

UC Davis

UC Davis Previously Published Works

Title

Extending the Developmental Milestones Checklist for use in a different context in Sub-Saharan Africa

Permalink

<https://escholarship.org/uc/item/4wz535rn>

Journal

Acta Paediatrica, 103(4)

ISSN

0803-5253

Authors

Prado, Elizabeth L
Abubakar, Amina A
Abbeddou, Souheila
[et al.](#)

Publication Date

2014-04-01

DOI

10.1111/apa.12540

Peer reviewed

Published in final edited form as:

Acta Paediatr. 2014 April 1; 103(4): 447–454. doi:10.1111/apa.12540.

Extending the Developmental Milestones Checklist for use in a different context in Sub-Saharan Africa

Elizabeth L Prado¹, Amina A Abubakar^{2,3,4}, Souheila Abbeddou¹, Elizabeth Y Jimenez⁵, Jérôme W Somé^{1,6}, Jean-Bosco Ouédraogo⁶, and members of the iLiNS-Zinc Study Team*

¹Department of Nutrition, University of California Davis, Davis, CA, USA ²Department of Cross-Cultural Psychology, Tilburg University, Tilburg, Netherlands ³Department of Child and Adolescent Studies, Utrecht University, Utrecht, Netherlands ⁴Centre for Geographic Medicine Research-Coast, Kenya Medical Research Institute, Kilifi, Kenya ⁵Departments of Individual, Family, and Community Education and Family and Community Medicine, University of New Mexico, Albuquerque, NM, USA ⁶Institut de Recherche en Sciences de la Santé/DRO, Bobo-Dioulasso, Burkina Faso

Abstract

Aim—Sub-Saharan Africa bears a disproportionate amount of global diseases related to neurodevelopmental delays in infancy, including malnutrition, malaria and HIV. Evaluating interventions to prevent such delays requires developmental assessment tools appropriate for Sub-Saharan Africa. This study aimed to develop and evaluate such a tool.

Methods—The Developmental Milestones Checklist (DMC) was developed in Kenya to provide motor, language and personal-social scores for children aged from 3 to 24 months. We developed an extended version (DMC-II) in Burkina Faso, West Africa, and then evaluated the reliability and sensitivity of the scores to age and nutritional and environmental measures.

Results—The internal, interinterviewer and test–retest reliability of the DMC-II scores were >0.7. In 214 children aged 11.6–25.4 months, each score correlated with age ($r_s > 0.7$). In 1123 children aged 16.8–19.9 months, the scores were sensitive to stunting, wasting and underweight (effect sizes 0.31–0.87 *SD*). The scores also showed expected correlations with measures of play materials in the home and activities with caregivers ($r_s = 0.13$ –0.41).

Conclusion—The DMC-II is easily used by trained fieldworkers with no previous experience in developmental assessment. It is a practical, reliable and sensitive tool for evaluating motor, language and personal-social development in different contexts in Sub-Saharan Africa.

Keywords

Child Development; Developmental assessment; Infant Development; Neurodevelopment; Sub-Saharan Africa

Correspondence: Elizabeth L Prado, PhD, Program in International and Community Nutrition, University of California at Davis, 3253 Meyer Hall, One Shields Ave, Davis, CA, 95616, USA. Tel: +1 530 752 1992 | Fax: +1 530 752 3406 | elprado@ucdavis.edu.

*Other members of the iLiNS-Zinc Study Team who contributed to the design, data acquisition and/or analysis for this paper were the following: Hess, Sonja Y¹; Vosti, Stephen¹; Ouédraogo, Zinéwendé P⁶; Brown, Kenneth H¹

Introduction

Sub-Saharan Africa lacks adequately standardised and culturally appropriate measures of child development (1). While the use of measures developed in Western countries may provide an easy and cost-effective approach to assessing child development in Africa, much evidence indicates that the use of Western measures could lead to systematic bias, which may threaten the validity of the data (2,3). Problems with imported Western measures can be caused by a lack of familiarity with the stimulus materials and test requirements (4). Problems with validity could also arise due to differences in construct definition, as well as differences in the test content, and items needed to adequately assess certain skills in a different cultural context (5,6). Variations in developmental trajectories between cultures may present yet another source of problems. For instance, Vierhaus et al. (2011) reported that children in Cameroon were unable to perform less advanced tasks on the Bayley Scales of Infant Development (BSID) motor scale, but succeeded on more advanced tasks. Following the BSID discontinuation rule would have resulted in the children being under-rated in their performance (7).

Given this problematic scenario, there have been recent efforts to develop culturally appropriate measures for use in Sub-Saharan Africa (5,8). One such measure is the Developmental Milestone Checklist (DMC), which was developed in Kenya (9). This tool uses a caregiver interview to assess the motor, language and personal-social development of children aged from 3 to 24 months. Carefully constructed caregiver report instruments have been found to be a valid way to assess child development, when compared to direct assessment of the child, for the DMC itself in Kenya (9) as well as other tools in other parts of the world (10,11).

While previous research shows that certain developmental skills are context-specific, such as climbing stairs and looking in a mirror, others, such as walking and sitting, are universal (12,13). Items in the DMC were identified through a review of several published measures, including the Griffiths Mental Developmental Scale for Infants (14) and Vineland Adaptive Behavior Scales (15). Context-appropriate items, that is, those that could be transferred from the original test to the new context, were identified through focus groups and pilot testing. The final checklist contained 66 items. Responses were given on a three-point scale (0: not observed, 1: emerging behaviour, 2: established behaviour). The DMC was found to be reliable when administered by a community health worker with minimal training, during an interview with a caregiver. For details on its development and psychometric properties, see Abubakar et al. (9).

The aim of the current study was to develop and evaluate an extended version of the DMC (DMC-II) for a new context in Sub-Saharan Africa, in the International Lipid-Based Nutrient Supplements (iLiNS) Project in Burkina Faso (iLiNS-Zinc). Specifically, the current study aimed to: (i) modify the DMC for use in a different context and expand the items to ensure that it resulted in sufficient variability in scores at the target age (18 months), (ii) evaluate the reliability (internal consistency, interinterviewer reliability and test-retest reliability) of the DMC-II scores and (iii) evaluate the validity (sensitivity to developmental progression, malnutrition and environmental variables) of the DMC-II scores.

Ideally, one would validate a new tool against a gold standard measure. In our context, we are not aware of any developmental measures that have been previously validated in the study area. Therefore, we examined whether the DMC-II scores showed expected relations with theoretically related variables (convergent validity). As children grow older, their developmental abilities increase. Consequently, scores on developmental measures such as the DMC-II were expected to increase with age, demonstrating sensitivity to maturation. Additionally, earlier studies indicate that children with growth restrictions, such as low height for age and weight for age, consistently show lower scores, on average, than children who are growing normally (16,17). We therefore expected that children with growth restriction would score significantly lower on the DMC-II. Moreover, environmental variables, such as maternal education, toys available in the home and how a child interacts with caregivers, have been consistently related to developmental scores in previous studies (18,19), although in Sub-Saharan Africa, not all studies have shown a relationship between maternal education and developmental scores (20,21). Therefore, we expected a positive relation between the DMC-II scores and the variety of toys available in the home and children's activities with caregivers, and possibly also with maternal years of education.

Methods

Study setting

This study was conducted in 34 villages of the Dandé health district in rural south-western Burkina Faso, a West African country that covers 3528 km², 172 villages and 214 470 individuals. Young children in the area are affected by holoendemic malaria transmission, food insecurity and a high prevalence of stunting and underweight (22).

General design

The study was carried out in two phases. In the first phase, we applied the necessary modifications to the DMC. In the second phase, we evaluated the psychometric properties of the new tool: the DMC-II. Ethical approval for the study procedures was obtained from the University of California Davis Institutional Review Board and the Comité d'Ethique Institutionnel du Centre Muraz, Bobo Dioulasso. All the participants provided informed consent.

Phase 1: modifications to the DMC

Initial modifications: The modifications we made to the DMC are described in detail in Appendix S1 and briefly explained here. First, we changed the administration procedure of the DMC to allow flexible administration by observations as well as caregiver interviews. As in the first version of the DMC, the caregiver's response is obtained for all items and determines the item score. In the new version, we added space on the form for the interviewer to record his or her observations on each item as well. There are several cases in which it may be useful for the interviewer to additionally administer the item to observe whether the child is able to perform the skill, for example, if the caregiver does not know if the child can perform the skill or to ensure that the caregiver has understood the question and to verify his or her response (for additional details see Appendix S1). In the iLiNS-Zinc project, we targeted children aged 18 months. Based on our experience with the DMC, and

published data on developmental achievements at this age, we selected nine items that were key to evaluating motor development at the age of 18 months. The data collectors were instructed to observe these nine key items for every child. In addition, we developed a pictorial flip chart to help the caregiver understand the meaning of the questions (Fig. 1). Finally, we added and modified certain items, presented in detail in Appendix S1.

Pilot test and additional modifications: The modified form and manual were translated into French by two bilingual French and English speakers. Four data collectors were trained to administer the DMC-II. These four data collectors verbally agreed on how the items were translated into the local languages in which the interview was administered (Dioula and Moré). They then conducted an initial pilot test with nine children aged 15–24 months. Through this pilot test, we evaluated cultural appropriateness (do the items reflect an activity or behaviour familiar to the respondents?), age appropriateness and clarity of the translations.

After making the modifications described in Appendix S1, all items were found to be appropriate for the study area. However, after collecting data from the full sample of 1123 children, three additional items were eliminated after a large percentage of caregivers responded ‘don’t know’. These were: ‘Opens door by turning and pulling doorknob’ (57%), ‘Uses spoon to feed self with spillage’ (27%) and ‘Uses spoon to feed self no spillage’ (27%). These caregivers generally responded that they did not have doorknobs in their homes and rarely used spoons. These items were not included in the calculation of scores in Phase two.

Phase two: psychometric evaluation

Sample and design: In the iLiNS-Zinc study, children were assigned to receive nutrient supplements and community-based treatment of malaria and diarrhoea from the ages of 9–18 months (immediate intervention group) or 18–27 months (delayed intervention group). At the age of 18 months, 1123 children (16.8–19.9 months, mean = 18.3, SD = 0.4, 567 boys, 376 in the delayed intervention group) were evaluated using the DMC-II. Height and weight were measured at the same visit, and an interview was administered concerning toys and activities in the home. We used the data from this sample to evaluate the internal consistency and sensitivity to growth restriction and environmental variables of the DMC-II scores. We evaluated interinterviewer and test–retest reliability using data from a subset of 16 children (seven boys) aged 17.6–19.0 months (mean = 18.3, SD = 0.5). We evaluated developmental sensitivity using data from 214 children (103 boys) in the immediate intervention group. These 214 children were divided into four groups, each tested at a different time point: 57 children (28 boys) mean age 12.2 (SD = 0.4) months, 56 children (28 boys) mean age 15.3 (0.4) months, 50 children (28 boys) mean age 21.2 (0.4) months and 51 children (19 boys) mean age 24.4 (0.5) months. A subset of 72 of the 214 children (35 boys) was also tested at a mean age 18.2(0.4) months. These 72 children included 16 children who were also tested at 12 months, 22 children also tested at 15 months, 20 children also tested at 21 months and 24 children also tested at 24 months.

Methods for collection of anthropometry, maternal education, and toys and activities in the home:

Two length and weight measurements were collected using a portable infant length board (± 0.1 cm; Seca model 417, Hamburg, Germany) and frequently standardised electronic scale (± 10 g; Seca model 383, Hamburg, Germany). For any given child, if the two length measurements differed by more than 0.5 cm or two weight measurements differed by more than 0.1 kg, a third measurement was taken. The average of the two closest measurements was used for the z-score calculation. Z-scores were calculated using the SAS (version 9.3, Cary, NC) macros for the World Health Organization Child Growth Standards (2006). Maternal educational attainment was collected via interview at enrolment.

Caregivers were asked whether seven types of toys were available for their children to play with in the home: things that make or play music, things for drawing and writing, picture books for children, things designed for stacking or building, things for moving around (balls, bats), toys for learning shapes and colours and things for pretending. They were also asked if any adult had engaged in six activities with the child in the past 3 days: read books or looked at picture books, told stories, sang songs, took the child outside the home, played with the child and spent time naming, counting, or drawing things. All items were taken from the Family Care Indicators interview (23,24) developed by United Nations Children's Fund (UNICEF) and validated in Bangladesh (25). The score for each scale was the sum of the item scores (0 = no, 1 = yes).

Data analysis: A small percentage of DMC-II item scores (0.57%) were missing. We imputed these missing scores based on the other items in the same subscale, using the method described in Raghunathan et al. (26). We calculated the score for each subscale as the sum of the item scores in that subscale (32 motor, 16 language and 28 personal-social) and the total score as the sum of all 76 item scores. To evaluate internal consistency, we computed Cronbach's alpha for each subscale and for the full set of 76 items. To evaluate test-retest reliability, we calculated the Pearson's correlation (r) and the intraclass correlation coefficient (ICC) between the scores from the first and second testing. Both testings were conducted by the same interviewer with a test-retest interval of 7-9 days. To evaluate interinterviewer reliability, we calculated the Pearson's correlation and the ICC between the scores from the two interviewers. In this case, both interviewers interviewed the caregiver separately on the same day.

We evaluated developmental sensitivity in two ways. First, we conducted a cross-sectional analysis by calculating the Pearson's correlation between the child's age (in months) and each DMC-II score. We also tested for differences in scores between each of the five age groups using ANOVA. Second, we conducted a longitudinal analysis examining the change in scores between the first and second interview for the subset of 72 children who were tested at two time points using paired t-tests.

We evaluated sensitivity to group differences by looking at the differences in scores between children with growth restrictions (stunting, wasting and underweight) and those without. Stunting was defined as length for age z-score (LAZ) < -2 SD below the mean according to the World Health Organization (WHO) growth standards (27). Wasting and underweight were defined as weight for height z-score (WHZ) and weight for age z-score (WAZ) < -2 SD

below the mean according to the WHO growth standards, respectively. For maternal education, children were classified as those whose mothers had no formal education versus those whose mothers had one or more years of formal education. We used t-tests to test for differences between group means. For the play materials and activities in the home, we calculated the Spearman's rank correlations with the DMC-II scores, following Hamadani et al. (25).

Results

Reliability

The internal consistency of all three subscales and the total scale was >0.7 (Table 1), providing evidence that all items in each subscale measured a unitary construct. The interinterviewer reliability for each score ranged from 0.81 to 0.93 (Table 1), demonstrating that the interviewers asked the questions consistently and caregivers responded consistently to different interviewers. The agreement between interviewers for the nine obligatory observation items was also high (97% agreement). The test-retest reliability for each score ranged from 0.77 to 0.96 (Table 1), demonstrating that scores were consistent when caregivers were interviewed again after a period of 7–9 days.

Developmental sensitivity

The cross-sectional analysis showed that each DMC-II subscale score was strongly related to child age, with both Pearson correlations and ANOVAs resulting in highly significant results. For the motor score, the correlation with age in months was $r(284) = 0.80$ ($p < 0.0001$), for the language score it was $r(284) = 0.80$ ($p < 0.0001$) and for the personal-social score it was $r(284) = 0.74$ ($p < 0.0001$). The differences in means between the five age groups (12, 15, 18, 21 and 24 months) are presented in Figure 2. The ANOVAs revealed the following results: for the motor score $F_{4,281} = 150.3$ ($p < 0.0001$), for the language score $F_{4,281} = 134.6$ ($p < 0.0001$) and for the personal-social score $F_{4,281} = 93.9$ ($p < 0.0001$). The longitudinal analysis also showed that scores within children increased significantly with age (Table 2).

Sensitivity to group differences

Of the 1123 children in the sample, 361 (32%) were stunted, 118 (11%) were wasted, 257 (23%) were underweight and 155 (14%) had mothers with one or more years of education. Children who were stunted, wasted or underweight scored significantly lower on all three subscales compared to children who were not stunted, wasted or underweight, respectively (Table 3). Children of mothers with no education scored significantly lower on the language subscale only. Of the three subscales, the motor score showed the largest differences between groups for the three nutritional indicators, with effect sizes 0.69 *SD*, 0.87 *SD* and 0.81 *SD* for stunting, wasting and underweight, respectively, while the effect sizes for language and personal-social development ranged from 0.31 to 0.41. The Spearman's rank correlations with the variety of toys in the home were $r(1119) = 0.41$ for the motor score, $r(1119) = 0.32$ for the language score and $r(1119) = 0.27$ for the personal-social score ($ps < 0.0001$). The correlations with activities with caregivers in the past 3 days were $r(1119) =$

0.19 for the motor score, $r(1119) = 0.15$ for the language score and $r(1119) = 0.13$ for the personal-social score ($ps < 0.0001$).

Discussion

We created the DMC-II by expanding the DMC, previously developed in Kenya, to a new context in Burkina Faso. Specifically, we added and modified a number of items, changed the administration procedure to permit data collection by observation as well as caregiver reports and developed a pictorial flip chart to help illustrate the meaning of the questions. Our results indicate that these modifications were implemented without changing the initial strengths of the DMC. The DMC-II remained easy to quickly implement (10–20 min per caregiver) by trained fieldworkers with no previous experience in developmental assessment.

The added element of observation was useful in several ways. First, the observation, as well as the flip chart, aided the data collectors in prompting the caregiver's memory and ensuring that she understood each question clearly. Second, the inclusion of observation as well as caregiver response allowed us to expand the item pool beyond skills for which all caregivers could easily recall their child's ability. For example, a caregiver might not know if a child can scribble with a pen or throw a ball if these play materials are not available in the household. However, we expect children to develop the motor coordination to do these activities even if they have not previously attempted the activity. Through observation, we were able to determine whether the child could do the skill even if he or she had never attempted it before. We also administered an interview to gather information concerning the types of play materials available in the home as well as other aspects of the home environment, which is important information for the interpretation and analysis of developmental scores.

Moreover, the DMC-II retained good psychometric properties, similar to those reported for the DMC. We observed adequate internal, test–retest and interinterviewer reliability, which are a prerequisite of a good standardised measure (28). The high internal consistency of the 76 item scores indicates that the three subscales form a unitary construct 'developmental level' providing justification for use of both a total score and the subscale scores. Correlation with age and progression in scores as children grow older indicates the sensitivity of the DMC-II to developmental changes and shows that the DMC-II may be a valuable tool in monitoring child development over time. Future work needs to investigate the possibility of incorporating this tool into routine monitoring strategies for children in Burkina Faso and other similar settings.

Due to the absence of a locally appropriate gold standard developmental assessment or diagnosis of developmental disorders, we could not evaluate many aspects of validity. However, we evaluated the extent to which the DMC-II is sensitive to true group differences, in this case, whether it can discriminate between children with and without compromised anthropometric status. We observed that children with compromised anthropometric status had lower scores, on average, which is consistent with what has been reported for the earlier DMC and also with other measures in Africa (5,8). Moreover, the effect sizes we found were similar to the effect sizes reported in a recent study which used the Bayley Scales of Infant

Development 2nd edition (BSID-II) to assess motor development in children under the age of 2 years in Tanzania (29). The BSID-II motor scale yielded effect sizes of 0.47 for stunting, 0.89 for wasting and 0.85 for underweight, while the DMC-II motor scale yielded effect sizes of 0.69, 0.87 and 0.81, respectively.

The DMC-II scores also showed the expected relations with the variety of play materials in the home and activities with caregivers. A recent study in Bangladesh analysed the correlations between these same measures and the BSID-II scores. The Spearman's rank correlation between the BSID-II motor score and the variety of play materials was 0.20 and with play activities was 0.19, while the DMC-II motor score resulted in correlations of 0.41 and 0.19, respectively. The Spearman's rank correlation between the BSID-II mental score and the variety of play materials was 0.27 and with play activities was 0.29, while the DMC-II language score resulted in correlations of 0.32 and 0.15, respectively. Thus, for environmental variables, just as for markers of undernutrition, the DMC-II yielded effect sizes similar to the BSID-II, which is a tool designed for children in high-income countries and which requires more extensive tester training and time for administration than the DMC-II.

While children of mothers with no education had slightly lower DMC-II motor and language scores compared to those with at least 1 year of formal education, only the language score was significantly different between groups. This is consistent with the results from the earlier analysis of the DMC and other results from Africa (9,20,21). In resource poor settings, additional maternal education and knowledge may have limited impact on child development if households lack access to the necessary resources to enhance child development. The weak effect of maternal education may also be explained by the generally low levels of education in this sample. The average number of years of education was only 0.6 years. Of the 1123 mothers in this study, only 155 (14%) had been to school and only 19 (2%) had attended school beyond grade six. A few years of elementary education may not be enough for the emergence of an association between maternal education and children's developmental attainment.

One limitation of our study was the small size of the sample in which test-retest reliability and interinterviewer reliability were evaluated ($n = 16$). The large study area and the heavy workload of the data collectors made it difficult to conduct a large number of repeat visits in this study. Future work should examine test-retest and interinterviewer reliability of the DMC-II in a larger number of children and in children at different ages.

Our report not only presents the results of the development of a new measure, but also highlights some of the procedural considerations that must be taken into account when one is developing measures for use in a context where there have previously been no or limited psychological assessments. Consistent with other studies (5,8,9,30), we illustrate how the use of systematic adaptation procedures leads to the development of measures that are culturally and contextually relevant with sound psychometric properties.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Kathryn Dewey, iLiNS Project Director, supervised the postdoctoral work of Elizabeth Prado, including this study. The iLiNS-Zinc study was designed and executed by the iLiNS-Zinc team, led by Kenneth Brown, Steve Vosti, Sonja Hess, and Jean-Bosco Ouedraogo. Rosemonde Guissou contributed to the coordination of the iLiNS-ZINC project. Vincent Gnimassou, Déborah Siry, Mariâtre Rayaisse and Elise Kéré conducted the pilot test and collected the DMC-II and anthropometric data. Bernadin Sanou, Aminata Léga, Brigitte Yé and Aicha Diallo conducted the family care indicators interview. Janet Peerson performed the randomisation for sample selection. Penny Holding was involved in the development of the original DMC and advised on the development of the DMC-II. This publication is based on research funded by a grant to the University of California, Davis, from the Bill & Melinda Gates Foundation. The findings and conclusions contained within are those of the authors and do not necessarily reflect positions or policies of the Bill & Melinda Gates Foundation.

Abbreviations

BSID	Bayley scales of infant development
DMC	Developmental milestones checklist
LAZ	Length for age z-score
iLiNS	International lipid-based nutrient supplements project, comprised of four randomised trials in three countries
iLiNS-Zinc	International lipid-based nutrient supplements project in Burkina Faso, one of four randomised trials in the iLiNS project
ICC	Intraclass correlation coefficient
UNICEF	United nations children's fund
WAZ	Weight for age z-score
WHZ	Weight for height z-score

References

1. Holding PA, Taylor HG, Kazungu SD, Mkala T, Gona J, Mwamuye B, et al. Assessing cognitive outcomes in a rural African population: development of a neuropsychological battery in Kilifi District, Kenya. *J Int Neuropsychol Soc.* 2004; 10:246–60. [PubMed: 15012845]
2. van de Vijver FJR, Tanzer NK. Bias and equivalence in cross-cultural assessment: an overview. *Eur Rev Appl Psychol.* 1997; 47:263–79.
3. Greenfield PM. You can't take it with you: why ability assessments don't cross cultures. *Am Psychol.* 1997; 52:1115–24.
4. Malda M, Van de Vijver DA, Temane QM. Rugby versus soccer in South Africa: content familiarity contributes to cross-cultural differences in cognitive test scores. *Intelligence.* 2010; 38:582–95.
5. Gladstone M, Lancaster GA, Umar E, Nyirenda M, Kayira E, van den Broek NR, et al. The Malawi Developmental Assessment Tool (MDAT): the creation, validation, and reliability of a tool to assess child development in rural African settings. *PLoS Med.* 2010; 7:e1000273. [PubMed: 20520849]

6. Kambalmetore S, Hartley S, Lansdown R. An exploration of the Malawian perspective on children's everyday skills: implications for assessment. *Disabil Rehabil.* 2000; 22:802–7. [PubMed: 11194621]
7. Vierhaus M, Lohaus A, Kolling T, Teubert M, Keller H, Fassbender I, et al. The development of 3- to 9-month-old infants in two cultural contexts: bayley longitudinal results for Cameroonian and German infants. *Eur J Dev Psychol.* 2011; 8:349–66.
8. Abubakar A, Holding PA, Van Baar A, Newton CRJC, Van de Vijver FJR. Monitoring psychomotor development in a resource-limited setting: an evaluation of the Kilifi Developmental Inventory. *Ann Trop Paediatr.* 2008; 28:217–26. [PubMed: 18727851]
9. Abubakar A, Holding P, van de Vijver FJ, Bomu G, Van Baar A. Developmental monitoring using caregiver reports in a resource-limited setting: the case of Kilifi, Kenya. *Acta Paediatr.* 2010; 99:291–7. [PubMed: 20353499]
10. Fenson, L, Marchman, VA, Thal, D, Dale, PS, Reznick, JS, Bates, E. *The MacArthur-Bates communicative development inventories, second edition: user's guide and technical manual.* Baltimore: Paul H. Brookes Publishing Co; 2007.
11. Glascoe FP. Screening for developmental and behavioral problems. *Ment Retard Dev D R.* 2005; 11:173–9.
12. Abubakar A, van de Vijver FJ, Mithwani S, Obiero E, Lewa N, Kenga S, et al. Assessing developmental outcomes in children from Kilifi, Kenya, following prophylaxis for seizures in cerebral malaria. *J Health Psychol.* 2007; 12:417–30. [PubMed: 17439993]
13. Gladstone M, Lancaster GA, Jones AP, Maleta K, Mtitimila E, Ashorn P, et al. Can Western developmental screening tools be modified for use in a rural Malawian setting? *Arch Dis Child.* 2008; 93:23–9. [PubMed: 17379661]
14. Griffiths, R. *The abilities of young children.* High Wycombe, UK: The Test Agency; 1976.
15. Sparrow, S, Balla, D, Cicchetti, D. *Vineland adaptive behavior scales.* Circle Pines, MN: AGS Publishing; 1984.
16. Grantham-McGregor S. A review of studies of the effect of severe malnutrition on mental development. *J Nutr.* 1995; 125:2233S–8S. [PubMed: 7542705]
17. Grantham-McGregor S, Baker-Henningham H. Review of evidence linking protein and energy to mental development. *Public Health Nutr.* 2005; 8:1191–201. [PubMed: 16277829]
18. Duncan G, Brooks-Gunn J, Klebunov P. Economic deprivation and early childhood development. *Child Dev.* 1994; 65:296–318. [PubMed: 7516849]
19. Bradley RH, Corwyn RF. Caring for children around the world: a view from HOME. *Int J Behav Dev.* 2005; 29:468–78.
20. Aboud FE, Alemu T. Nutrition, maternal responsiveness and mental development of Ethiopian children. *Soc Sci Med.* 1995; 41:725–32. [PubMed: 7502104]
21. Richter LM, Grieve KW. Home environment and cognitive development of black infants in impoverished South African families. *Infant Ment Health J.* 1991; 12:88–102.
22. Zeba AN, Sorgho H, Rouamba N, Zongo I, Rouamba J, Guiguemde RT, et al. Major reduction of malaria morbidity with combined vitamin A and zinc supplementation in young children in Burkina Faso: a randomized double blind trial. *Nutr J.* 2008; 7:7. [PubMed: 18237394]
23. Kariger P, Frongillo EA, Engle P, Britto PM, Sywulka SM, Menon P. Indicators of family care for development for use in multicountry surveys. *J Health Popul Nutr.* 2012; 30:472–86. [PubMed: 23304914]
24. Fernald LC, Kariger P, Hidrobo M, Gertler PJ. Socioeconomic gradients in child development in very young children: evidence from India, Indonesia, Peru, and Senegal. *Proc Natl Acad Sci USA.* 2012; 109(Suppl 2):17273–80. [PubMed: 23045688]
25. Hamadani JD, Tofail F, Hilaly A, Huda SN, Engle P, Grantham-McGregor SM. Use of family care indicators and their relationship with child development in Bangladesh. *J Health Popul Nutr.* 2010; 28:23–33. [PubMed: 20214083]
26. Raghunathan TE, Lepkowski JM, Van Hoewyk J, Solenberger P. A multivariate technique for multiply imputing missing values using a sequence of regression models. *Surv Methodol.* 2001; 27:85–95.

27. WHO Multicentre Growth Reference Study Group. 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; 2006.
28. Kline, P. The handbook of psychological testing. London: Routledge; 1993.
29. McDonald CM, Manji KP, Kupka R, Bellinger DC, Spiegelman D, Kisenge R, et al. Stunting and wasting are associated with poorer psychomotor and mental development in HIV-exposed tanzanian infants. *J Nutr.* 2013; 143:204–14. [PubMed: 23256148]
30. Prado EL, Hartini S, Rahmawati A, Ismayani E, Hidayati A, Hikmah N, et al. Test selection, adaptation, and evaluation: a systematic approach to assess nutritional influences on child development in developing countries. *Br J Educ Psychol.* 2010; 80:31–53. [PubMed: 19772714]

Key notes

- Finding solutions to prevent developmental delays caused by malnutrition and other conditions prevalent in Sub-Saharan Africa requires developmental assessment tools appropriate for the region.
- We expanded the Developmental Milestones Checklist, previously developed in Kenya, to a new West African context without changing its strengths.
- The DMC-II demonstrated high reliability and sensitivity to age, malnutrition and environmental variables and trained fieldworkers with no previous experience found it quick and easy to implement.



Figure 1.
An example of a flip chart picture that supported the question ‘Does your child hold a pen in any way with the intention to draw or write?’

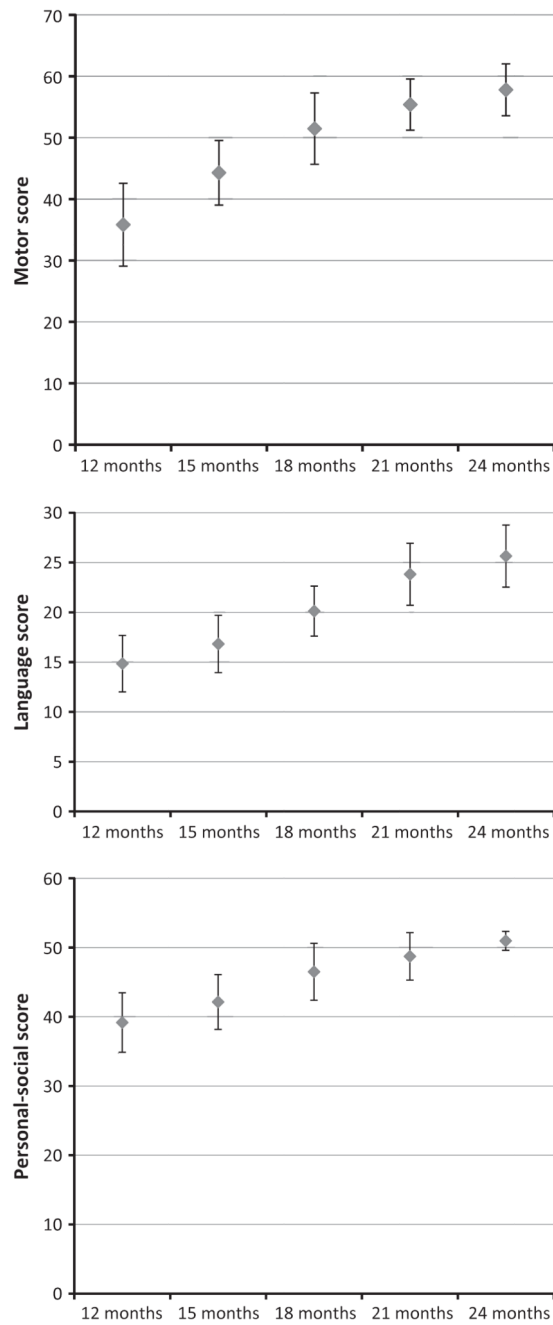


Figure 2. Mean score for each age group for each DMC-II subscale. Error bars show the standard deviation of the mean.

Table 1
Internal consistency, interinterviewer reliability and test–retest reliability of each DMC-II score

	Internal consistency		Interinterviewer reliability			Test–retest reliability		
	n	Cronbach's alpha	n	r	ICC	N	r	ICC
Motor	1123	0.84	15	0.88	0.88	16	0.96	0.95
Language	1123	0.70	15	0.85	0.81	16	0.77	0.77
Personal-social	1123	0.74	15	0.91	0.90	16	0.93	0.91
Total score	1123	0.88	15	0.93	0.93	16	0.96	0.96

Table 2
Increase in scores within children tested at two time points

	12 and 18 months n = 16		15 and 18 months n = 22		18 and 21 months n = 20		18 and 24 months n = 14	
	Mean increase (SD)	t	Mean increase (SD)	t	Mean increase (SD)	t	Mean increase (SD)	t
Motor	17.3 (6.3)	10.9 ^{***}	6.9 (5.4)	6.0 ^{***}	5.5 (6.8)	3.6 ^{**}	7.1 (6.9)	3.9 ^{**}
Language	4.9 (2.2)	8.7 ^{***}	3.5 (2.6)	6.3 ^{***}	4.3 (3.6)	5.4 ^{***}	5.7 (3.3)	6.4 ^{***}
Personal-social	7.9 (5.3)	5.9 ^{***}	5.1 (4.3)	5.6 ^{***}	3.3 (3.5)	4.2 ^{***}	5.1 (5.1)	3.8 ^{**}

^{***}
p < 0.001.

^{**}
p < 0.01.

Table 3
Differences in DMC-II subscale scores based on stunting, wasting, underweight and maternal education in children age 18.3 (+/- 0.4) months

	n	Raw score mean (SD)	z-score mean (SD)	n	Raw score mean (SD)	z-score mean (SD)	t-value
Stunting (LAZ <-2)	Not stunted			Stunted			
Motor	762	52.4 (4.6)	0.22 (0.76)	361	48.2 (7.8)	-0.47 (1.28)	11.25***
Language	762	20.5 (2.8)	0.12 (0.94)	361	19.5 (3.2)	-0.24 (1.08)	5.64***
Personal-social	762	46.7 (4.0)	0.10 (0.92)	361	45.4 (4.9)	-0.21 (1.13)	4.89***
Wasting (WHZ <-2)	Not wasted			Wasted			
Motor	1005	51.6 (5.3)	0.09 (0.86)	118	46.2 (10.0)	-0.78 (1.63)	9.20***
Language	1005	20.3 (2.9)	0.04 (0.96)	118	19.2 (3.8)	-0.34 (1.25)	3.92***
Personal-social	1005	46.4 (4.1)	0.04 (0.95)	118	45.1 (5.8)	-0.27 (1.33)	3.15**
Underweight (WAZ <-2)	Not underweight			Underweight			
Motor	866	52.1 (4.7)	0.18 (0.77)	257	47.2 (8.6)	-0.62 (1.40)	11.93***
Language	866	20.5 (2.8)	0.09 (0.94)	257	19.3 (3.4)	-0.31 (1.14)	5.79***
Personal-social	866	46.6 (4.1)	0.07 (0.94)	257	45.3 (5.1)	-0.23 (1.17)	4.35***
Maternal education	One or more years			No maternal education			
Motor	155	51.7 (5.5)	0.11 (0.90)	963	50.9 (6.3)	-0.02 (1.02)	1.54
Language	155	20.7 (3.0)	0.16 (0.99)	963	20.1 (3.0)	-0.02 (1.00)	2.07*
Personal-social	155	46.2 (4.1)	-0.02 (0.96)	963	46.3 (4.4)	0.00 (1.01)	0.29

p < 0.001.

**
p < 0.01.

*
p < 0.05.