary observational prospective community-based study at 8 sites in countries with a high incidence of diarrheal disease and malnutrition [15]. The MAL-ED study investigates the hypothesis that infection or coinfection with specific enteropathogens contributes to malnutrition by causing intestinal inflammation and/or by altering intestinal barrier and absorptive function which, in turn, leads to growth faltering, stunting, deficits in cognitive development, and reduced immunogenicity of childhood vaccines. The study aims to identify the periods during the first 2 years of life when specific infections are associated with the greatest effect on growth and development. Factors evaluated for their effects include enteric and other infections, micronutrient levels, dietary intake, socioeconomic status, maternal depressive symptoms, and the home environment. The MAL-ED Network of investigators established technical subcommittees comprised of representatives from all institutions associated with the study and focused on the following major areas of investigation: disease surveillance (including anthropometry), gut function, microbiology, nutrition, vaccine response, and cognitive development. We report here the methods used and lessons learned by the Cognitive Development Technical Subcommittee, including the theoretical framework that guided the selection of measurements to assess impact on development, test selection, and strategies for comparability across sites with respect to training, adaptation, implementation, and quality assurance.

METHODS

Setting and Participants

Data collection took place in 8 field sites: Dhaka, Bangladesh (BGD), Fortaleza, Brazil (BRF), Vellore, India (INV), Bhaktapur, Nepal (NEB), Loreto, Peru (PEL), Naushero Feroze, Pakistan (PKN), Venda, South Africa (SAV), and Haydom, Tanzania (TZH). Specifics about each study site, including the participant characteristics, can be found elsewhere in this supplement.

The Cognitive Development Technical Subcommittee was comprised of a leader trained in nutrition and psychology with experience in conducting cognitive testing in LMICs, senior individuals with training in psychology/child development from each site, 2 psychometricians, a developmental pediatrician/epidemiologist, a data analyst, and an infectious disease physician/epidemiologist. The subcommittee's purpose was to ensure the training and harmonization of the cognitive protocol across sites, and used weekly teleconferences to share experiences on test adaptation, piloting, administration, and discussion of issues and problems.

Theoretical Framework

We developed a theoretical framework to conceptualize the interactions between infection with enteropathogens and postnatal child development (Figure 1). Diarrhea is related to nutrient loss, and relations between altered nutrient status and child

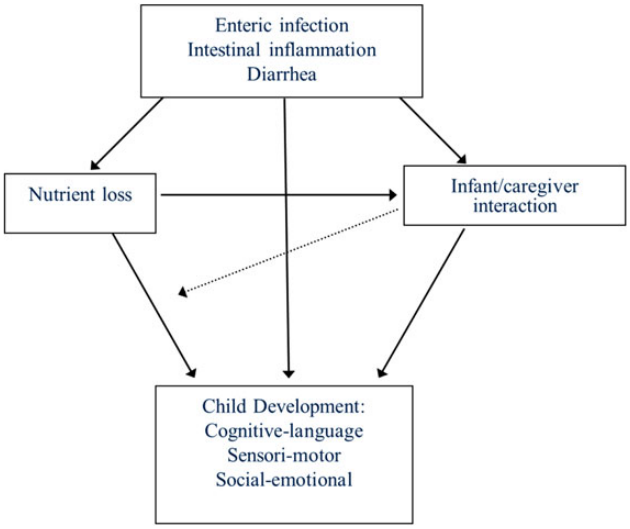

Figure 1. Theoretical framework for the relation between infection with enteropathogens and child development (only postnatal events are illustrated).

development are well established [2]. Diarrhea may alter infant/caregiver interaction (quality and/or quantity), and optimal infant/caregiver interaction is integral to child growth and development [16–18]. Alternately, repeated enteric infections in early life may have a direct effect on child development. Nutrient status may alter infant/caregiver interaction [19–21] and infant/caregiver interaction can modify the extent to which nutrient status affects child development [22] (depicted by the dashed arrow in Figure 1).

Test Selection

The Cognitive Development Technical Subcommittee was responsible for selecting appropriate and comprehensive assessments of child developmental outcomes in multiple domains while accounting for the cross-cultural nature of the MAL-ED study as well as other demands on study subjects. We were challenged to ensure that our assessments would result in a comprehensive analysis of the child's development while trying to limit the burden on participants. Keeping in mind the overarching hypothesis that repeated enteric infections in early life contribute to malnutrition which, in turn, leads to altered child development, test selection was based on published studies [23, 24] showing an association between our outcomes of interest and nutrient deficiencies likely to be encountered at MAL-ED study sites as well as the potential for administration in a harmonious manner across all 8 sites.

Few tests that assess development have been produced specifically for LMIC settings. Tests that are produced and standardized in 1 language or culture are not automatically valid in a setting that differs from the original population [25, 26]. It was challenging to find tests that accurately assess development at these young ages and are also predictive of later development in our domains of interest. Many of the common infant development assessments have been criticized for their poor predictive validity in terms of intelligence at older ages. However, poor predictive validity does not necessarily imply low sensitivity to high-risk conditions for impaired development. Additionally, several of the common infant assessments have been revised in recent years to include items that have been shown to predict later cognitive competence [27].

We chose a global test of development as our primary outcome and supplemented it with a more specific test of language development as early language development may have good predictability for later school achievement [28]. Although most global development tests typically include measurement of language development, best practices for assessing child development emphasize the collection of data from multiple sources (eg, observer and caregiver). Therefore, for our specific test of language development, we chose a parent report questionnaire.

The subcommittee sought to identify factors known to affect child development to carefully control for these variables in our analyses. A review of the literature examining the relation between nutrient deficiencies and development revealed that many studies have not accounted for environmental factors that may impact development [2]. Although some studies have collected data on the physical environment of the child, most have not accounted for maternal reasoning skills, child temperament, parenting, or maternal depressive symptoms. Studies have consistently revealed a strong association between infant temperament, the quality of parental engagement, and child development [29, 30]. Parenting plays a crucial role in child development [31, 32] and is particularly influential during early development [31, 33]. Maternal depressive symptoms are related to alterations in mother–infant interactions and insecure mother–infant attachments [34–36] that may precipitate child behavioral, developmental, and cognitive delays lasting years beyond infancy [36–40]. Finally, maternal intelligence has long been acknowledged as influential on child development [41–43] but is largely ignored in studies assessing malnutrition and child development. Therefore, we considered it important to account for each of these factors in our study.

Although prenatal maternal factors (nutrient deficiencies, prepregnancy weight, gestational weight gain) may exert significant influence on postnatal child development, enrollment in the MAL-ED study occurred after the birth of the child, so prenatal factors were not considered as part of this study.

As part of the MAL-ED study, information was collected on parental education and age, breastfeeding practices, child nutrient status, and household environmental measures and socioeconomic status that will all be used in our final analyses. Our

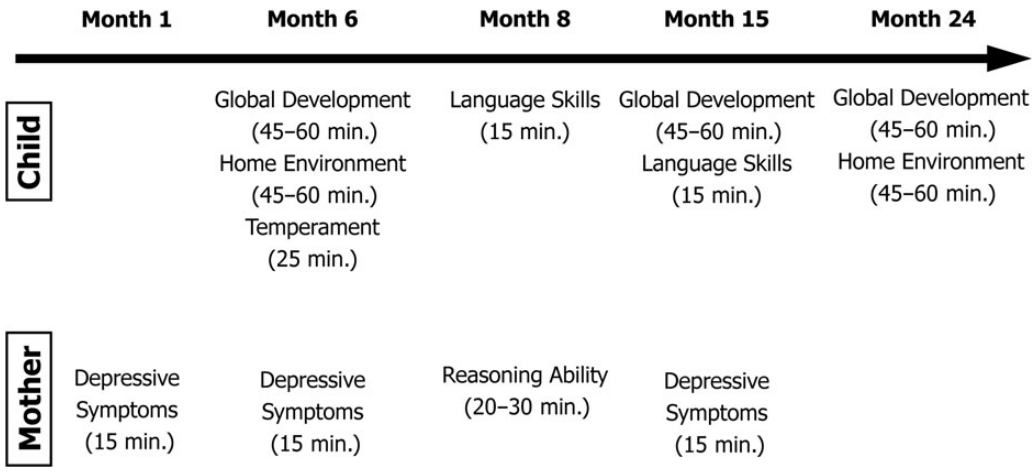

Figure 2. Cognitive data collection scheme; domain tested (time to administer).

testing battery included the following child and maternal measures, and psychosocial environment measure (Figure 2).

Child Measurements

Global Development

The Bayley Scales of Infant and Toddler Development (BSID-III) [27] were chosen to assess global child development at 6, 15, and 24 months of age. Assessment timing was chosen to correspond with critical points in development as well as scheduled blood collection to facilitate the linkage between developmental measures and nutrient status at specific time points. The BSID-III was used as the primary cognitive development outcome measure and administered to all children. It contains 5 individually administered subscales: Cognitive, Fine Motor, Gross Motor, Receptive Language, and Expressive Language; and 2 parent/guardian-completed rating scales: Social-Emotional and Adaptive Behavior. We decided against the collection of the Adaptive Behavior questionnaire for 2 reasons: first, the time commitment required and second, significant overlap with items on the temperament scale that we chose for use in this study. The assessments took approximately 45–60 minutes and, in most circumstances, were conducted in a central location identified by study field site coordinators that was quiet and comfortable for the child, with few distractions. However, some of the BSID-III assessments needed to be done at the child's home, in which case the deviation was recorded so that we could account for location of administration in our analyses.

The BSID have been criticized for their poor predictive validity for later measures of intelligence [44–46]. However, this is not surprising given that the early versions of the BSID generally assessed sensory capacities, motor achievements, and affective responses such as orienting, reaching, and smiling. Tests administered to older children target very different skills, usually those related to language, reasoning, and memory. Despite criticisms of poor predictive validity, the BSID have been shown to be sensitive to accurate measurement of child development [27]. The most recent revision of the BSID (used in this study) took many of these criticisms into account and added as well as deleted some items to improve the scales. Among the items added were habituation and novelty preference. Infant performance on these types of tasks has been shown to predict later cognitive competence [47].

The criterion for selection of our primary cognitive development outcome measure was that we wanted a well-known scale that would assess global child development in various domains. Several alternative scales to assess global child development exist [48, 49]; however, reasons for excluding these included the time required for administration as well as poor sensitivity and specificity (ie, providing less useful information for continuous tracking of specific changes in child development [50]). Also, several of the MAL-ED sites had used the BSID in previous studies [51] (A. Yousafzai, personal communication, 2013). Finally, the BSID are the most widely used assessment of child development worldwide, thereby facilitating comparison with prior studies conducted around the globe.

Language Skills

The MacArthur Communicative Development Inventory (CDI) [52], Words and Gestures, was used to assess verbal fluency at 8 and 15 months. The earliest age at which the CDI can be completed is 8 months. The 15-month time point was chosen to align with administration of the BSID-III. The CDI is a parent report of the child's ability to understand and/or say specific words. Typically, the CDI is completed by the primary caregiver; however, given the wide variation in caregiver literacy at the MAL-ED sites, the CDI was administered by a trained

interviewer who presented each item to the primary caregiver orally and then recorded the primary caregiver's answers. The assessment took approximately 15 minutes and was administered in the child's home.

Early child language development (in the first year of life) has been associated with both cognitive functioning and academic achievement at 11 years of age [28]. Thus, these measures have been shown to have good predictability for later school achievement. We also chose to measure language development as it is important for school readiness and a child's ability to learn in school [53]. Interestingly, it has been shown that velocity and acceleration in development of verbal fluency early in life are predictive of later vocabulary, especially in children from low-socioeconomic-status backgrounds [53].

We chose the MacArthur CDI scale as it is very easy to administer and takes a minimum amount of time, has been used worldwide, and is a caregiver report of a child's language abilities. Use of the CDI in conjunction with the BSID-III allows us to compare caregiver and researcher assessments of a child's language ability. Finally, at least 1 of the MAL-ED sites had previously used the principles from the MacArthur CDI when developing a language inventory for a previous study [54] and, therefore, we believed it could be adapted and used by the other MAL-ED study sites.

Temperament

Given the known association between infant temperament, parent–child interactions, and child development [29, 30], we believed it was vital to collect data on infant temperament to properly interpret the child development outcomes. Initially, we selected the Infant Behavior Questionnaire [55]; however, after piloting at the field sites, it quickly became evident that this test was not feasible for use at the MAL-ED sites due to questionnaire length and caregiver confusion with the response categories. After further discussion, the Cognitive Development Technical Subcommittee chose the Infant Temperament Scale (ITS), authored by Dr Theodore Wachs (T. Wachs, modified for use by MAL-ED, personal communication, 2010) and originally developed for research purposes in a LMIC (Peru). We believed that the ITS could be adapted to other LMIC countries, including our MAL-ED sites. The original version of the scale contains 112 items assessing 8 dimensions: activity, positive emotionality, negative emotionality, sociability, fear, attention, cooperativeness, and soothability. Items are designed to be answered by the caregiver using a 5-point scale ranging from "my child is always like this" to "my child is never like this." Due to its length, using the ITS in its entirety was not feasible, and 2 dimensions were dropped (fear and cooperativeness). The final administered scale contained 47 items assessing 6 dimensions. This adapted ITS was administered by a trained interviewer who presented each item to the primary caregiver, and then recorded the primary caregiver's answers. The questionnaire was administered at the child's home and took approximately 25 minutes to complete. A 6-month time point for administration was chosen to coordinate with the first BSID-III assessment.

Maternal Measures

Reasoning Ability

Maternal reasoning ability was assessed because it is known to be highly correlated with the child's mental abilities [41–43]. A combination of the Raven's Coloured Progressive Matrices and Standard Progressive Matrices (RCM) [56] was administered to the biological mothers of enrolled children at 6–8 months of age. The RCM is a nonverbal test of reasoning ability that measures the ability to form comparisons, reason by analogy, and organize spatial perceptions into systematically related wholes. The problems in each set build upon the previous solution logic and become progressively more difficult. This allows for the assessment of the consistency of an individual's intellectual activity across successive lines of thinking. We selected this particular test as the measures taken are suitable for populations who have not participated in a formal education system [57, 58]. RCM test content was also considered to be less culturally specific, and thus may be administered in diverse settings with relative ease [59, 60]. Finally, nearly all of the MAL-ED study sites had prior experience with the RCM. The assessment took approximately 20–30 minutes and, in most circumstances, was conducted in a central location that was quiet and comfortable with few distractions. As some assessments needed to be done at the subject's home, this deviation was recorded so that we could account for location of administration in our analyses. The RCM was administered by a trained tester who recorded the mothers' answers after the mothers pointed to their answer of choice.

Depressive Symptoms

Maternal psychological disturbances, specifically, depressive symptoms, are a known risk factor for poor child development outcome [33–39]. We decided to collect information on psychological disturbances in the biological mothers of study children at ages 1, 6, 15, and 24 months. The 1-month time point was chosen to capture psychological disturbances in the early postpartum period and the later time points were chosen to correspond with the BSID-III assessments. The Self Reporting Questionnaire–20 [61] (SRQ-20), a scale developed by the World Health Organization specifically for use in developing countries [62], was selected. The SRQ-20 is comprised of 20 items designed to assess psychological adjustment related to depressive symptoms. Questions are answered with a simple "yes" or "no" and the instrument may be self-administered or interview-administered.

One criticism of the SRQ-20 is that no global, generally applicable cutoff score is recommended. This is because the score

used depends on the language or culture of administration, method of administration, population being studied, and the needs of the particular research study (eg, sensitivity, specificity) [61]. However, the Cognitive Development Technical Subcommittee believed the SRQ-20 is the most appropriate questionnaire for the purposes of the MAL-ED study. This questionnaire was interview-administered at the subject's home, and took approximately 15 minutes to complete. Given the sensitive nature of the questions, the interviewers were careful to ensure the privacy of the mother's answers and, where culturally acceptable, no one (other than the mother and the interviewer) was allowed in the room while the questions were being asked. At the completion of the questionnaire, if the interviewer was concerned about the mother's mental health (based on answers that she gave to the questions), she or he immediately reported this concern to the field supervisor. The field supervisor then proceeded with culturally appropriate follow-up/referral. Each study site identified a locally relevant SRQ-20 cutoff score to identify respondents who met the criteria for referral. A positive answer to having suicidal thoughts was also grounds for referral, regardless of the total SRQ-20 score. The referral support services available varied by site according to the local standard of care.

Psychosocial Environment

A disadvantaged environment is among the many factors identified as influencing child cognition and development [63, 64]. We chose to collect information about the child's environment to account for the effect that it has on child development. The Home Observation for the Measurement of the Environment (HOME) scale [65] (Infant/Toddler version) was selected by the Cognitive Development Technical Subcommittee. One of our primary reasons for choosing to use the HOME scale is that studies have shown the HOME scale to be a predictor of later school achievement [66, 67].

This internationally recognized scale has been used successfully in >100 countries [64, 65] and validated in Bangladesh and Brazil, among others (A. Yousafzai, personal communication, 2013). It measures both the quality and the quantity of stimulation and support that is available to a child in his/her home environment. The version we used was previously adapted for use in Bangladesh [68]. The modified version of the HOME scale is comprised of 48 items clustered into 6 subscales: (1) responsiveness to parent, (2) avoidance of restriction and punishment, (3) organization of the environment, (4) appropriate play materials, (5) parental involvement, and (6) variety in daily stimulation. The items were selected to provide information from the child's perspective on stimuli that have been found to affect children's cognitive development. A 45- to 60-minute observation period in the home was completed for each child at 6 and 24 months of age. The observation occurred when the child was awake and engaged in activities typical for that time of day. The information was collected from observations supplemented by caregiver interview. These time points were chosen to correspond with the 6- and 24-month BSID-III assessments.

Test Translation, Adaptation, and Piloting

Most of the selected assessments were not produced in LMIC settings; thus, adaptation, piloting, and validity assessments were necessary at each study site. Initial translation was done by bilingual MAL-ED field staff familiar with the local language and culture, and difficult items and wording were decided by group consensus. Back-translation by bilingual staff different from the person conducting the original translation was then done. Testing materials were reviewed for culturally irrelevant stimuli, and potential substitutions were identified and discussed. Following best practices guidelines, the MAL-ED adaptations were intended to be relevant and easily understood, and to match the difficulty level of the original item [69].

In adapting the Cognitive, Receptive Language, and Expressive Language scales of the BSID-III, we were careful to match the style of the original item such that photographs were replaced with photographs, cartoons with cartoons, etc. For example, a cartoon drawing in the original BSID-III showing a boy building a snowman was adapted by the BGD team to show a boy building a house out of mud or sand, thereby making it more relevant to our sites that do not have any snowfall. In another adaptation, a photograph of a man vacuuming was replaced with a photograph of a woman sweeping for our study sites where vacuum cleaners are not used.

Extensive pilot testing of the MacArthur CDI revealed that many of the words that were common or "easy" in an English or North American context were not commonly used in the MAL-ED study sites. For example, *duck*, *kitty*, and *juice*, which were easier words in the United States, were not commonly used in some MAL-ED sites. Furthermore, *chicken*, which is considered a "medium" difficulty word for 12-month-olds in the United States, was found to be "easy" for 12-month-olds in INV and NEB, but "difficult" for the toddlers in TZH (Table 1). Fruits used in the scale were translated based on what was locally available at each site. For example, some sites used apple, whereas others used banana or jackfruit.

Adaptations to the selected tests under consideration were discussed and agreed upon during weekly teleconferences. Adapted tests were piloted on approximately 20–30 children (depending on the test) at each site who were the target age but who were not enrolled in the MAL-ED study. Pilot data were then assessed by each site to determine acceptability of the adaptations made.

Psychometric Analyses

To ensure that our chosen assessments are reliably measuring our target developmental outcomes and that these constructs

Table 1. Examples of Translated Words With Level of Difficulty[a]

	United States	BGD (80)	BRF (40)	INV (30)	NEB (50)	PEL (NA) [b]	PKN (30)	SAV (40)	TZH (29)
Word	Chicken	Murghi	Galinha	Kozhi	Kukhura/Kha	Gallina/Pollo	Choozo	Khuhu	Konki
Difficulty level	Medium	Medium	Medium	Easy	Easy	Medium	Medium	Medium	Difficult
Word	Milk	Dudh	Leite	Paal	Dudh/Duru	Leche	Kheer	Mafhi	Ilwa'a
Difficulty level	Easy	Easy	Difficult	Easy	Easy	Medium	Easy	Easy	Difficult

The number of children on which words were piloted is shown in parentheses next to each site abbreviation.

Abbreviations: BGD, Dhaka, Bangladesh; BRF, Fortaleza, Brazil; INV, Vellore, India; NA, not available; NEB, Bhaktapur, Nepal; PEL, Loreto, Peru; PKN, Naushero Feroze, Pakistan; SAV, Venda, South Africa; TZH, Haydom, Tanzania.

[a] Difficulty level for countries other than United States was obtained on a pilot of approximately thirty 12-month-old children from each site.

[b] Data from 15-month-old children.

are being measured equivalently across sites, the data are being evaluated for their psychometric properties [70]. The translated assessments in our study require testing for evidence of construct, metric, and functional equivalence to the original test [71]. Analyses include the Item Response Theory analyses, internal structural congruence, and validation by nomological network. The results of these analyses will inform the Cognitive Development Technical Subcommittee as to the validity and reliability of our measures as well as the extent to which they can be compared across MAL-ED sites.

Study Personnel

At the outset of the study, the Cognitive Development Subcommittee agreed on criteria to select individuals as eligible to collect the cognitive development data for the MAL-ED study at each site. These criteria included a minimum master's-level education in psychology or child development. Although most study sites were able to find qualified individuals locally, some were not. In TZH, this was a significant challenge due to the lack of qualified personnel in the country, as until very recently (2012), no masters-level psychology training was available in Tanzania. For TZH site staff, it was not possible to find anyone who had the combination of local language knowledge (Iraqw) and training in psychology; therefore, local field staff were trained to complete the testing.

Training Activities for Data Collection

To ensure that all MAL-ED sites collected cognitive development information similarly and according to protocol, regional on-site training was provided by central MAL-ED personnel at the following sites: SAV, TZH (2 sessions), PEL (2 sessions for PEL and BRF staff), INV (for INV and BGD staff), and NEB (for NEB and PKN staff). Training sessions lasted for 5–6 days, and attendees from each site included 1 to 2 trainers with expertise in clinical psychology/nutrition and experience in test administration, and the individuals who would be conducting the chosen cognitive development assessments in the field, as well as individuals supervising those collecting the data. The basic format for the training sessions included introduction to the MAL-ED study with an emphasis on the activities of the cognitive development team, 3 days of training on the BSID-III (including practice with infants), 1 day of training on the RCM and MacArthur CDI (including practice with mothers and peers), and 1 day of training on the HOME scale, the SRQ-20, and the ITS (including peer-mediated practice). The teams responsible for collecting the cognitive development data at each site were then given a practice schedule to follow to become proficient on each of the study instruments. This entailed administration of the tests to children who were not enrolled in MAL-ED until they felt that mastery was achieved. After completing the practice scheduled, testers from most sites videorecorded the administration of the tests for review by the leader of the Cognitive Development Technical Subcommittee and/or the on-site field supervisors of the cognitive data. The videos were evaluated for accuracy of test administration as well as scoring. Once a tester mastered standard test administration (ie, no administration or scoring mistakes) and scored the complete test battery on 5 different children, he or she was "certified" as able to officially collect data for the project.

Subsequently, cognitive development supervisors at each field site monitored the progress of each of their team members and reported any questions/difficulties back to the Cognitive Development Technical Subcommittee leader (L. E. M.-K.) for discussion and recommendation of a course of action. Refresher training for the administration of the 15-month BSID-III was conducted by central MAL-ED staff in PEL (for PEL and BRF staff) and TZH. Other field sites conducted refresher training by senior psychologists on site (BGD, INV, NEB, PKN, SAV).

Quality Assurance/Quality Control

Because the MAL-ED study was implemented at 8 culturally diverse sites, it was important to ensure that harmonized

protocols were used at each site. Different types of quality assurance/quality control (QA/QC) strategies were implemented. Those central to all sites included:

- Weekly conference calls: Weekly calls ensured that all sites had common understanding of standard operating procedures, and that any change was communicated to each site. Calls also provided the opportunity for study sites to discuss any emerging field issues.
- Central video reviews by the Cognitive Development Technical Subcommittee leader (L. E. M.-K.): Overall, 275 of 5800, or 5.1% (1.7%–11.0%) of all BSID-III and RCM assessments were filmed and sent to L. E. M.-K. by each site. This included 231 of 3991, or 6.3% (2.3%–13.8%) of BSID-III, and 44 of 1809, or 2.5% (0%–7.0%) of RCM assessments. Trained personnel in Dr Murray-Kolb's lab reviewed the videos and sent feedback regarding test administration back to each site. The feedback provided included both positive comments as well as areas for improvement, with a focus on any drift that might be occurring in data collection. In subsequent videos, the reviewers would look for evidence that the testers had incorporated the previous feedback into their testing sessions. TZH was unable to participate in this video review process until December 2012 due to stipulations of its institutional review board.

Other QA/QC strategies employed by the MAL-ED study sites are listed in Table 2 and below:

- Random form checking: At least 10% of forms were checked by senior psychologists at each site to ensure consistency of scoring.
- Surprise visits by senior psychologists: Random surprise visits were made by the supervisor to observe assessments and provide guidance to field psychologists if needed.
- Interrater visits by the senior psychologists: This strategy was similar to the above, but it also included simultaneous scoring of the tool by the senior psychologist. The scores were then matched for disagreements and followed up in discussions. This strategy was particularly useful for observation tools such as the HOME inventory.
- Discussion meetings by the senior psychologist on site or via Skype: In these meetings, emerging issues from the field site were discussed and a common strategy for addressing the issues was formulated. Meetings could also include quick testing of understanding of tools through questions, administration of items, role-plays of interview tools, or review of a topic included in the training (eg, risk factors for child development).
- Video reviews: Videos were reviewed internally by the senior psychologist, and feedback was sent via email and discussed during the field visits. The reviewed video was studied again by the field team to note errors identified. At certain sites, as the field team developed, internal reviews were completed by the field psychologists.
- Shadowing by senior members of the field psychology team of another child development project on the site: This strategy was used in PKN, which was particularly helpful given that PKN had problems with retention of psychologists. New field team members were shadowed by senior members for support and quality enhancement until they were able to conduct assessments independently.

DISCUSSION

Over the past 5 years, the Cognitive Development Technical Subcommittee has gained invaluable experience with respect to implementing a harmonized protocol for the assessment of

Table 2. On-site Quality Assurance/Quality Control Strategies

Site	Random Form Checking	Random Field Visits by Senior Staff	Interobserver Reliability	Daily or Weekly Team Discussions of Emerging Issues	On-Site Video Review	Refresher Training Sessions	Shadowing by Senior Field Psychologists
BGD	√	√	√	√ (W)	√	√	
BRF	√	√		√	√		
INV	√	√		√	√	√	
NEB	√	√	√	√ (W)	√	√	
PEL[a]	√	√					
PKN	√		√	√	√	√	√
SAV	√	√	√	√	√	√	
TZH	√	√		√[b]	√[c]	√	

Abbreviations: BGD, Dhaka, Bangladesh; BRF, Fortaleza, Brazil; INV, Vellore, India; NEB, Bhaktapur, Nepal; PEL, Loreto, Peru; PKN, Naushero Feroze, Pakistan; SAV, Venda, South Africa; TZH, Haydom, Tanzania; W, weekly.

[a] PEL had only 1 psychologist, so many of these measures were not applicable.

[b] Every third month.

[c] By 2 expatriate psychologists during April–September 2011 and by a specially trained psychology student October–December 2011.

child development in 8 LMICs. Here we discuss important challenges, lessons learned, strengths, and limitations.

Adaptation of Assessments

Many of the assessments used in the MAL-ED study were created and normed in a North American context. As discussed earlier, several of the testing items were not appropriate to LMICs, and thus needed translation and adaption to local cultures. Some questions on the HOME scale were not relevant in all sites (eg, not all sites had homes that included windows for ventilation). In addition, some of the practices inquired about in the ITS were not relevant to all sites; for example, one question asks about infant behavior during nail cutting, but in some of the sites, nails were only cut while the child was sleeping. Adjustments had to be made when comparing across sites for several of the ITS items.

While translating the assessments into the 10 languages used across the 8 sites, we learned the importance of having study personnel who speak the local languages fluently. Not only was this helpful for translation, but also for making appropriate adaptations and conducting video review. At 3 sites (NEB, INV, and TZH), 2 local languages were spoken. In INV, some participants spoke a combination of 2 different local languages (with different scripts), which created challenges for translation. At the TZH site, finding an individual who spoke one of the local languages (Iraqw) and English fluently and who could easily translate the assessments was difficult. In translating the MacArthur CDI, we had to use words that were comparable across sites and cultures. A further challenge was translating the tests into local languages that have limited written traditions. The psychometric analyses to evaluate the effectiveness of these adaptations are under way.

The literacy rates of mothers and caregivers varied by site. In addition to having all assessments administered by study personnel rather than self-administered by study subject, some sites found it important to use visual rating scales. INV found that a scale depicting a bottle with various levels of liquid in it was well-accepted and useful for depicting answer choices ranging from "my child is always like this" to "my child is never like this" (Figure 3).

Adequate time for test adaptation needs to be built into project timelines, with common agreed-upon protocols at the start. Ideally, several months prior to beginning the study, sufficient time to adapt the tests, pilot the adaptations, and perform metrics on the findings to formally validate the instruments would have been beneficial. For the MAL-ED study, the time allotted to adapt and pilot tests prior to study initiation was extremely challenging. Simultaneous testing of one age group while piloting adaptations for the upcoming age group was an added complexity and burden on study site personnel. The process of adaptation requires substantial efforts with identification of the right population, data collection, and analyses. Future studies would do well to have dedicated funding and time for an adaptation, piloting, validation phase that occurs before the formal data collection starts.

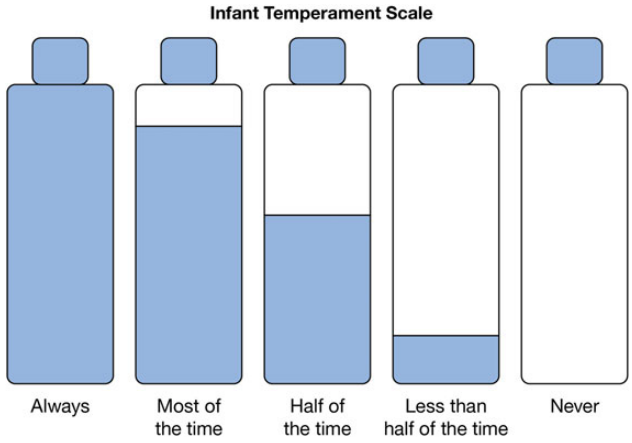

Figure 3. Visual scale used to depict answer choices for the Infant Temperament Scale.

Our team approach allowed experts in early childhood assessment from different areas of the world to review each site's adaptations and make suggestions. Having broad expertise from many different cultures and languages added to the ability to make adaptations in a relevant and useful manner. We found frequent communication (eg, weekly phone calls, a central collaboration website, and email) to be essential to collaboration for adaptation of the tests chosen.

Training

Training was especially important when working together on a large study that spanned sites in 8 different countries. Training >1 person per site was useful so teams could problem-solve together. Psychologists and test administrators need checks and reminders to prevent drift, and coming together in the same location was helpful for consistency. Although gathering everyone from 10 nations in one place annually was difficult and costly, it was extremely worthwhile to promote and strengthen collaboration within and across sites.

Our psychologists generally reported that in-person training was preferable to video training. One site (INV) recorded the in-person trainings and then used the videos for reference and refresher training later in the study. Most of our testing materials and instructions were available only in English. Several sites (PEL, BRF, PKN, NEB, TZH) translated administration instructions or forms into local languages. Of all instruments used, the BSID-III required the most training and practice due to its complexity and length.

At many sites, it was difficult to find psychologists willing to participate in the study and be trained. A few sites had excellent

child development centers already in place with experienced child psychologists and developmental pediatricians available to collect data. It is important to locate available psychologists before a study begins and to ensure there are senior psychologists who can provide assistance and monitoring to those with less experience. Furthermore, site capacity should be an important consideration for collaboration, as having at least 1 person with a combination of child development background, community-based research experience, and leadership skills to lead the on-site component is advantageous.

A single initial basic training was insufficient, and had to be followed by regular supplementary trainings. When planning for training, the supplementary trainings need to be adequately budgeted and planned for as well. If the study budget does not allow for central trainings, on-site supervisors could deliver these trainings. The trainings can be evaluated and feedback can be shared with the larger group.

Test Administration

Some study families found it difficult to come to the study center due to various cultural or logistical challenges. Consequently, some assessments were done in participant homes (a small percentage of BSID, RCM). Because we found that crowded space and distractions were often a challenge, in-home assessments were noted for further analysis. At some sites, it was culturally unacceptable for unaccompanied male staff to administer tests to mothers and children; thus, male testers had to be accompanied by a female staff member for all home visits.

Occasionally we found that fathers made decisions for the family and thus, even if the mother consented to study procedures, the father had to give consent as well. We also had to be flexible with who was considered the primary caregiver over the life of the child. At some sites, mothers returned to work or school in a different location, and the children were cared for by other relatives.

In administering the RCM to mothers, a decision was made by the Cognitive Development Technical Subcommittee to start in the middle (with set C from the Standard Progressive Matrices) and, if the items were too challenging, to then move to the Coloured Progressive Matrices. However, some sites discovered that mothers were frustrated if they did not know the answers and had to go backward, and other sites would have preferred to start at the beginning. Although the RCM is created to measure reasoning and problem solving without the need for language or formal education, several mothers in the study did not want to participate, citing lack of any schooling as the reason.

Quality Assurance and Control

Central procedures and measures that seek to keep assessments consistent across sites, such as central video reviews, are key to promoting reliability of test administration. It was equally important for each study site to have its own QA/QC measures in place. Communication between sites about how to perform quality checks facilitated the maintenance of high standards and helped pinpoint the need for refresher trainings if difficulties were found in test assessments.

CONCLUSIONS

The MAL-ED cohort study is a large, groundbreaking study that has much to offer other multisite studies. The measures of cognitive function presented here are part of a comprehensive study on enteric infections, gut function, and nutrition and their impact on child development. This study provides an important opportunity to investigate risk factors for cognitive impairment. However, the sheer size and scale of the study was also a limitation in terms of the burden on the families and study staff involved. Ideally, we would have chosen several additional tests that assessed specific developmental abilities and cognitive functioning; however, we had to be sensitive to placing additional demands on the study participants as part of the larger study.

Children in this study were followed from just after birth until 24 months of age, and ideally, we would have collected information on prenatal factors such as maternal prepregnancy mental health status and micronutrient deficiencies. Additionally, environmental measures such as indoor air pollution and tests of drinking water were outside the scope of this study but would have provided useful information to our assessment of child development. Nevertheless, the measures collected such as child global development, language development, maternal depressive symptoms, maternal reasoning abilities, child temperament, and home environment ensure that the MAL-ED study is a comprehensive investigation of child development, and the most diverse multisite collaboration in LMICs to date. The Cognitive Development Technical Subcommittee followed best practices guidelines, when available, in choosing the set of assessments as well as for adapting and evaluating the measures [72–74].

The MAL-ED study's Cognitive Development Technical Subcommittee has developed a portfolio of tools translated into 10 languages and enhanced capacity at the study sites. A thorough psychometric analysis of each of our selected measures, which is a critical step often overlooked in other studies, is in progress [75]. Although challenges abound with a study of this magnitude, this in-depth and thorough look at early childhood development in the context of malnutrition and enteric disease will provide crucial information on early determinants of growth, learning, and cognitive development in children in LMICs.

Supplementary Data

Supplementary materials are available at *Clinical Infectious Diseases* online (http://cid.oxfordjournals.org). Supplementary materials consist of data

provided by the author that are published to benefit the reader. The posted materials are not copyedited. The contents of all supplementary data are the sole responsibility of the authors. Questions or messages regarding errors should be addressed to the author.

Notes

Acknowledgments. We thank the staff and participants of the MAL-ED Network Project for their important contributions.

Financial support. The Etiology, Risk Factors, and Interactions of Enteric Infections and Malnutrition and the Consequences for Child Health and Development Project (MAL-ED) is carried out as a collaborative project supported by the Bill & Melinda Gates Foundation, the Foundation for the National Institutes of Health, and the National Institutes of Health, Fogarty International Center.

Supplement sponsorship. This article appeared as part of the supplement "The Malnutrition and Enteric Disease Study (MAL-ED): Understanding the Consequences for Child Health and Development," sponsored by the National Institutes of Health and the Foundation for the National Institutes of Health.

Potential conflicts of interest. All authors: No reported conflicts.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Grantham-McGregor S, Cheung YB, Cueto S, Glewwe P, Richter L, Strupp B; International Child Development Steering Group. Developmental potential in the first 5 years for children in developing countries. Lancet **2007**; 369:60–70.
2. Walker SP, Wachs TD, Gardner JM, et al; International Child Development Steering Group. Child development: risk factors for adverse outcomes in developing countries. Lancet **2007**; 13:145–57.
3. Hall JE, Sammons P, Sylva K, et al. Measuring the combined risk to young children's cognitive development: an alternative to cumulative indices. Br J Dev Psychol **2010**; 28:219–38.
4. Tomalski P, Johnson MH. The effects of early adversity on the adult and developing brain. Curr Opin Psychiatry **2010**; 23:233–8.
5. Stiles J. The fundamentals of brain development: integrating nature and nurture. Brentwood, CA: Harvard University Press, **2008**.
6. Committee on the Prevention of Mental Disorders and Substance Abuse Among Children, Youth, and Young Adults. In: O'Connell ME, Boat T, Warner KE, eds. Preventing mental, emotional, and behavioral disorders among young people: progress and possibilities. Washington, DC: National Academy Press, **2009**.
7. Committee on Integrating the Science of Early Childhood Development. In: Shonkoff JP, Phillips DA, eds. From neurons to neighborhoods: the science of early childhood development. Washington, DC: National Academy Press, **2000**.
8. Britto PR, Ulkuer N. Child development in developing countries: child rights and policy implications. Child Dev **2012**; 83:92–103.
9. Grand Challenges Canada. Saving brains, **2011**. Available at: http://www.grandchallenges.ca/wp-content/uploads/savingbrains_economicRFP_2011Oct12_EN.pdf. Accessed 14 April 2014.
10. Georgieff MK. Nutrition and the developing brain: nutrient priorities and measurement. Am J Clin Nutr **2007**; 85:614–20S.
11. Cusick SE, Georgieff MK. Nutrient supplementation and neurodevelopment: timing is the key. Arch Pediatr Adolesc Med **2012**; 166:481–2.
12. Martinez M. Myelin lipids in the developing cerebrum, cerebellum, and brain-stem of normal and undernourished children. J Neurochem **1982**; 39:1684–92.
13. Rice D, Barone S. Critical periods of vulnerability for the developing nervous system: evidence from humans and animal models. Environ Health Perspect **2000**; 108:511–33.
14. Thompson RA, Nelson CA. Developmental science and the media—early brain development. Am Psychol **2001**; 56:5–15.
15. The MAL-ED Network Investigators. The MAL-ED study: a multinational and multidisciplinary approach to understand the relationship between enteric pathogens, malnutrition, gut physiology, physical growth, cognitive development, and immune responses in infants and children up to 2 years of age in resource-poor environments. Clin Infect Dis **2014**; 59(suppl 4):S193–206.
16. Lamb ME, Hwang CP, Ketterlinus RD, Fracasso M. Parent-child relationships: development in the context of the family. In: Bornstein MHLamb ME, eds. Developmental psychology: an advanced textbook. 4th ed. Mahwah, NJ: Lawrence Erlbaum Associates, **1999**; 411–50.
17. Beck CT. The effects of postpartum depression on child development: a meta-analysis. Arch Psychiatr Nurs **1998**; 12:12–20.
18. Field TM, Healy B, Goldstein S, Guthertz M. Behavior-state matching and synchrony in mother-infant interactions in nondepressed versus depressed dyads. Dev Psychol **1990**; 26:7–14.
19. Armony-Sivan R, Kaplan-Estrin M, Jacobson SW, Lozoff B. Iron-deficiency anemia in infancy and mother-infant interaction during feeding. J Dev Behav Pediatr **2010**; 31:326–32.
20. Murray-Kolb LE, Beard JL. Iron deficiency and child and maternal health. Am J Clin Nutr **2009**; 89:946S–50S.
21. Lozoff B, Klein NK, Nelson EC, McClish DK, Manuel M, Chacon ME. Behavior of infants with iron-deficiency anemia. Child Dev **1998**; 69:24–36.
22. Gardner JM, Powell CA, Baker-Henningham H, Walker SP, Cole TJ, Grantham-McGregor SM. Zinc supplementation and psychosocial stimulation: effects on the development of undernourished Jamaican children. Am J Clin Nutr **2005**; 82:399–405.
23. Lozoff B, Beard J, Connor J, Barbara F, Georgieff M, Schallert T. Long-lasting neural and behavioral effects of iron deficiency in infancy. Nutr Rev **2006**; 64:S34–43.
24. Castillo-Duran C, Perales CG, Hertrampf ED, Marin VB, Rivera FA, Icaza G. Effects of zinc supplementation on development and growth of Chilean infants. J Pediatr **2001**; 138:229–35.
25. Greenfield PM. You can't take it with you: why ability assessments don't cross cultures. Am Psychol **1997**; 52:1115–24.
26. Rogler LH. Methodological sources of cultural insensitivity in mental health research. Am Psychol **1999**; 54:424–33.
27. Bayley N. Bayley scales of infant and toddler development III. San Antonio, TX: Harcourt Assessment, **2005**.
28. Hohm E, Jennen-Steinmetz C, Schmidt MH, Laucht M. Language development at ten months: predictive of language outcome and school achievement ten years later? Eur Child Adolesc Psychiatry **2007**; 16:149–56.
29. Vaughn BE, Bost KK. Attachment and temperament: redundant, independent or interacting influences on interpersonal adaptation and personality development? In: Cassidy J, Shaver PR, eds. Handbook of attachment: theory, research, and clinical applications. New York: Guilford Press, **1999**:198–225.
30. Rothbart MK, Ahadi SA, Hershey KL, Fisher P. Investigations of temperament at three to seven years: the Children's Behavior Questionnaire. Child Dev **2001**; 72:1394–408.
31. Cole PM, Michel MK, Teti LO. The development of emotion regulation and dysregulation: a clinical perspective. Monogr Soc Res Child Dev **1994**; 59:73–100.
32. Bornstein MH, Tamis-LeMonda CS, Haynes OM. First words in the second year: continuity, stability, and models of concurrent and predictive correspondence in vocabulary and verbal responsiveness across age and context. Infant Behav Dev **1999**; 22:65–85.
33. Guralnik MJ. The effectiveness of early intervention. Baltimore, MD: Brookes, **1997**.
34. Holden JM. Postnatal depression: its nature, effects, and identification using the Edinburgh Postnatal Depression Scale. Birth **1991**; 18:211–21.
35. Cooper P, Murray L. Postnatal depression. BMJ **1998**; 316:1884–86.
36. Field T, Healy B, Goldstein S, et al. Infants of depressed mothers show "depressed" behavior even with non-depressed adults. Child Dev **1988**; 59:1569–79.

37. Cox AD, Puckering C, Pound A, Mills M. The impact of maternal depression in young children. J Child Psychol Psychiatry **1987**; 28:917–28.
38. Whiffen VE, Gotlib IH. Infants of postpartum depressed mothers: temperament and cognitive status. J Abnorm Psychol **1989**; 98:274–9.
39. Stein A, Gath D, Bucher J, Bond A, Day A, Cooper P. The relationship between post-natal depression and mother-child interaction. Br J Psychiatry **1991**; 158:46–52.
40. Teti DM, Gelfand DM, Messinger DS, Isabella R. Maternal depression and the quality of early attachment: an examination of infants, preschoolers, and their mothers. Dev Psychol **1995**; 31:364–76.
41. Hunt J. Intelligence and experience. New York: Ronald Press, **1961**.
42. Plomin R, Daniels D. Why are children in the same family so different from each other? Behav Brain Sci **1987**; 10:1–16.
43. Tong S, Baghurst P, Vimpani G, McMichael A. Socioeconomic position, maternal IQ, home environment, and cognitive development. J Pediatr **2007**; 151:284–8.
44. Fagan JF, Singer LT. Infant recognition memory as a measure of intelligence. In: Lipsitt LP, ed. Advances in infancy research. Vol 2. Norwood, NJ: Ablex, **1983**:31–78.
45. Honzik MP. Measuring mental abilities in infancy: the value and limitations. In: Lewis M, ed. Origins of intelligence: infancy and early childhood. 2nd ed. New York: Plenum Press, **1983**:67–106.
46. Lewis M, Sullivan MW. Infant intelligence and its assessment. In Wolman B, ed. Handbook of intelligence: theories, measurements, and applications. New York: Wiley, **1985**: 505–99.
47. Rose SA, Feldman JF, Futterweit LR, Jankowski JJ. Continuity in tactual-visual cross-modal transfer: infancy to 11 years. Dev Psychol **1998**; 34:435–40.
48. Griffiths R. The abilities of babies. London: University of London Press, **1967**.
49. Frankenburg WK, Dodds J, Archer P, Shapiro H, Bresnick B. The Denver II: a major revision and restandardization of the Denver Developmental Screening Test. Pediatrics **1992**; 89:91–7.
50. Drachler Mde L, Marshall T, de Carvalho Leite JC. A continuous-scale measure of child development for population-based epidemiological surveys: a preliminary study using Item Response Theory for the Denver Test. Paediatr Perinat Epidemiol **2007**; 21:138–53.
51. Tofail F, Kabir Q, Hamadani JD, et al. Supplementation of fish-oil and soy-oil during pregnancy and psychomotor development of infants. J Health Popul Nutr **2006**; 24:48–56.
52. Fenson L, Dale PS, Reznick JS, et al. MacArthur Communicative Development Inventories. User's guide and technical manual. Baltimore: Paul H. Brookes, **1994**.
53. Rowe ML, Raudenbush SW, Goldin-Meadow S. The pace of vocabulary growth helps predict later vocabulary skill. Child Dev **2012**; 83:508–25.
54. Hamadani JD, Baker-Henningham H, Tofail F, Mehrin F, Huda SN, Grantham-McGregor SM. Validity and reliability of mothers' reports of language development in 1-year-old children in a large-scale survey in Bangladesh. Food Nutr Bull **2010**; 31:S198–206.
55. Gartstein MA, Rothbart MK. Studying infant temperament via the Revised Infant Behavior Questionnaire. Infant Behav Dev **2003**; 26:64–86.
56. Raven J, Raven JC, Court JH. Manual for Raven's Progressive Matrices and Vocabulary Scales. San Antonio, TX: Harcourt Assessment, **2003**.
57. Raven J. The Raven's Progressive Matrices: change and stability over culture and time. Cogn Psychol **2000**; 41:1–48.
58. Rushton JP, Skuy M, Bons TA. Construct validity of Raven's Advanced Progressive Matrices for African and non-African engineering students in South Africa. Int J Select Assess **2004**; 12:220–9.
59. Cattell RB, Cattell AKS. Culture fair intelligence test. Champaign, IL: Institute for Personality and Ability Testing, **1963**.
60. Jensen AR. Bias in mental testing. New York: Free Press, **1980**.
61. Beusenberg M, Orley JA. Users guide to the Self Reporting Questionnaire (SRQ). Geneva. Switzerland: World Health Organization, **1994**.
62. Harding TW, De Arango MV, Baltazar J, et al. Mental disorders in primary health care: a study of their frequency and diagnosis in four developing countries. Psychol Med **1980**; 10:231–41.
63. Bradley RH, Corwyn R, Whiteside-Mansell L. Life at home: same time, different places, an examination of the HOME Inventory in different cultures. Early Dev Parenting **1996**; 5:251–69.
64. Bradley RH, Corwyn RF. Caring for children around the world: a view from HOME. Int J Behav Dev **2005**; 29:468–78.
65. Caldwell B, Bradley R. Home observation for the measurement of the environment. Little Rock: University of Arkansas, **1984**.
66. Vandoorninck WJ, Caldwell BM, Wright C, Frankenburg WK. The relationship between 12-month home stimulation and school-achievement. Child Dev **1981**; 52:1080–3.
67. Molfese VJ, Modglin A, Molfese DL. The role of environment in the development of reading skills: a longitudinal study of preschool and school-age measures. J Learn Disabil **2003**; 36:59–67.
68. Black MM, Baqui AH, Zaman K, et al. Iron and zinc supplementation promote motor development and exploratory behavior among Bangladeshi infants. Am J Clin Nutr **2004**; 80:903–10.
69. Solarsh B, Alant E. The challenge of cross-cultural assessment—the Test of Ability to Explain for Zulu-Speaking Children. J. Commun Disord **2006**; 39:109–38.
70. American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. Standards for educational and psychological testing. Washington, DC: American Educational Research Association, **1999**.
71. Fernandez K, Boccaccini MT, Noland RM. Professionally responsible test selection for Spanish-speaking clients: a four-step approach for identifying and selecting translated tests. Prof Psych Res Practice **2007**; 38:363–74.
72. 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**; 80:31–53.
73. Hambleton RK, Kanjee A. Increasing the validity of cross-cultural assessments: use of improved methods for test adaptations. Eur J Psychol Assess **1995**; 11:147–57.
74. 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 years of life. Washington, DC: World Bank, **2009**.
75. Prince M. Measurement validity in cross-cultural comparative research. Epidemiol Psichiatr Soc **2008**; 17:211–20.