

**Adaptation of the Mullen Scales of Early Learning for use among infants aged 5-  
24-months in rural Gambia**

Bosiljka Milosavljevic<sup>1\*</sup>, Perijne Vellekoop<sup>2\*</sup>, Helen Maris<sup>1</sup>, Drew Halliday<sup>3</sup>, Saikou  
Drammeh<sup>4</sup>, Lamin Sanyang<sup>4</sup>, Momodou K. Darboe<sup>4</sup>, Clare Elwell<sup>5</sup>, Sophie E. Moore<sup>6</sup> &  
Sarah Lloyd-Fox<sup>1</sup>

<sup>1</sup>Centre for Brain and Cognitive Development, Birkbeck, University of London

<sup>2</sup>The Global Health and Tropical Medicine Training Institute, The Netherlands

<sup>3</sup>Department of Psychology, University of Victoria

<sup>4</sup>Medical Research Council, The Gambia, at The London School of Hygiene and Tropical  
Medicine

<sup>5</sup>Department of Medical Physics and Biomedical Engineering, University College London

<sup>6</sup>Department of Women and Children's Health, King's College London

\*Joint first authors

**Corresponding Author:**

Bosiljka Milosavljevic  
Centre for Brain and Cognitive Development, Birkbeck, University of London  
Malet Street, London WC1E 7HX

Tel: +44 207 079 0863

Email: b.milosavljevic@bbk.ac.uk

**Conflict of interest:** The authors confirm that they have no conflict of interest

**Acknowledgements:**

We would like to thank the parents and infants who took part in this study as well as the field staff at the MRC Keneba Field Station without whom this work would not have been possible. This study was supported by Bill & Melinda Gates Foundation Grants OPP1061089 and OPP1127625, and core funding MC-A760-5QX00 to the International Nutrition Group by the Medical Research Council UK and the UK Department for International Development (DfID) under the MRC/DfID COncordant agreement.

**Research Highlights:**

- A population-specific adaptation of the Mullen Scales of Early Learning (MSEL) developed for use and tested in a rural area of The Gambia
- Infants in this setting had lower MSEL scores compared to US norms and performance declined from the ages of 5 to 24 months
- Reduced performance was observed across multiple domains of cognitive development, including perceptual, motor and language scales
- Performance on the Mullen Scales of Early Learning was significantly associated with measures of physical growth

### Abstract

Infants in low-resource settings are at heightened risk for compromised cognitive development due to a multitude of environmental insults in their surroundings. However, the onset of adverse outcomes and trajectory of cognitive development in these settings is not well understood. The aims of the present study were to adapt the Mullen Scales of Early Learning (MSEL) for use with infants in a rural area of The Gambia, to examine cognitive development in the first 24-months of life and to assess the association between cognitive performance and physical growth. In phase 1 of this study, the adapted MSEL was tested on 52 infants aged 9-24 months (some of whom were tested longitudinally at two time points). Further optimization and training were undertaken and phase 2 of the study was conducted, where the original measures were administered to 119 newly recruited infants aged 5-24 months. Infant length, weight and head circumference were measured concurrently in both phases. Participants from both phases were split into age categories of 5-9m ( $N=32$ ), 10-14m ( $N=92$ ), 15-19m ( $N=53$ ) and 20-24m ( $N=43$ ) and performance was compared across age groups. From the age of 10-14m, Gambian infants obtained lower MSEL scores than US norms. Performance decreased with age and was lowest in the 20-24m old group. Differential onsets of reduced performance were observed in the individual MSEL domains, with declines in visual perception and motor performance detected as early at 10-14 months, while reduced language scores became evident after 15-19 months of age. Performance on the MSEL was significantly associated with measures of growth.

**Keywords:** Global Mental Health; Cognitive Development; Infancy; Growth; Mullen Scales of Early Learning; The Gambia

The United Nations Sustainable Development Goals (UN, 2015) have identified the reduction of poor child development in low- and middle-income countries (LMICs) as a key priority for global health research and interventions. A particular area of concern is compromised cognitive development within these settings. Grantham-McGregor et al. (2007) suggest that over 200 million children worldwide fail to meet their cognitive potential due to poverty, undernutrition and inadequate care. This has significant impact on individuals' quality of life, as well as the nation's workforce and economic development (Alderman, Behrman, Grantham-McGregor, Lopez-Boo, & Urzua, 2014). Multiple interventions have attempted to improve cognitive outcomes of children who are at-risk, which have had varied effectiveness (for review see Perkins et al., 2017). Discrepancies are likely due to outstanding questions about the nature and trajectory of compromised cognitive development in LMICs. The age at which interventions are most effective and domains of cognitive functioning most at-risk are still under investigation (Perkins et al., 2017). Moreover, the appropriateness of using established measures of cognitive development in these settings is uncertain (Isaacs & Oates, 2008).

The present study aims to address these issues by adapting a widely used measure of child development, the Mullen Scales of Early Learning (MSEL; Mullen, 1995), for use with infants in a rural area of The Gambia, West Africa. The objective is to examine the applicability and challenges of using a population-specific adaptation of the MSEL within this context. Furthermore, we examine age-related trajectories of cognitive development within the first 24-months of life in this setting to identify key periods of vulnerability, across multiple domains. We also assess the association between cognitive performance and measures of growth.

### *Cognitive Development in Infancy and Periods of Sensitivity*

Within the first 24-months of life, neuronal networks develop most rapidly and exhibit high plasticity (Fox, Calkins, & Bell, 1994; Fox, Levitt, & Nelson, 2010; Kolb & Gibb, 2011). Environmental factors in early life, such as home environment, caregiver responsiveness, infection and diet, have important and lasting impacts on neural development (e.g. Murray-Kolb et al., 2014). Consequently, the infant brain is highly vulnerable to compromised development during this time. Yet, it is also during this period that interventions may be most effective (Kolb & Gibb, 2011). Thus, many interventions targeted at reducing developmental delay are administered in the first years of life, with the aim of minimising the impact of adverse factors before poorer developmental outcomes are fully manifest (Petrenko, 2013).

It is, however, important to note that the ontogenetic events of neural maturation occur at different periods during this early time frame and that they have cascading effects (Grantham-McGregor et al., 2007; Thompson & Nelson, 2001). For example, visual reception begins to develop at birth and by approximately 6-months of age, infants exhibit similar patterns of visual preference and visuospatial organisation as adults (Deen et al., 2017; Emberson, Richards, & Aslin, 2015). On the other hand, advances in communicative development begin to emerge later and are most prominent by 12-months of age (e.g. Levine, Strother-Garcia, Golinkoff, & Hirsh-Pasek, 2016). Yet, evidence suggests that infants' early perceptual abilities influence the development of speech perception and production later on (Bruderer, Danielson, Kandhadai, & Werker, 2015). Similarly, healthy development of motor abilities is posited to assist infants in engaging with their environment, thus providing the opportunity to hone other cognitive skills (Iverson, 2010). Among other populations at-risk for cognitive delay, such as infants at familial risk for Autism Spectrum Disorder (ASD), delays in sensorimotor development emerge very early in infancy and precede atypicalities in language performance (Leonard et al., 2014; Nishimura, Takei, Tsuchiya, Asano, & Mori,

2016). Thus, we argue that it is insufficient to solely measure global cognitive development, or development during later childhood only, when trying to identify a key period of developmental risk. Rather, it is important to track age-related changes in performance across specific perceptual, motor and language domains to observe more fine-grained periods when development may be sensitive to insults.

### ***Measuring Infant Development in LMICs***

There is an emerging body of research that has investigated cognitive development in LMICs. This research has provided important insights into how environmental risk factors such as undernutrition, lack of stimulation, poverty and disease contribute to compromised cognitive development (Fernald, Kariger, Hidrobo, & Gertler, 2012; Fernald et al., 2006; McCoy et al., 2016). However, there is considerable variability in the types of cognitive assessment methods used, outcomes measured, and age of children tested in the studies. Many studies use school performance as a measure of cognitive outcome (Fernald, Kariger, Hidrobo, & Gertler, 2012). Others use standardised measures of language and IQ (e.g. Schady et al., 2015), which provide useful information but are not suitable for young, pre-verbal children. Thus, if studies are to measure cognitive development during early life, assessment methods suitable for this age need to be identified.

The MSEL and the Bayley Scales of Infant and Toddler Development (Bayley, 1995) are key measures used to assess child developmental level in the first 5 and 3 years of life, respectively. These assessments have been used widely in research and measure development across perceptual, motor and language domains. A key limitation of these measures is that they have been designed for and normed in Western, English speaking populations. Consequently, their utility in LMICs is limited without adaptation to the specific target population and language (Isaacs & Oates, 2008). Several such adaptations already exist,

providing a guide for future researchers attempting to use infant development measures cross culturally.

Hamadani et al. (2014) measured the association between poverty and infant cognitive development in Bangladesh using the Bayley scales, among other measures, at 7-, 18- and 64-months of age. Their study reports that children from the lowest socioeconomic status (SES) group in the sample exhibited reduced cognitive performance from the age of 7-months. Furthermore, the cognitive deficit observed in this group increased with age. One limitation of this study is that it did not use the same measure at all time points (Bayley scales were only used at 18-months), leaving unclear the effectiveness and applicability using a single measure of cognitive ability for longitudinal follow-up. Koura et al. (2013) adapted the MSEL for use with 12-month old infants in Benin and found that a substantial proportion of children at this age exhibited developmental delays. Scores from the MSEL were also significantly associated with parent-report measures of infant development, suggesting good validity of this adapted scale. Bodeau-Livinec et al. (2018) utilised the MSEL that had been adapted for Benin (Koura et al., 2013) and report reasonable construct validity for both raw and standardised scores among children aged 3-6 years in this setting. Likewise, cognitive performance was significantly associated with measures of risk factors (e.g. maternal depression).

Multiple studies have adapted the MSEL and Bayley scales for use in South Africa, all reporting satisfactory utility of the scales within this setting. Ballot et al. (2017) used the Bayley scales in a longitudinal study and reports that infants exhibited a decline in language scores after the age of 12-months. In contrast, no change in global cognitive performance was detected, highlighting the importance of examining development across several age points and cognitive domains. Bornman, Sevcik, Ronski & Kyeong Pae (2010) describe in detail an adaptation of the MSEL for use with Afrikaans-speaking children in South Africa. They

report that standardised scores on the MSEL corresponded with parent-report of developmental milestones, but that no age-related differences in these scores were observed among children aged 3-6 years. Bornman et al. (2018) assessed age-related changes in MSEL performance in a cross-sectional study with children aged 21-68 months. Their study reports that MSEL scores increased with age, but it is important to note that only raw scores (and not scores corrected for age) were examined, which are expected to increase since older children complete more items. Finally, Boivin (2013) reports increases in MSEL performance from the age of 6- to 18-months after a parent-led intervention, among Ugandan infants exposed to HIV. A number of these studies also report sex differences, with girls scoring higher than boys, particularly in language subscales (e.g. Koura et al., 2013).

Taken together, these studies suggest that population-specific adaptations of key developmental assessments can be applied in LMICs. They demonstrate appropriateness for use in the contexts that they were designed for, are sensitive to changes in performance with age, and are able to detect improvement after intervention. However, Perkins et al. (2017) warn that many adaptations are designed to be suitable among children of higher SES. This is evident in some of the studies described above, where the measures were translated into the country's official European language (e.g. French in Benin, Koura et al., 2013) while a large proportion of the population in these countries have a local language as their first language (Wright, 2016). Many of these adaptations have been tested in urban or semi-urban settings, while children most highly at-risk are likely to be living in rural areas (Baulch, 2011). Moreover, these studies did not describe in detail the challenges of tailoring and administering the adapted assessments in their respective settings. This would have been highly useful information for future research aiming to adapt infant behavioural measures for similar contexts.

*Association between cognitive development and growth*



Nutritional intake both pre- and postnatally is posited to have important consequences for the structural development of the infant brain, which, in turn, lays the foundations for cognitive development (Benton, 2008; Nurliyana, Mohd Shariff, Mohd Taib, Gan, & Tan, 2016). Human autopsy studies and animal models support this hypothesis, suggesting that undernutrition is associated with compromised neural composition in both infants and adults (Prado & Dewey, 2014). Furthermore, deficiencies in specific micronutrients and malnutrition, prenatally and in infancy, are associated with reduced cognitive performance and developmental delays (Nyaradi, Li, Hickling, Foster, & Oddy, 2013).

There is a growing literature that directly examines the relationship between physical growth (an indicator of nutritional status) and cognitive development in young children in LMICs. Using a meta-analytic approach, Sudfeld et al. (2015) examined the association between pooled adjusted standardised mean performance in cognitive ability and infant length. Among children under the age of 2 years, a single unit increase in length was associated with significantly higher cognitive performance. This relationship was also observed among children older than 2 years, but to a lesser extent. This study also reports that, longitudinally, a unit increase in length under the age of 2 years is associated with improved cognitive ability at the age of 5-11 years.

The impact that nutritional interventions and “catch-up” growth have on the reversibility of insults to cognitive development remain under investigation (Perkins et al., 2017). Numerous reports suggest that interventions involving early nutritional supplementation have shown beneficial effects on both physical growth and cognitive performance (Stewart et al., 2018; Tofail et al., 2018). Nutritional interventions have been shown to be most effective at minimising the impact of insults to cognitive development when administered during the first two years of life (Ip et al., 2017). However, there is vast heterogeneity in this domain, since other studies concluded that improving nutritional intake

does not result in better cognitive performance (Sokolovic et al., 2014). Perkins et al. (2017) argue that these discrepant findings could be due to heterogeneity in the type of intervention administered, child sex and family demographics. Furthermore, these associative findings are largely drawn by examining early life growth measures and later pre-school or school aged cognitive markers. The parallel investigation of physical growth and cognitive development during early childhood would enable us to better disentangle the relative impact that early-life risk factors have on development.

### ***Context and Setting of Present Study***

The current study preceded, and contributed to, the development of a large scale prospective longitudinal study now underway in The Gambia to measure early brain and cognitive development across the first two years of life: The Brain Imaging for Global Health (BRIGHT) Study – [www.globalfnirs.org/the-bright-project](http://www.globalfnirs.org/the-bright-project). This project aims to implement brain imaging and neurocognitive developmental methods (including functional Near Infra-Red Spectroscopy (fNIRS), electroencephalography (EEG) and MSEL) to model longitudinal changes in brain function and cognitive development, identify critical time points, and moderators and mediators of compromised development within this rural Gambian population. In preparation for this large-scale study, multiple feasibility and pilot studies have been conducted combining neuroimaging (fNIRS, EEG), behavioural (MSEL) and growth measures in a longitudinal and cross-sectional design. Results of the early fNIRS studies have already been published (Begus et al., 2016; Lloyd-Fox et al., 2017, 2014; Papademetriou et al., 2014). The present study describes the outcomes of a series of pilot studies using the key cognitive development behavioural assessment of the BRIGHT study - the MSEL.

This research was undertaken at the Medical Research Council (MRC) field station in the village of Keneba, in the rural West Kiang region of The Gambia ([www.mrc.gm](http://www.mrc.gm); [www.ing.mrc.ac.uk](http://www.ing.mrc.ac.uk)). The community in this region relies predominantly on subsistence

farming and, thus, eating patterns and income vary greatly between the annual wet and dry seasons (Hennig et al., 2015; van der Merwe et al., 2013). Additionally, there is a high prevalence of infectious disease (van der Merwe et al., 2013). Taken together, these factors pose important threats to healthy development, which is reflected in the finding that a majority of the children in this setting exhibit moderately severe growth faltering from the age of 3 months (van der Merwe et al., 2013).

### *Aims and Hypotheses*

The aims of this study were to assess the applicability of an environmentally and linguistically adapted MSEL for use among infants aged 5- to 24-months in this setting. Likewise, we sought to examine cognitive development and age-related changes in performance on the MSEL to identify crucial periods of sensitivity in both overall cognitive performance and on individual visual, motor and language domains. Finally, we examined the association between MSEL scores and measures of growth. Specifically, we aimed to test the following hypotheses:

1. Given prior findings that early life exposure to environmental risk is associated with reduced cognitive ability from an early age, and that performance worsens with age (e.g. Hamadani et al., 2014), we predict that our sample will exhibit reduced MSEL scores (compared to US norms) from the age of 5-months. We also expect that performance will worsen with increasing age.
2. Furthermore, since different domains of cognition develop at different periods within the first 24-months (Thompson & Nelson, 2001), we expect that reduced performance will be apparent at different ages for individual domains. Since, perceptual and motor skills develop very early in infancy, we predict that declines in visual reception and motor domains will become apparent first, followed by declines in language development.

3. Finally, growth in early life has been associated with long-term cognitive ability (e.g. Sudfeld et al., 2015) thus we predict that measures of growth will be positively associated with cognitive development in our sample. Given that reduced cognitive performance and growth faltering become more apparent with age (Perkins et al., 2017), we also predict that the strength of the association between cognitive performance and growth will increase with age.

## **Method**

### ***Participants***

Two phases of testing were undertaken between February 2013 and April 2014 with a total of 171 children born in the West Kiang region. A total of 192 infants were recruited for the study but not all completed the assessments (see results).

According to The Gambia Bureau of Statistics (2011), a majority of the population in this setting live below the poverty line, with an average income of below \$2/day. Literacy rates, particularly among women, are low but increasing numbers of children attend primary school. However, fewer than half transition to secondary education and there is minimal opportunity to attend preschool. A typical diet consists of staple foods (e.g. white rice, millet or maize) with limited additional ingredients, such as sauce made from ground nuts, vegetable or palm oil, fish and vegetables (Dominguez-Salas et al., 2013). Rates of exclusive breastfeeding are high, and infants are predominantly breastfed until the age of 24-months (Eriksen et al., 2017). Weaning foods are commonly nutritionally insufficient, leading to widespread growth faltering (Nabwera, Fulford, Moore, & Prentice, 2017).

The residents of the West Kiang region have access to primary care services, provided for by the MRC, as well as local healthcare facilities run by the Gambian Ministry of Health (Hennig et al., 2015; Rayco-Solon, Moore, Fulford, & Prentice, 2004). The MRC provides a wide range of nutrition-specific interventions, antenatal and child care. These have led to a

significant decrease in infant mortality, but undernutrition remains highly prevalent (Nabwera, Fulford, Moore, & Prentice, 2017). While English is the official language in The Gambia, almost 80% of the population in the West Kiang region have Mandinka as their primary language (Hennig et al., 2015). All infants in this study came from households where Mandinka was spoken.

*Phase 1:* Between February and November 2013, 66 infants (32 female, 34 male) aged 9- to 24-months were recruited using the West Kiang Demographics Surveillance System (Hennig et al., 2015). All were born full-term (37+ weeks gestation) and had normal birth weight. Within this phase, there were two groups of participants. The first group consisted of 42 infants (18 female, 24 male) who were tested longitudinally. Baseline was at age 9- to 17-months and the follow-up was conducted 3 months later, when they were aged 12- to 20-months. The second group consisted of 24 infants (14 female, 10 male), tested cross-sectionally between the ages of 18- to 24-months.

*Phase 2:* Between February and April 2014, 126 infants (68 female, 58 male) aged 5- to 24-months were recruited (by random selection) from the Early Nutrition and Immune Development trial (ENID; Moore et al., 2012). These children from the ENID cohort, were a subset of the full cohort, and were included in this phase if they reached the age of 6-, 12-, 18- or 24-months (+/- 1 month) during the time of testing.

For both phases, ethics approval was granted by the joint Gambia Government/MRC Unit The Gambia Ethics Committee. Written informed consent for participation and further use of videotapes was obtained from parents.

### ***Demographic Characteristics***

Basic demographic data was available on all participants, collected as part of the West Kiang Demographic Surveillance System (Hennig et al., 2015). The ENID trial (Moore et al., 2012) collected more detailed socio-demographic data from all participants aged 12-months

or older. This included maternal and paternal age at the time of the infant's birth, whether parents attended school and the duration (in years) spent in education, family annual household income, and number of full siblings (with the same mother and father as target infant). For the present study, we obtained this ENID-collected demographic data for all participants recruited from that study (in phase 2), who were aged 12+ months at the time of testing. Other than parental age at time of birth, there were no socio-demographic data available for participants who were younger than 12-months in phase 2 and socio-demographic information was not collected for participants tested in phase 1.

### ***Measures of Cognitive Development***

*The Mullen Scales of Early Learning* (MSEL; Mullen, 1995) is an assessment battery that measures child developmental level from birth to 68-months of age. The MSEL consists of 5 subscales that measure cognitive development across the domains of Visual Reception (VR), Fine Motor (FM) skills, Gross Motor (GM) skills, Receptive Language (RL) and Expressive Language (EL). Each subscale consists of a set of performance-based items presented in hierarchical order of difficulty, where children are rated on whether they successfully complete the task in each item. Raw scores on each subscale can be converted to age-normed t-scores ( $M=50$ ,  $SD=10$ ), based on a U.S. sample. Furthermore, a global cognitive t-score ( $M=200$ ,  $SD=30$ ; hereafter "Cog T") can be computed by summing t-scores from the VR, FM, RL and EL subscales. While GM scores are not included in the Cog-T score and are not considered to constitute cognitive development, we included this scale in analyses in order to examine performance on the complete MSEL. The MSEL also provides descriptive categories that correspond to overall scores; *Very High* (Cog T  $\geq 260$ ), *Above Average* (Cog T = 231-259), *Average* (Cog T = 169-230), *Below Average* (Cog T = 138-168) and *Very Low* (Cog T  $\leq 137$ ). The original MSEL studies (Mullen, 1995) report that the measure has good psychometric properties across the different subscales and total score. For example, the 5

subscales scales exhibit satisfactory internal consistency, with median values across age groups ranging from  $r=.75$  to  $r=.83$ . Similarly, the scales had high test re-test reliability, with scores ranging from  $r=.782$  to  $r=.85$  for the cognitive scales and  $r=.96$  for the GM scale.

### ***Linguistic Translation and Adaptation of the MSEL***

Adaptation of the MSEL was undertaken with the help of a panel that consisted of the co-investigator and local principal investigator (MKD) and two senior MRC field workers (SD, LS), who are all native Mandinka speakers. Specific details about the adaptations are provided in the supplementary material.

#### ***Linguistic Translation***

MSEL instructions, item cues and language scales were translated into Mandinka. Peña (2007) provides guidelines for translating measures into different languages and suggests that, in addition to undertaking a pure linguistic translation, functional, cultural and metric equivalence of the translated items need to be considered. In this way, researchers can ensure that the translated measure captures subtleties in meaning, while maintaining an equivalent level of difficulty.

Given that Mandinka is not a written language and is, thus, without an official dictionary (e.g. Gamble, 1987), there can be substantial variability in both spoken and written Mandinka. Therefore, translation was undertaken in collaboration between the 3 panel members and a consensus translation was reached for each item. Subsequently, an additional, Mandinka speaking, field assistant, reviewed the translated MSEL to assess accuracy and word choice. Where discrepancies were noted, the field assistant provided the panel with suggestions for alternatives. The panel reviewed these suggestions and made necessary adjustments to develop an updated translation. Finally, the updated version underwent blind-back translation to English by a field assistant who was not involved in the original translation. The back translated version was compared to the original English MSEL by the

panel to assess accuracy. Discrepancies were discussed among the panel members and resolved by committee consensus.

As outlined by Peña (2007), the linguistic and cultural differences between English and Mandinka did not allow for pure linguistic equivalence in translation. Instead, the aim was to achieve functional and cultural equivalence. The phrasing of item cues was altered to adhere to local speaking patterns. When changing the phrasing, care was taken to ensure that the intended outcome of the instructions remained the same. For example, some items ask the participant to identify a toy baby doll by telling them “pat baby”, in Mandinka we asked the children to “take baby” because requesting a child to pat a doll is less common in that setting. Due to the stark environmental differences, it was not always possible to achieve cultural equivalence and certain stimuli had to be replaced (see below). Finally, efforts were made to maintain metric equivalence of the two measures by ensuring that the adapted material had words and instructions that had the similar difficulty, length, frequency of use and level of guidance as in the English-language version.

#### *Substitution of culturally inappropriate stimuli and materials*

Toys and stimuli that were deemed inappropriate or unfamiliar to children in this setting were identified. These included items that were not commonly observed (e.g. a TV) or were visually different (e.g. a western style house) in the rural Gambian context, or were objects that children were reluctant to engage with (e.g. a light-skinned doll). When selecting replacements for these stimuli, efforts were made to maintain the category of the item (e.g. household object, animal). The supplementary material describes the adaptations in full detail.

#### *Training*

Field workers, already familiar with gross and fine motor assessments of infants and toddlers, were trained in MSEL administration by experienced researchers from the UK



(SLF, PV, HM). Training consisted of watching videos and live demonstrations of item administration and discussion of scoring criteria. Trainees were given opportunity to rehearse administration with each other first. Subsequently, they administered the assessment to child participants. Supervision was provided both by their peers in the training group and supervisors. Testing sessions were video-recorded, and trainees were given in-depth feedback on their administration and scoring. A second phase of refresher training was conducted prior to data collection for phase 2.

### ***Anthropometric Measures***

Measurement of length, weight and head circumference (HC) was performed on all infants. Measurements were taken by trained fieldworkers using calibrated tools. Length was measured using a Harpenden Infantometer length board (Holtain Ltd) to a precision of 0.1cm. Weight was obtained using an electronic baby scale (model 336, SECA) to a precision of 0.01kg. Finally, HC was measured around the maximum circumference of the head (forehead to occiput) using stretch-proof measuring tape (model 201, SECA) to the nearest 0.1cm. Each measure was taken in triplicate and the mean of the 3 measures was used in analyses.

Anthropometric measures were converted to age and sex adjusted z-scores that are based on World Health Organization normative growth data (WHO Multicentre Growth Reference Study Group, 2006). Height-for-Age (HAZ), Weight-for-Height (WHZ) and Head Circumference (HCZ) z-scores were computed. Children categorised as “stunted” or “wasted” were identified using WHO criteria based on HAZ and WHZ scores, respectively. Severity of stunting/wasting is categorised as -2SD for moderate and -3SD for severe.

### ***Procedure and Optimization***

#### ***Testing Procedure***

Children were tested in the morning in a quiet room. Breakfast was provided for all participants to minimise risk of hunger or fatigue. In the case of infant refusal or disinterest in

the task, testing was interrupted and continued later in the morning. Caregivers were present during all testing sessions and were asked to positively encourage their children during testing if they were hesitant to interact with the administrator.

Each participant was tested by a single field assistant, who underwent the aforementioned training, and all sessions were video recorded. Additionally, one of the researchers from the UK (SLF, HM, PV), who are all trained on the MSEL and have 3-5 years of experience in administration, were present during every testing session. Scoring of the assessments was done concurrently by both the administrator and the researcher present. Consensus on the scores was compared after administration and, in the case of incongruent scoring between assessors, tapes were reviewed, and mutual agreement was reached. Assessors were also instructed to record all items where the infant refused to respond and all items that were scored based on parent-report rather than task administration. A similar scoring method, whereby the assessment is concurrently scored by an examiner and a trained observer, has been applied in a variety of research using behavioural measures with infants (e.g. Gammer et al., 2015).

### *Optimization*

After phase 1 of data collection, it became apparent that a substantial number of items were completed using parent-report for items where this was not a valid response option. Thus, the administrators underwent additional training, to reduce parental report before commencing phase 2 of testing. Additionally, the children's reactions to the adapted stimuli in phase 1 were reviewed and any stimuli or toys that were found to be inappropriate during testing were re-evaluated and replaced by more suitable ones for use in phase 2.

### *Analyses*

Participants from phases 1 and 2 were analysed as one combined group and divided into 4 age categories; 5-9 months, 10-14 months, 15-19 months, and 20-24 months. Data

collected from participants who were tested longitudinally was included separately in the age category that corresponded with their age at the baseline and follow-up visits.

The proportion of males and females in each age category was evaluated using Chi-Square analyses. Furthermore, given prior reports of sex differences in cognitive assessments (e.g. Koura et al., 2013), we tested for sex differences in our sample using Analysis of Variance (ANOVA) to compare Cog T and all subscale t-scores across sex. Significant sex differences emerged (see results), thus sex was controlled for in all further analyses.

Socio-demographic characteristics obtained from the ENID trial (Moore et al., 2012) in phase 2 were compared between the age groups using ANOVA and Chi-square, where appropriate.

Likewise, differences in Cog T and subscale t-scores in the two phases of testing were compared, using ANOVA. Again, significant differences emerged between the two phases (see results) and phase was also controlled for in further analyses.

The number of items completed by parent-report was compared between the two phases. Given that participants do not always complete the same number of items on the MSEL, a ratio of items scored through parent-report out of the total number of items completed was computed. These scores then underwent arcsine transformation and were compared between phases 1 and 2 using the Mann-Whitney U test. To determine whether parent-report had an impact on the total score, Spearman Rho correlation was run between number of items scored by parent-report and Cog T scores.

#### *Comparison of Gambian MSEL scores and US Norms*

To assess the first hypothesis, that the Gambian sample would have reduced MSEL scores, one-sample t-tests were used to compare Cog T scores from our sample with US-normed Cog T scores ( $M=200$ ,  $SD=30$ ). These were performed separately for each age group and Bonferroni adjusted  $p$ -value ( $.05/4=.0125$ ) was used to account for multiple testing.

Likewise, we compared the number of participants who were in the different descriptive categories in each age group. To reduce the number of tests, the descriptive categories were collapsed to create categories that comprise normative scores (Very High, Above Average and Average; hereafter '*Average and Above*') and those that indicated delay (Below Average and Very Low, hereafter '*Below Average*'). The proportion of participants that fell into these categories in each age group was assessed using Chi-Square analysis. Post hoc tests were performed following the method suggested by Beasley and Schumacker (1995), where the adjusted standardized residuals were analysed and converted to  $p$ -values. Bonferroni adjusted  $p$ -values ( $.05/4=.001$ ) were used to account for multiple testing.

#### *Age-Related Changes in MSEL Performance*

To assess for age-related changes in MSEL scores, Multivariate ANOVA (MANOVA) was performed to compare Cog T and subscale t-scores across the 4 age groups, controlling for phase of testing and sex. Where significant differences emerged, pairwise comparisons were used to identify age groups that significantly differed from each other in each of the scales. Bonferroni correction was applied to correct for multiple testing.

To further examine age-related trajectories in MSEL scores, additional analyses were performed in the group that was tested longitudinally. Repeated Measures ANOVA was used to compare Cog T and all subscale t-scores between the baseline and follow-up visits, controlling for infant sex.

We also examined change in descriptive category between the two visits in the longitudinal group. McNemar's test was used to compare the proportion of participants that had shifted between the *Average and Above* and *Below Average* categories from baseline to follow-up. When a significant result emerged, post-hoc analyses were done by performing McNemar's test in the *Average and Above* and *Below Average* categories separately, with Bonferroni adjusted  $p$ -values ( $.05/2=.01$ ) to account for multiple testing.

*Association between MSEL Scores and Measures of Growth*

Age and sex differences in the growth measures were assessed first. MANOVA was used to compare mean WHZ, HAZ, and HCZ across the 4 age groups and between males and females. Where significant differences emerged, pairwise comparisons were performed to identify which groups significantly differed from each other. Bonferroni correction was applied to account for multiple testing.

Subsequently, we sought to examine the association between Cog T scores WHZ, HAZ and HCZ. Spearman Rho correlations were performed between Cog T scores and each of the growth z-scores to identify significant associations. Significant correlations emerged between Cog T and WHZ and HAZ scores (see results). These correlations were followed up by linear regression to determine whether the associations would remain when taking into account participant age and if the strength of these associations increased as children became older. Two separate regressions were run with WHZ and HAZ as predictors in each. Cog T scores were entered as the dependant variable, WHZ/HAZ, age and an interaction term between age\*WHZ/ age\*HAZ were entered as predictors. WHZ, HAZ and age were mean centred to reduce the risk of multicollinearity. Furthermore, sex and phase of testing were controlled for in both regressions.

All analyses were performed in SPSS Version 24 (IBM Corp, 2016). Effect sizes are reported as Cohen's  $d$ ,  $\eta^2$ , and  $r^2$ , where appropriate.

## **Results**

### ***Demographic Characteristics***

Table 1 summarises the ages and sex ratio of the entire sample by age group, as well as within phases. From the overall sample, 21 participants did not complete the MSEL due to refusal to participate on the day of testing. Phase 1 had data missing from 13 participants, 8 of these were among participants tested longitudinally (5 missed the baseline visit, 2 missed

the follow-up and 1 did not complete either visit) and 5 participants who were tested cross-sectionally refused to participate. In Phase 2, only 8 participants did not complete the MSEL. There were no significant differences in the proportion of males and females in Phases 1 and 2 ( $X^2=1.36, p=.24$ ) or across the 4 age groups ( $X^2(3)=5.61, p=.13$ ).

In phase 2 (see Table 2), there were significant differences in maternal age between the 4 age groups, with mothers of infants in the 20-24m group being older when then infant was born than mothers of infants in both the 5-9m ( $p<.001, d=1.09$ ) and 10-14m ( $p=.02, d=.83$ ) groups. Likewise, fathers of infants in the 20-24m group were older when the infant was born than were fathers in the 5-9m group ( $p=.02, d=1.01$ ). Additionally, infants in the 20-24m group had more siblings than those in the 10-14m group ( $p=.04, d=.98$ ). There were also trend-level differences between the age groups in the number of fathers who attended school ( $X^2(2)=5.24, p=.07$ ). There were no significant differences in any of the other socio-demographic variables.

Significant sex differences were observed in the Cog T score ( $F(1, 207)=7.41, p=.01, d=.36$ ) and the GM ( $F(1, 207)=8.73, p=.003, d=.41$ ), VR ( $F(1, 207)=5.42, p=.01, d=.32$ ), and FM ( $F(1, 207)=8.18, p=.01, d=.40$ ) scales. In each of these subscales males outperformed females (see Table 3 for summary of scores). Furthermore, males scored higher than females in the RL scale, but this only reached trend-level significance ( $F(1, 207)=3.15, p=.08, d=.24$ ). Because of these significant sex differences, sex was controlled for in further analyses of MSEL scores.

Significant differences were also observed between the two Phases of testing in the Cog T score ( $F(1, 207)=6.35, p=.01, d=.35$ ), the GM scale ( $F(1, 207)=16.32, p<.001, d=.55$ ), the FM scale ( $F(1, 207)=16.98, p<.001, d=.56$ ) and the EL scale ( $F(1, 207)=4.53, p=.03, d=.30$ ). Scores were higher in Phase 1 for every scale where a significant difference was detected (Table 3). Consequently, Phase was controlled for in all further analyses.

In Phase 1, an average of 17.2% ( $SD=16.1$ ) of items were completed using parent-report, while in phase 2 this decreased to 3.6% ( $SD=2.98$ ). This decrease was statistically significant,  $U=1282.50$ ,  $p<.001$ . On the other hand, the number of items that were completed using parent-report was not significantly associated with Cog T scores,  $r_s(209)=-.06$ ,  $p=.43$ .

### ***Comparison of Gambian MSEL scores and US Norms***

At 5-9m there was no significant difference between the Cog T scores obtained by the Gambian sample and the US norms ( $t(31)=-2.20$ ,  $p=.04$ ,  $d=.33$ ), with a Bonferonni-corrected  $p$ -value applied. However, at the ages of 10-14m ( $t(89)=-5.69$ ,  $p<.001$ ,  $d=.60$ ), 15-19m ( $t(47)=-8.74$ ,  $p<.001$ ,  $d=1.23$ ) and 20-24m ( $t(38)=-9.96$ ,  $p<.001$ ,  $d=1.55$ ) the scores obtained from the Gambian infants were significantly lower than the US norms (see Figure 1 for summary of mean scores at each age point and Figure 2 for individual participant scores) .

Furthermore, examining the proportion of participants that were *Average and Above* and *Below Average* revealed significant differences between the age groups. At 5-9 months, there was a significantly greater proportion of participants scoring *Average and Above* (84.38%) than *Below Average* (15.63%),  $p<.001$ . At 10-14 months, a larger number of participants scored *Average or Above* (66.67%) than *Below Average* (33.33), however this did not reach significance after Bonferroni correction ( $p=.02$ ). By 15-19 months, a significantly greater portion of participants (58.33%) scored *Below Average* than those scoring *Average or Above* (41.67%),  $p=.01$ . At 20-24 months, this pattern was similar, with a significantly greater percentage of participants scoring *Below Average* (66.67%) compared with those who scored *Average or Above* (33.33%),  $p<.001$ .

### ***Age-Related Changes in MSEL Performance***

Table 4 shows a summary of the MSEL scores in each age group. After controlling for sex and phase of testing, there were significant differences in the Cog T ( $F(1, 203)=12.40$ ,  $p<.001$ ,  $\eta^2=.16$ ), EL ( $F(1, 203)=5.95$ ,  $p=.001$ ,  $\eta^2=.08$ ), RL ( $F(1, 203)=3.86$ ,  $p=.01$ ,  $\eta^2=.05$ ),

FM ( $F(1, 203)=8.92, p<.001, \eta^2=.12$ ) and VR ( $F(1, 203)=22.26, p<.001, \eta^2=.25$ ) scales.

Pairwise comparisons revealed multiple group differences within each of the scales. Figure 3 shows the MSEL subscale scores at each age, with reference to the US normed scores.

On the Cog T scores, the 5-9m group had significantly higher scores than both the 15-19m ( $p<.001, d=1.06$ ) and 20-24m ( $p<.001, d=1.43$ ) groups. Similarly, the 10-14m group also had better performance than both the 15-19m ( $p=.01, d=.62$ ) and 20-24m ( $p<.001, d=.94$ ) groups.

On the VR scale, the 5-9m group scored significantly higher than the 10-14m ( $p=.002, d=1.01$ ), 15-19m ( $p<.001, d=1.57$ ) and 20-24m ( $p<.001, d=1.92$ ) groups. Likewise, the 10-14m group scored higher than the 15-19m ( $p=.001, d=.65$ ) and 20-24m ( $p<.001, d=.88$ ) groups. On the FM scale, both the 10-14m ( $p<.001, d=1.28$ ) and 15-19m ( $p=.02, d=.88$ ) groups had higher scores than the 20-24m olds.

Finally, on the RL subscale, the 5-9m old group scored significantly higher than the 15-19m olds ( $p=.01, d=.67$ ), but no other significant effects emerged. On the EL scale, the 5-9m olds scored significantly higher than both the 15-19m ( $p=.02, d=.92$ ) and the 20-24m ( $p<.001, d=1.21$ ) groups.

### ***Longitudinal Changes in MSEL Performance***

Repeated measures ANOVA (controlling for infant sex) in the sample that was tested longitudinally revealed that there were significant declines in MSEL performance with age, in both the Cog T and individual subscale scores. Table 5 summarises the MSEL scores at the baseline and follow-up visits. The baseline MSEL scores were significantly higher than the follow-up for the Cog T ( $F(1, 32)=5.50, p=.03, d=.65$ ), VR ( $F(1, 32)=4.51, p=.04, d=.57$ ) and RL ( $F(1, 32)=16.30, p<.001, d=1.18$ ) scales.

Subsequently, the number of participants in the categories of *Average and Above* and *Below Average* were examined. At visit 1, 29 participants (85.29%) scored *Average or*



*Above*, while 5 participants (14.71%) scored *Below Average*. At visit 2, 19 participants (55.88%) scored *Average or Above*, while 15 scored (44.12%) scored *Below Average*.

McNemar's test was used to compare the number of participants that had changed category between *Average and Above* and *Below Average* from visit 1 to visit 2, the overall test was significant,  $p=.01$ . Post hoc analyses revealed that 11 participants who were *Average of Above* at visit 1 were categorised as *Below Average* at visit 2,  $p=.01$ . On the other hand, only 1 participant that was *Below Average* at visit 1 moved up to *Average and Above* at visit 2 and this did not reach statistical significance ( $p=1.00$ ).

### ***Association between MSEL Scores and Measures of Growth***

Table 6 presents a summary of the anthropometric measures in each of the age groups. The children showed significant faltering across the period of study, with significantly higher WHZ ( $F(3, 216)=4.98, p=0.02, \eta^2=.07$ ), HAZ ( $F(3, 216)=2.86, p=0.04, \eta^2=.04$ ) and HCZ ( $F(3, 216)=3.48, p=0.02, \eta^2=.05$ ) at the earlier time points. Pairwise comparisons revealed that in the WHZ measures, the 15-19m group had significantly lower z-scores than the 5-9m ( $p=.002, d=.83$ ) and 10-14m ( $p=.03, d=.32$ ) groups. For HAZ measures, the 5-9m group had significantly higher z-scores than the 10-14m group ( $p=.04, d=.53$ ). Finally, in the HCZ measures, the 5-9m group had significantly higher z-scores than the 10-14m group ( $p=.01, d=.63$ ).

There were significant sex differences in HCZ scores ( $F(1, 224)=7.65, p=.01, \eta^2=.03$ ), with girls having higher scores across age ( $M=-.81, SD=.95$ ) than boys ( $M=-1.16, SD=.99$ ),  $d=.36$ . There was also a trend-level sex difference in HAZ ( $F(1, 224)=2.89, p=.09, \eta^2=.01$ ), where girls ( $M=-1.06, SD=1.13$ ) had higher scores than boys ( $M=-1.34, SD=1.37$ ),  $d=.22$ . On the contrary, there was no significant difference in WHZ ( $F(1, 224)=1.27, p=.26, \eta^2=.01$ ).

Correlation analysis revealed a significant association between Cog T scores and HAZ ( $r_s(209)=.21, p=.002$ ) and WHZ ( $r_s(209)=.17, p=.02$ ). On the other hand, there was no association between HCZ and Cog T scores  $r_s(209)=-.01, p=.92$ .

Regression analysis revealed that, after adjusting for sex and phase of testing, both age ( $\beta=-.36, t(203)=-5.81, p<.001$ ) and HAZ ( $\beta=.26, t(203)=4.05, p<.001$ ) significantly predicted Cog T performance. The HAZ and age interaction showed trend-level significance ( $\beta=.11, t(203)=1.70, p=.09$ ). Similarly, in the model where WHZ was used as a predictor, both age ( $\beta=-.35, t(203)=-5.39, p<.001$ ) and WHZ ( $\beta=.16, t(203)=2.35, p=.02$ ) significantly predicted Cog T scores, while the interaction between age and WHZ did not ( $\beta=.02, t(203)=.34, p=.74$ ).

### Discussion

The present study was the first to implement a population-specific, tailored version of the MSEL for use with infants in a rural area of The Gambia. The sample, aged 5-24-months, obtained lower MSEL scores than US norms and scores worsened with increasing age. Additionally, a significant association emerged between measures of growth, specifically height-for-age and weight-for-height, and cognitive performance. These findings suggest that the adapted MSEL was sensitive enough to detect age and risk related trajectories of development within this sample. Furthermore, with these findings, we contribute to a growing body of research, which suggests that infants growing up in LMICs are at heightened risk for adverse cognitive outcomes.

#### *Performance of the Gambian Sample and Change in Scores Across Age*

The difference between the Gambian infants' MSEL scores and US norms was observed as early as 10-14-months of age and continued to be evident through to 24-months. These findings are consistent with reports in prior literature (Fernald, Kariger, Hidrobo, & Gertler, 2012; Hamadani et al., 2014; Koura et al., 2013), which suggest that infants growing

up in low-resource environments begin to exhibit reduced cognitive ability within the first year of life. Similarly, we observed that a substantial proportion of participants in the older age categories (15-24-months) had MSEL scores that would be categorised as below average, suggesting that these children may be at-risk for developmental delay.

When MSEL performance was compared across age groups, there was a significant decline in scores on both global cognitive performance and on individual subscales. Importantly, the decline in overall cognitive ability started to be evident in the 15-19-month old group, but there was variation in the trajectories for individual subscales. For example, reduced performance on the VR scale was evident at a much younger age, with the 5-9-month old group outperforming all of the older age groups and with evident decline in performance among the older age groups. On the other hand, declines in language performance only became evident among the 15-19-month and 20-24-month olds. Visual processing and fine motor skills begin to develop early in infancy (Thompson & Nelson, 2001). On the other hand, while there is evidence of language comprehension from as early as 6-9 months (Bergelson & Swingley, 2012), receptive language skills develop rapidly around 12-months of age (Swanson et al., 2017). Likewise, children start producing their first words around their first birthday but undergo a vast increase in vocabulary by 24- to 36-months (Swanson et al., 2017). The maturation of early visual and motor skills is fundamental for the development of language ability later on. For example, visual reception is crucial for label mapping and motor skills are important for providing an infant with opportunities to explore their environment (Iverson, 2010). Thus, the differential onset of decline observed in the different cognitive domains is consistent with the age when milestones in these domains are expected to occur and there appears to be an accumulation in derailment with increasing age.

The regression in cognitive ability becomes starker when the group of infants tested longitudinally are taken into consideration. Firstly, the infants' cognitive scores significantly decreased within a 3-month time period. Additionally, this reduction was so pronounced among some participants, that a significant proportion changed from having scores in the average category to falling below average.

The MSEL has been used to map trajectories of cognitive development among a variety of risk groups. Similar declines in cognitive performance are observed within the first 12-24-months of life among infants with increased familial risk for ASD (Brian et al., 2014) and infants born pre-term (Yaari et al., 2018). Taken together, these findings highlight a developmental period where infants at-risk for cognitive delay begin to exhibit the first behavioural manifestations of adverse outcomes. This is unsurprising, given that this is the period when infants begin to actively engage with their environment and, in turn, environmental demands increase (Karasik, Tamis-Lemonda, & Adolph, 2014). However, the earlier onset of decline in the VR subscale also suggests that atypicalities actually start to emerge in the perceptual domains first, on which later development cascades.

#### ***Association between Cognitive Performance and Measures of Growth***

The population tested in this study was selected because of the high prevalence of undernutrition and difficulties in growth observed among children in this setting (Hennig et al., 2015; van der Merwe et al., 2013). Within our sample, the average WHZ, HAZ and HCZ scores all fell in the range of normal growth according to WHO criteria (WHO Multicentre Growth Reference Study Group, 2006), however for both HAZ and HCZ from 10-months onwards the means fell around 1 z-score below the world-wide average. There was evidence of wasting and stunting among a substantial proportion of the children tested. Likewise, all three measures of growth evidenced decline with age from 5- to 24-months.

Consistent with prior research (e.g. Sudfeld et al., 2015), greater HAZ scores were significantly associated with improved MSEL scores, as were WHZ scores. On the contrary, HCZ was not associated with cognitive performance. Stunting (indicated by reduced height for age) is considered to be an indicator of chronic undernutrition, while wasting (reduced weight for age) is more transient and is, thus, an indicator of acute malnutrition (Briend, Khara, & Dolan, 2015). This is perhaps why a majority of studies that examine the association between growth and cognitive development focus primarily on stunting (Perkins et al., 2017; Sudfeld et al., 2015). However, our results contribute to a sparser body of literature (e.g. McDonald et al., 2013), which suggests that reduced weight-for-height is also associated with poorer cognitive outcomes. The parallel onset of decline in physical and cognitive development may be partially attributed to the introduction of weaning foods at approximately 4-6 months of age (Eriksen et al., 2017; Onofiok & Nnanyelugo, 1998). Lack of nutrient-rich weaning foods and contamination result in children getting less energy and micronutrient dense foods and increased exposure to pathogens (Arpadi et al., 2009).

Given that HC is posited to be a measure of brain growth (Bartholomeusz, Courchesne, & Karns, 2002) it is surprising that an association between HCZ and MSEL scores was not observed in our sample. However, Lira et al. (2010) report that growth in head circumference in the first 6-months of life significantly predicts cognitive outcomes, while subsequent growth does not. Thus, it is possible that the children in our sample were too old to exhibit this association and future research should begin to examine head growth at even younger ages.

There are a multitude of environmental factors that can impact on both infant physical growth and cognitive development (Perkins et al., 2017). In this present study, we did not include measures of other key variables, which limits our ability to make firm conclusions about the nature of this association or any causal inferences. Therefore, we propose that

findings from the present study be interpreted as signalling the presence of an association, upon which further research can build. In the aforementioned BRIGHT project, we have included a multitude of measures that are relevant to child physical growth and cognitive development, such as SES, measures of parent-infant interaction, and maternal mental health, among others.

### ***Strengths and challenges of the adapted MSEL and implications for future research***

Findings from this study suggest that the use of standardised cognitive assessments with infants in LMICs is both feasible and pressing. The adapted MSEL used has several important advantages; it was developed for and tested among children living in a rural setting and was translated into a local language. However, one of the goals of the current study was to highlight the challenges that we encountered while attempting to implement the MSEL into the rural Gambian setting and several key issues emerged.

Infants in this setting are often unfamiliar with the general set-up used in behavioural testing. For example, toys and books are not readily available in this environment and most play takes place outdoors (Bradley & Corwyn, 2005). Furthermore, children in LMICs are taught to regard adults as authority figures and are discouraged from interacting with strangers (Bradley & Corwyn, 2005). Whilst this has not been studied or documented in The Gambia, in neighbouring rural regions in Senegal, adults are also discouraged from talking to pre-verbal infants, as cultural beliefs posit that children so young cannot understand speech and that interacting with them is futile (Weber, Fernald, & Diop, 2017). As a consequence of these circumstances, an activity that involves spending a prolonged period indoors, engaging with novel objects on a table top, interacting with an unfamiliar adult experimenter, and being video-recorded, is an unusual experience for infants in this setting. We observed that the participants were very reticent during testing, even after receiving encouragement from caregivers. The administration time would regularly reach double the expected time from

similar aged infants and toddlers undertaking this assessment in our UK studies: this became more apparent as participant age increased. It is possible that girls in this rural environment are even more greatly discouraged from exploring novel environments and interacting with strangers. This could partly explain our finding that girls had reduced performance, which is contradictory to prior literature (Koura et al., 2013). It is also possible that an age-related decline in GM skills was not observed because physical play outdoors allows children to hone their motor abilities.

Infant non-compliance and reticence to engage were often reasons why experimenters resorted to parent-report to score an item. It is encouraging that the number of items scored through parent-report was not significantly associated with overall scores on the MSEL. Koura et al. (2013) reported that MSEL scores of infants in Benin were significantly associated with a parent-report measure of cognitive development. Coupled with these findings, our results suggest that parent-report measures may be useful to supplement behavioural measures in low-resource settings.

A limitation of this study was that, in the absence of a control group, we relied solely on US-normed t-scores. The MSEL was originally normed among typically-developing children in the US and its validity for testing more diverse populations has been called into question (e.g. Akshoomoff, 2006). Furthermore, since items become more advanced and require more instruction as children get older, the measure could be more culturally sensitive among older children and this could have contributed to the age-related differences observed in this sample. Conversely, it also is possible that the sensitivity of behavioural measures such as the MSEL increases with age. For example, Nishimura, Takei, Tsuchiya, Asano & Mori (2016) report that the MSEL was better able to discriminate children with developmental delay when they were older than 12-months than during the first few months of life. Therefore, we cannot rule out that the decrease in performance among older age

groups in our sample could be due to improved effectiveness of the measure with increasing age.

Given the challenges outlined above, it is possible that the standardised scores may not necessarily be suitable criteria against which to compare children from LMICs. While this posed less of a problem in the present study that was examining longitudinal changes in a single population, we propose comparing performance on raw scores when possible. Additionally, if measures such as the MSEL are to be used to identify children at-risk for cognitive delay, it is crucial to create norms for the specific population, or to identify comparable populations for comparison. While we made every effort to optimise the assessment for this population, it remains difficult to ascertain whether the adaptations of the test materials that we selected were truly optimal for children in this setting. A remaining caveat to the current findings is that it is not possible for us to disentangle whether the decline in cognitive performance with age is also confounded by the appropriateness of this cognitive development scale, for infants and toddlers of this age, within this population.



**References**

- Akshoomoff, N. (2006). Use of the Mullen Scales of Early Learning for the assessment of young children with Autism Spectrum Disorders. *Child Neuropsychology*, *12*(4–5), 269–277. <https://doi.org/10.1080/09297040500473714>
- Alderman, H., Behrman, J. R., Grantham-McGregor, S., Lopez-Boo, F., & Urzua, S. (2014). Economic perspectives on integrating early child stimulation with nutritional interventions. *Annals of the New York Academy of Sciences*, *1308*, 129–138. <https://doi.org/10.1111/nyas.12331>
- Arpadi, S., Fawzy, A., Aldrovandi, G. M., Kankasa, C., Sinkala, M., Mwiya, M., Thea, D. M., & Kuhn, L. (2009). Growth faltering due to breastfeeding cessation in uninfected children born to HIV-infected mothers in Zambia. *The American Journal of Clinical Nutrition*, *90*(2), 344–353. <https://doi.org/10.3945/ajcn.2009.27745>
- Ballot, D. E., Ramdin, T., Rakotsoane, D., Agaba, F., Davies, V. A., Chirwa, T., & Cooper, P. A. (2017). Use of the Bayley Scales of Infant and Toddler Development, Third Edition, to assess developmental outcome in infants and young children in an urban setting in South Africa. *International Scholarly Research Notices*, *2017*, 1–5. <https://doi.org/10.1155/2017/1631760>
- Bartholomeusz, H. H., Courchesne, E., & Karns, C. M. (2002). Relationship between head circumference and brain volume in healthy normal toddlers, children, and adults. *Neuropediatrics*, *33*(5), 239–241. <https://doi.org/10.1055/s-2002-36735>
- Baulch, B. (2011). *Why Poverty Persists: Poverty Dynamics in Asia and Africa*. Northampton, MA: Edward Elgar Publishing Limited.
- Bayley, N. (1995). *Manual For the Bayley Scales of Infant Development*. New York: Psychological Corporation.
- Beasley, T. M., & Schumacker, R. E. (1995). Multiple Regression Approach to Analyzing

- Contingency Tables: Post Hoc and Planned Comparison Procedures. *The Journal of Experimental Education*, 64(1), 79–93. <https://doi.org/10.1080/00220973.1995.9943797>
- Begus, K., Lloyd-Fox, S., Halliday, D., Papademetriou, M., Darboe, M. K., Prentice, A. M., Moore, S. E., & Elwell, C. E. (2016). Using fNIRS to study working memory of infants in rural Africa. *Advances in Experimental Medicine and Biology*, 876, 273–279. [https://doi.org/10.1007/978-1-4939-3023-4\\_34](https://doi.org/10.1007/978-1-4939-3023-4_34)
- Benton, D. (2008). The influence of children's diet on their cognition and behavior. *European Journal of Nutrition*, 47(S3), 25–37. <https://doi.org/10.1007/s00394-008-3003-x>
- Bergelson, E., & Swingle, D. (2012). At 6-9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences of the United States of America*, 109(9), 3253–3258. <https://doi.org/10.1073/pnas.1113380109>
- Bodeau-Livinec, F., Davidson, L. L., Zoumenou, R., Massougbojji, A., Cot, M., & Boivin, M. J. (2018). Neurocognitive testing in West African children 3–6 years of age: Challenges and implications for data analyses. *Brain Research Bulletin*. <https://doi.org/10.1016/j.brainresbull.2018.04.001>
- Boivin, M. J., Bangirana, P., Nakasujja, N., Page, C. F., Shohet, C., Givon, D., Bass, J. K., Opoka, R. O., & Klein, P. S. (2013). A Year-long caregiver training program to improve neurocognition in preschool Ugandan HIV-exposed children. *Journal of Developmental & Behavioral Pediatrics*, 34(4), 269–278. <https://doi.org/10.1097/DBP.0b013e318285fba9>
- Bornman, J., Ronski, M., Tonsing, K., Sevcik, R., White, R., Barton-Hulsey, A., & Morwane, R. (2018). Adapting and translating the Mullen Scales of Early Learning for the South African context. *South African Journal of Communication Disorders*, 65(1), 9. <https://doi.org/10.4102/sajcd.v65i1.571>

- Bornman, J., Sevcik, R. A., Ronski, M., & Pae, H. K. (2010). Successfully translating language and culture when adapting assessment measures. *Journal of Policy and Practice in Intellectual Disabilities*, 7(2), 111–118. <https://doi.org/10.1111/j.1741-1130.2010.00254.x>
- Bradley, R. H., & Corwyn, R. F. (2005). Caring for children around the world: A view from HOME. *International Journal of Behavioral Development*, 29(6), 468–478. <https://doi.org/10.1177/01650250500146925>
- Brian, A. J., Roncadin, C., Duku, E., Bryson, S. E., Smith, I. M., Roberts, W., Szatmari, P., Drmic, I., & Zwaigenbaum, L. (2014). Emerging cognitive profiles in high-risk infants with and without autism spectrum disorder. *Research in Autism Spectrum Disorders*, 8(11), 1557–1566. <https://doi.org/10.1016/J.RASD.2014.07.021>
- Briend, A., Khara, T., & Dolan, C. (2015). Wasting and stunting - similarities and differences: Policy and programmatic implications. *Food and Nutrition Bulletin*, 36(1\_suppl1), S15–S23. <https://doi.org/10.1177/15648265150361S103>
- Bruderer, A. G., Danielson, D. K., Kandhadai, P., & Werker, J. F. (2015). Sensorimotor influences on speech perception in infancy. *Proceedings of the National Academy of Sciences*, 112(44), 13531–13536. Retrieved from <http://www.pnas.org/content/112/44/13531.abstract>
- Deen, B., Richardson, H., Dilks, D. D., Takahashi, A., Keil, B., Wald, L. L., Kanwisher, N., & Saxe, R. (2017). Organization of high-level visual cortex in human infants. *Nature Communications*, 8, 1–10. <https://doi.org/10.1038/ncomms13995>
- Dominguez-Salas, P., Moore, S. E., Cole, D., da Costa, K.-A., Cox, S. E., Dyer, R. A., Fulford, A. J., Innis, S. M., Waterland, R. A., Zeisel, S. H., Prentice, A. M., & Hennig, B. J. (2013). DNA methylation potential: dietary intake and blood concentrations of one-carbon metabolites and cofactors in rural African women. *The American Journal of*

*Clinical Nutrition*, 97(6), 1217–1227. <https://doi.org/10.3945/ajcn.112.048462>

Emberson, L. L., Richards, J. E., & Aslin, R. N. (2015). Top-down modulation in the infant brain: Learning-induced expectations rapidly affect the sensory cortex at 6 months.

*Proceedings of the National Academy of Sciences*, 112(31), 9585–9590.

<https://doi.org/10.1073/pnas.1510343112>

Eriksen, K. G., Johnson, W., Sonko, B., Prentice, A. M., Darboe, M. K., & Moore, S. E.

(2017). Following the World Health Organization's recommendation of exclusive breastfeeding to 6 months of age does not impact the growth of rural Gambian infants.

*The Journal of Nutrition*, 147(2), 248–255. <https://doi.org/10.3945/jn.116.241737>

Fernald, L. C. H., Kariger, P., Hidrobo, M., & Gertler, P. J. (2012). Socioeconomic gradients in child development in very young children: Evidence from India, Indonesia, Peru, and Senegal. *Proceedings of the National Academy of Sciences*, 109(Supplement\_2), 17273–17280. <https://doi.org/10.1073/pnas.1121241109>

Fernald, L. C. H., Neufeld, L. M., Barton, L. R., Schnaas, L., Rivera, J., & Gertler, P. J.

(2006). Parallel deficits in linear growth and mental development in low-income Mexican infants in the second year of life. *Public Health Nutrition*, 9(2), 178–186.

Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/16571171>

Fox, N. A., Calkins, S. D., & Bell, M. A. (1994). Neural plasticity and development in the

first two years of life: Evidence from cognitive and socioemotional domains of research.

*Development and Psychopathology*, 6(4), 677–696.

<https://doi.org/10.1017/S0954579400004739>

Fox, S. E., Levitt, P., & Nelson, C. A. (2010). How the timing and quality of early

experiences influence the development of brain architecture. *Child Development*, 81(1),

28–40. <https://doi.org/10.1111/j.1467-8624.2009.01380.x>

Gamble, D. P. (1987). *Intermediate Gambian Mandinka-English Dictionary*. Gamble.

- Gammer, I., Bedford, R., Elsabbagh, M., Garwood, H., Pasco, G., Tucker, L., Volein, A., Johnson, M. H., Charman, T., & BASIS Team. (2015). Behavioural markers for autism in infancy: Scores on the Autism Observational Scale for Infants in a prospective study of at-risk siblings. *Infant Behavior and Development*, *38*, 107–115.  
<https://doi.org/10.1016/j.infbeh.2014.12.017>
- Grantham-McGregor, S., Cheung, Y. B., Cueto, S., Glewwe, P., Richter, L., Strupp, B., & International Child Development Steering Group. (2007). Developmental potential in the first 5 years for children in developing countries. *Lancet*, *369*, 60–70.  
[https://doi.org/10.1016/S0140-6736\(07\)60032-4](https://doi.org/10.1016/S0140-6736(07)60032-4)
- Hamadani, J. D., Tofail, F., Huda, S. N., Alam, D. S., Ridout, D. A., Attanasio, O., & Grantham-McGregor, S. M. (2014). Cognitive deficit and poverty in the first 5 years of childhood in Bangladesh. *Pediatrics*, *134*(4), e1001–e1008.  
<https://doi.org/10.1542/peds.2014-0694>
- Hennig, B. J., Unger, S. A., Dondeh, B. L., Hassan, J., Hawkesworth, S., Jarjou, L., Jones, K. S., Moore, S. E., Nabwera, H. M., Ngum, M., Prentice, A., Sonko, B., Prentice, A. M., & Fulford, A. J. (2015). Cohort profile: The Kiang West Longitudinal Population Study (KWLPS) - A platform for integrated research and health care provision in rural Gambia. *International Journal of Epidemiology*, *13*, 1–12.  
<https://doi.org/10.1093/ije/dyv206>
- IBM Corp. (2016). IBM SPSS Statistics Version 24.0. Armonk, NY: IBM Corp.
- Ip, P., Ho, F. K. W., Rao, N., Sun, J., Young, M. E., Chow, C. B., Tso, W., & Hon, K. L. (2017). Impact of nutritional supplements on cognitive development of children in developing countries: A meta-analysis. *Scientific Reports*, *7*(1), 1–9.  
<https://doi.org/10.1038/s41598-017-11023-4>
- Isaacs, E., & Oates, J. (2008). Nutrition and cognition: assessing cognitive abilities in

children and young people. *European Journal of Nutrition*, 47(S3), 4–24.

<https://doi.org/10.1007/s00394-008-3002-y>

Iverson, J. M. (2010). Developing language in a developing body: The relationship between motor development and language development. *Journal of Child Language*, 37(2), 1–25.

<https://doi.org/10.1017/S0305000909990432>.Developing

Karasik, L. B., Tamis-Lemonda, C. S., & Adolph, K. E. (2014). Crawling and walking infants elicit different verbal responses from mothers. *Developmental Science*, 17(3), 388–395. <https://doi.org/10.1111/desc.12129>

Kolb, B., & Gibb, R. (2011). Brain plasticity and behaviour in the developing brain. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, 20(4), 265–276.

Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/22114608>

Koura, K. G., Boivin, M. J., Davidson, L. L., Ouédraogo, S., Zoumenou, R., Alao, M. J., Garcia, A., Massougbojji, A., Cot, M., & Bodeau-Livinec, F. (2013). Usefulness of child development assessments for low-resource settings in francophone Africa. *Journal of Developmental & Behavioral Pediatrics*, 34(7), 486–493.

<https://doi.org/10.1097/DBP.0b013e31829d211c>

Leonard, H. C., Bedford, R., Charman, T., Elsabbagh, M., Johnson, M. H., & Hill, E. L.

(2014). Motor development in children at risk of autism: A follow-up study of infant siblings. *Autism*, 18(3), 281–291. <https://doi.org/10.1177/1362361312470037>

Levine, D., Strother-Garcia, K., Golinkoff, R. M., & Hirsh-Pasek, K. (2016). Language development in the first year of life: What deaf children might be missing before cochlear implantation. *Otology & Neurotology*, 37(2), e56–e62.

<https://doi.org/10.1097/MAO.0000000000000908>

Lira, P. I. C., Eickmann, S. H., Lima, M. C., Amorim, R. J., Emond, A. M., & Ashworth, A. (2010). Early head growth: Relation with IQ at 8 years and determinants in term infants

of low and appropriate birthweight. *Developmental Medicine and Child Neurology*, 52(1), 40–46. <https://doi.org/10.1111/j.1469-8749.2009.03353.x>

Lloyd-Fox, S., Begus, K., Halliday, D., Pirazzoli, L., Blasi, A., Papademetriou, M., Darboe, M. K., Prentice, A. M., Johnson, M. H., Moore, S. E., & Elwell, C. E. (2017). Cortical specialisation to social stimuli from the first days to the second year of life: A rural Gambian cohort. *Developmental Cognitive Neuroscience*, 25, 92–104. <https://doi.org/10.1016/j.dcn.2016.11.005>

Lloyd-Fox, S., Papademetriou, M., Darboe, M. K., Everdell, N. L., Wegmuller, R., Prentice, A. M., Moore, S. E., & Elwell, C. E. (2014). Functional near infrared spectroscopy (fNIRS) to assess cognitive function in infants in rural Africa. *Scientific Reports*, 4(4740), 1–8. <https://doi.org/10.1038/srep04740>

McCoy, D. C., Peet, E. D., Ezzati, M., Danaei, G., Black, M. M., Sudfeld, C. R., Fawzi, W., & Fink, G. (2016). Early childhood developmental status in low- and middle-income countries: National, regional, and global prevalence estimates using predictive modeling. *PLOS Medicine*, 13(6), 1–18. <https://doi.org/10.1371/journal.pmed.1002034>

McDonald, C. M., Manji, K. P., Kupka, R., Bellinger, D. C., Spiegelman, D., Kisenge, R., Msamanga, G., Fawzi, W. W., & Duggan, C. P. (2013). Stunting and wasting are associated with poorer psychomotor and mental development in HIV-exposed Tanzanian infants. *Journal of Nutrition*, 143(2), 204–214. <https://doi.org/10.3945/jn.112.168682>

Moore, S. E., Fulford, A. J. C., Darboe, M. K., Jobarteh, M. L., Jarjou, L. M., & Prentice, A. M. (2012). A randomized trial to investigate the effects of pre-natal and infant nutritional supplementation on infant immune development in rural Gambia: the ENID trial: Early Nutrition and Immune Development. *BMC Pregnancy and Childbirth*, 12(107), 1–8. <https://doi.org/10.1186/1471-2393-12-107>

- Mullen, E. (1995). *Mullen Scales of Early Learning*. Circle Pines, NM: American Guidance Service.
- Murray-Kolb, L. E., Rasmussen, Z. A., Scharf, R. J., Rasheed, M. A., Svensen, E., Seidman, J. C., Tofail, F., Koshy, B., Shrestha, R., Maphula, A., Vasquez, A. O., da Costa, H. P., Yousafzai, A. K., Oria, R. B., Roshan, R., Bayyo, E. B., Kosek, M., Shrestha, S., Schaefer, B. A., Bessong, P., Ahmed, T., Lang, D., & MAL-ED Network Investigators. (2014). 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. *Clinical Infectious Diseases*, *59*(suppl 4), S261–S272. <https://doi.org/10.1093/cid/ciu437>
- Nabwera, H. M., Fulford, A. J., Moore, S. E., & Prentice, A. M. (2017). Growth faltering in rural Gambian children after four decades of interventions: a retrospective cohort study. *The Lancet. Global Health*, *5*(2), e208–e216. [https://doi.org/10.1016/S2214-109X\(16\)30355-2](https://doi.org/10.1016/S2214-109X(16)30355-2)
- Nishimura, T., Takei, N., Tsuchiya, K. J., Asano, R., & Mori, N. (2016). Identification of neurodevelopmental trajectories in infancy and of risk factors affecting deviant development: A longitudinal birth cohort study. *International Journal of Epidemiology*, *45*(2), 543–553. <https://doi.org/10.1093/ije/dyv363>
- Nurliyana, A. R., Mohd Shariff, Z., Mohd Taib, M. N., Gan, W. Y., & Tan, K. A. (2016). Early nutrition, growth and cognitive development of infants from birth to 2 years in Malaysia: A study protocol. *BMC Pediatrics*, *16*(160), 1–7. <https://doi.org/10.1186/s12887-016-0700-0>
- Nyaradi, A., Li, J., Hickling, S., Foster, J., & Oddy, W. H. (2013). The role of nutrition in children's neurocognitive development, from pregnancy through childhood. *Frontiers in Human Neuroscience*, *7*, 1–16. <https://doi.org/10.3389/fnhum.2013.00097>



- Onofiok, N. O., & Nnanyelugo, D. O. (1998). Weaning foods in West Africa : Nutritional problems and possible solutions. *Food and Nutrition Bulletin*, *19*(1), 27–33.  
<https://doi.org/10.1177/156482659801900105>
- Papademetriou, M. D., Lloyd-Fox, S., Everdell, N. L., Darboe, M. K., Moore, S. E., Prentice, A. M., & Elwell, C. E. (2014). Optical imaging of brain activation in Gambian infants. *Advances in Experimental Medicine and Biology*, *812*, 263–269.  
[https://doi.org/10.1007/978-1-4939-0620-8\\_35](https://doi.org/10.1007/978-1-4939-0620-8_35)
- Peña, E. D. (2007). Lost in translation: Methodological considerations in cross-cultural research. *Child Development*, *78*(4), 1255–1264. <https://doi.org/10.1111/j.1467-8624.2007.01064.x>
- Perkins, J. M., Kim, R., Krishna, A., McGovern, M., Aguayo, V. M., & Subramanian, S. V. (2017). Understanding the association between stunting and child development in low- and middle-income countries: Next steps for research and intervention. *Social Science & Medicine*, *193*, 101–109. <https://doi.org/10.1016/j.socscimed.2017.09.039>
- Petrenko, C. L. M. (2013). A review of intervention programs to prevent and treat behavioral problems in young children with developmental disabilities. *Journal of Developmental and Physical Disabilities*, *25*(6), 1–28.  
<https://doi.org/10.1109/TMI.2012.2196707.Separate>
- Prado, E. L., & Dewey, K. G. (2014). Nutrition and brain development in early life. *Nutrition Reviews*, *72*(4), 267–284. <https://doi.org/10.1111/nure.12102>
- Rayco-Solon, P., Moore, S. E., Fulford, A. J., & Prentice, A. M. (2004). Fifty-year mortality trends in three rural African villages. *Tropical Medicine and International Health*, *9*(11), 1151–1160. <https://doi.org/10.1111/j.1365-3156.2004.01325.x>
- Schady, N., Behrman, J., Araujo, M. C., Azuero, R., Bernal, R., Bravo, D., Lopez-Boo, F., Macours, K., Marshall, D., Paxson, C., & Vakis, R. (2015). Wealth gradients in early

- childhood cognitive development in five Latin American countries. *Journal of Human Resources*, 50(2), 446–463. <https://doi.org/10.3368/jhr.50.2.446>
- Sokolovic, N., Selvam, S., Srinivasan, K., Thankachan, P., Kurpad, A. V., & Thomas, T. (2014). Catch-up growth does not associate with cognitive development in Indian school-age children. *European Journal of Clinical Nutrition*, 68(1), 14–18. <https://doi.org/10.1038/ejcn.2013.208>
- Stewart, C. P., Kariger, P., Fernald, L., Pickering, A. J., Arnold, C. D., Arnold, B. F., Hubbard, A. E., Dentz, H. N., Lin, A., Meerkerk, T. J., Milner, E., Swarouth, J., Colford, J. M., & Null, C. (2018). Effects of water quality, sanitation, handwashing, and nutritional interventions on child development in rural Kenya (WASH Benefits Kenya): A cluster-randomised controlled trial. *The Lancet Child and Adolescent Health*, 2(4), 269–280. [https://doi.org/10.1016/S2352-4642\(18\)30025-7](https://doi.org/10.1016/S2352-4642(18)30025-7)
- Sudfeld, C. R., Charles McCoy, D., Danaei, G., Fink, G., Ezzati, M., Andrews, K. G., & Fawzi, W. W. (2015). Linear growth and child development in low- and middle-income countries: A meta-analysis. *Pediatrics*, 135(5), e1266–e1275. <https://doi.org/10.1542/peds.2014-3111>
- Swanson, M. R., Wolff, J. J., Elison, J. T., Gu, H., Hazlett, H. C., Botteron, K., Styner, M., Paterson, S., Gerig, G., Constantino, J., Dager, S., Estes, A., Vachet, C., Piven, J., & IBIS Network. (2017). Splenium development and early spoken language in human infants. *Developmental Science*, 20(2), e12360. <https://doi.org/10.1111/desc.12360>
- The Gambia Bureau of Statistics. (2011). *The Gambia- multiple indicator cluster survey 2010*. UNICEF.
- Thompson, R. A., & Nelson, C. A. (2001). Developmental science and the media. Early brain development. *The American Psychologist*, 56(1), 5–15. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11242988>

- Tofail, F., Fernald, L. C., Das, K. K., Rahman, M., Ahmed, T., Jannat, K. K., Unicomb, L., Arnold, B. F., Ashraf, S., Winch, P. J., Kariger, P., Stewart, C. P., Colford, J. M., & Luby, S. P. (2018). Effect of water quality, sanitation, hand washing, and nutritional interventions on child development in rural Bangladesh (WASH Benefits Bangladesh): A cluster-randomised controlled trial. *The Lancet Child and Adolescent Health*, 2(4), 255–268. [https://doi.org/10.1016/S2352-4642\(18\)30031-2](https://doi.org/10.1016/S2352-4642(18)30031-2)
- UN. (2015). *Sustainable Development Goals*. United Nations.
- van der Merwe, L. F., Moore, S. E., Fulford, A. J., Halliday, K. E., Drammeh, S., Young, S., & Prentice, A. M. (2013). Long-chain PUFA supplementation in rural African infants: A randomized controlled trial of effects on gut integrity, growth, and cognitive development. *American Journal of Clinical Nutrition*, 97(April), 45–57. <https://doi.org/10.3945/ajcn.112.042267.1>
- Weber, A., Fernald, A., & Diop, Y. (2017). When cultural norms discourage talking to babies: Effectiveness of a parenting program in rural Senegal. *Child Development*, 88(5), 1513–1526. <https://doi.org/10.1111/cdev.12882>
- WHO Multicentre Growth Reference Study Group. (2006). *WHO Child Growth Standards: Length/height-for-age, Weight-for-age, Weight-for-length, Weight-for-height, and Body Mass Index-for-age: Methods and Development*. Geneva: World Health Organization.
- Wright, S. (2016). *Language Policy and Language Planning: From Nationalism to Globalisation*. Palgrave Macmillan UK.
- Yaari, M., Mankuta, D., Harel- Gadassi, A., Friedlander, E., Bar-Oz, B., Eventov-Friedman, S., Maniv, N., Zucker, D., & Yirmiya, N. (2018). Early developmental trajectories of preterm infants. *Research in Developmental Disabilities*, 81, 12–23. <https://doi.org/10.1016/j.ridd.2017.10.018>

## Supplementary Material

In addition to linguistic translations on the MSEL a number of stimuli and test items were found to be inappropriate for the rural Gambian setting and were replaced. Below is a list of adaptations and replacements for items in each subscale.

### **Expressive Language:**

Item 15 – Names objects: Since many children are not exposed to picture books, the book was replaced with a toy shoe.

Item 18 - Picture Vocabulary: In this item, the infant is asked to label a series of black and white, drawn images. Multiple objects depicted in these images were either novel to infants in this setting or looked different in their environment. Thus, the following replacements were made:

- Telephone → Mobile phone
- Bed → Chair
- Television → Goat
- House → Redrawn to resemble houses in West Kiang villages
- Purse → Donkey
- Clock → Cup
- Chimney → Water pump
- Lamp → Fire
- Paintbrush → Pencil
- Flower → Tree

### **Receptive Language:**

Item 14 – Identifies objects: The baby doll was replaced by a dark-skinned baby doll.

Item 17 – Follows directions: It was difficult to find a suitable word in Mandinka to translate block. Thus, the block was replaced by a small, plastic donkey, which are very commonplace animals in this area, and most children were familiar with.

Item 22 – Auditory spatial awareness: As with RL item 17, the block was replaced by a toy donkey. For similar reasons, the plastic teddy bear was replaced by a small, plastic toy shoe.

Item 24 – Identifies object function: The scissors used in this item were replaced by a plastic knife.

**Visual Reception:**

Item 11 – Makes object association: The hairbrush was replaced by a small, plastic comb.

Items 17 a and b – Matches object with/without naming: A small, toy donkey was used to replace the toy block.

Item 23 – Memory for one picture: This item involves showing children a picture of a clown, which many children in this setting are unfamiliar with. Therefore, the clown was referred to as the “funny man” or “crazy man”.

**Fine Motor:**

Item 18 – Imitates a 4-block train: The idea that the blocks represented a train was changed to say they represented a car. Additionally, instead of saying “Choo-choo”, experimenters were instructed to say “vroom vroom” so that it sounded like a noise made by a car.