

RESEARCH ARTICLE

Open Access



# Adaptation and standardization of a Western tool for assessing child development in non-Western low-income context

Teklu Gemechu Abessa<sup>1,7\*†</sup>, Berhanu Nigussie Worku<sup>1,7</sup>, Mekitie Wondafrash Kibebew<sup>2,6</sup>, Jan Valy<sup>3</sup>, Johan Lemmens<sup>3</sup>, Herbert Thijs<sup>4</sup>, Wondwosen Kasahun Yimer<sup>5</sup>, Patrick Kolsteren<sup>6</sup> and Marita Granitzer<sup>7†</sup>

## Abstract

**Background:** Due to lack of culturally relevant assessment tools, little is known about children's developmental profiles in low income settings such as Ethiopia. The objective of this study was to adapt and standardize the Denver II for assessing child development in Jimma Zone, South West Ethiopia.

**Methods:** Culture-specific test items in Denver II were modified. After translation into two local languages, all test items were piloted and fine-tuned. Using 1597 healthy children 4 days to 70.6 months of age, the 25, 50, 75 and 90 % passing ages were determined for each test item as milestones. Milestones attainment on the adapted version and the Denver II were compared on the 90 % passing age. Reliability of the adapted tool was examined.

**Results:** A total of 36 (28.8 %) test items, mostly from personal social domain, were adapted. Milestones attainment ages on the two versions differed significantly on 42 (34 %) test items. The adapted tool has an excellent inter-rater on 123 (98 %) items and substantial to excellent test-retest reliability on 119 (91 %) items.

**Conclusions:** A Western developmental assessment tool can be adapted reliably for use in low-income settings. Age differences in attaining milestones indicate a correct estimation of child development requires a population-specific standard.

**Keywords:** Adaptation, Child development, Denver II-Jimma, Developmental assessment tool

## Background

Despite substantial child mortality reduction in Sub-Saharan Africa, many children under-five are still developmentally at risk because of poverty and related risk factors such as malnutrition, poor health and unstimulating home environments [1]. The magnitude of developmental problems is, however, unknown due to lack of culturally relevant tools for assessing development. In the absence of such tools, it is also difficult to correctly determine the developmental effects of interventions

targeting children at risk. In rare studies conducted on children at developmental risk, researchers have used tools originally created for technological societies of Europe and North America by either translating or adapting them with little validation [2–6]. Sometimes culture specific test items were totally dropped [7–11] or no adaptation was made [12–14]. Among a Western tool adapted and used worldwide is the American Denver Developmental Screening Test [15] or its revised version, the Denver II [16]. The Denver II is a revised version of the Denver Developmental Screening Test developed in 1967. It was standardized in 1989 on 2,096 American children and published in 1992. It is a screening tool used to identify children between birth and six years who have problems in personal-social (self-help skills and socialization with others), problems in fine

\* Correspondence: teklogemechu.abessa@uhasselt.be; teklogem@yahoo.com

†Equal contributors

<sup>1</sup>Department of Psychology, College of Behavioral Sciences and Education, Jimma University, Jimma, Oromia, Ethiopia

<sup>7</sup>REVAL Rehabilitation Research Center, Biomedical Research Institute, Faculty of Medicine and Life Sciences, University of Hasselt, Hasselt, Belgium

Full list of author information is available at the end of the article



motor (eye-hand co-ordination, and manipulation of small objects), problems in language (production of sounds, ability to recognize, understand, and use language), and problems in gross motor (large muscle movements such as sitting, walking, jumping). The Denver II has been used in other countries such as Georgia, Singapore and Sri Lanka by adapting and standardizing it [17–19]. Though it is a simple, quick and feasible to use at institution and home settings to identify children at developmental risks [20], Denver II has not been adapted and validated for use in many low income countries of Africa such as Ethiopia. An indigenous tool similar to it, in style, however, was created for children in Malawi [21]. By using Denver II as a prototype, new test items that were more culturally relevant for Malawian children were created from the Denver Developmental Screening Test, the Denver II and the Griffiths Mental Development Scales.

The main objective of this research, therefore, was to adapt and standardize the Denver II on children between birth and six years of age in the low income context of Jimma Zone of Ethiopia for a more realistic assessment of their development.

## Methods

### Study setting

The study was conducted in Jimma Zone, South West Ethiopia. Within this zone, the population was estimated to be 2.8 million. Jimma Town is the Zonal Capital having about 149, 166 inhabitants [22]. The town is home to more than nine ethnic and linguistic communities communicating mostly in a federal language, Amharic, and a regional language, Afan Oromo. With a mixture of both urban and rural life styles, Jimma town represents the diverse socio-economic, multicultural and multi-lingual Ethiopian society.

### Adaptation process of the Denver II

The Denver II [16] comprises 125 test items grouped into four domains of child development: 25 personal-social (PS), 29 fine motor (FM), 39 language (LA) and 32 gross motor (GM). These test items are administered using a bell, glass bottle, set of 10 blocks, rattle, pencil, tennis ball, yarn, raisins, cup, white doll, white paper, and baby bottle. Adaptations involved identifying culture specific test items, test objects or materials and then modifying or replacing them to make them culturally relevant. In some cases, instructions for test item administration and criteria of passing were modified.

### Classifying test items under 'cross-cultural' and 'culture-specific' categories

All test items were first categorized into culture-specific and cross cultural items. Cross cultural relevance of tasks

in the test items was assessed using International Classifications of Functions [23]. Culture specific items related to movements (e.g. running, jumping, hopping) were identified using taxonomy of movement skills [24]. Other specific movement skills related to sport, complex movement skills and functional movement skills such as activities of daily living, work, and games are culture-specific. Cross-culturality of items other than movements was assessed using cross cultural psychology [25]. Within this process a local team (psychologists, a special educator and pediatricians) and a Belgian team (child psychiatrist, a pediatrician/nutritionist, a neuroscientist, a physiotherapist and occupational therapists) worked together.

After translations into the dominant languages (Amharic and Afan Oromo) dialect appropriateness was checked.

### Pilot studies and draft versions

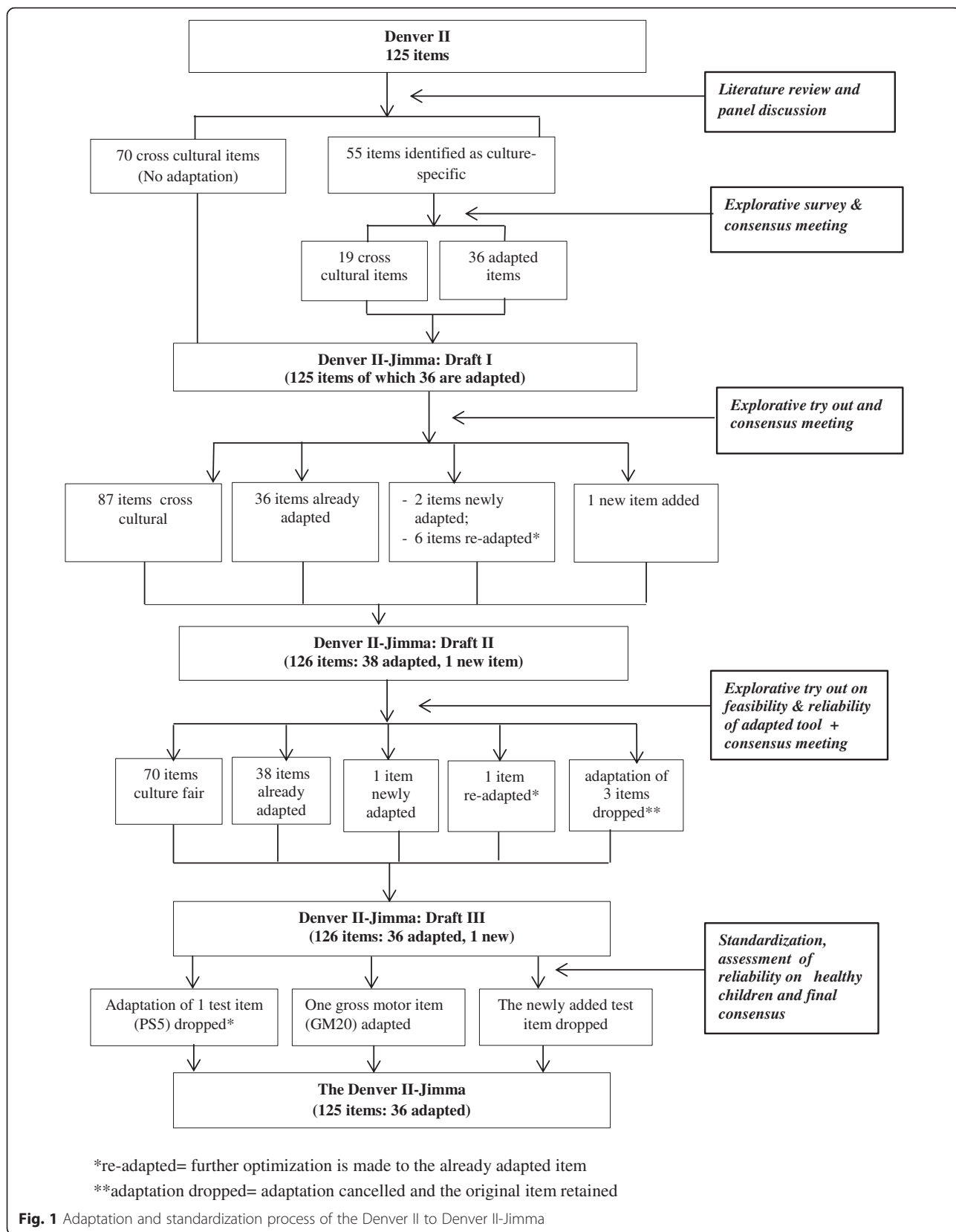
The test items were then piloted on apparently healthy children of accessible parents who consented orally to participate in the study. Draft I emerged based on a survey conducted in 2009 on 19 households. Four urban and 15 rural families were interviewed about the items which were identified as culture-specific (see Fig. 1). Draft I was tried out on eight urban kindergarten children (26–60 months of age; mean = 42.9; SD = ±14.1). Three local study team members, trained in Denver II test item administration, did the testing and the problematic items were discussed at the multidisciplinary team meeting. Re-adaptations resulted in Draft II which was further explored in 2010 for feasibility and reliability on 24 urban kindergarten children (mean age = 51.4 months, SD ± 8.2 months). Testing was conducted by seven trained kindergarten teachers. Further adaptation resulted in Draft III. Figure 2 summarizes the adaptation process.

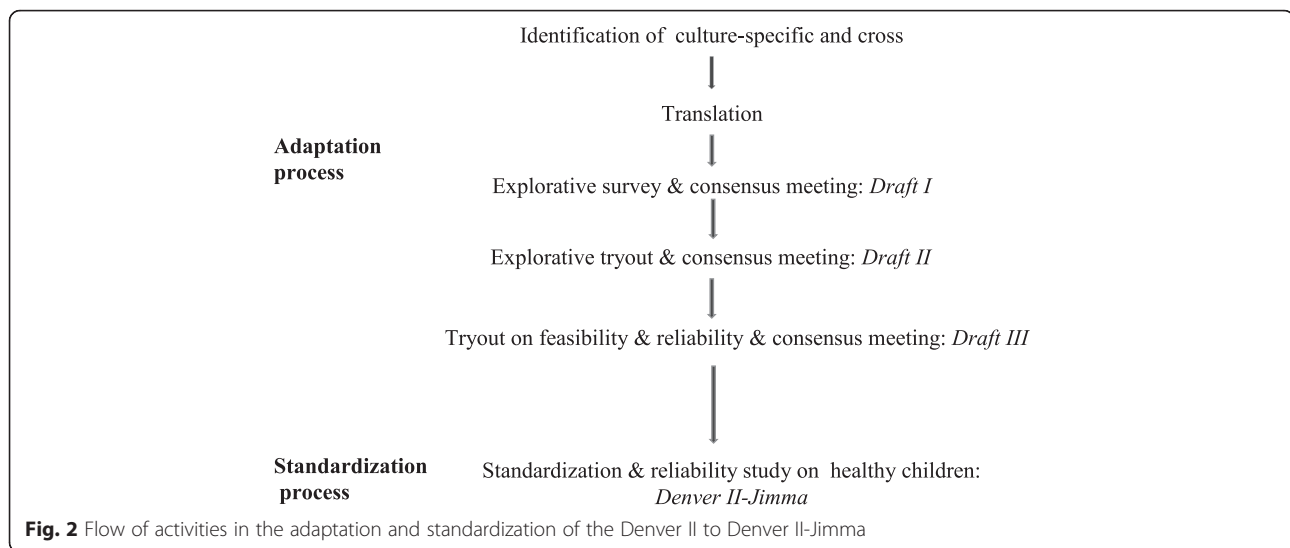
### Large sample study and standardization

#### Sampling, inclusion and exclusion criteria

Trained nurses collected data using the third draft. Under-six children in Jimma town whose parents could afford to pay preschool education fees were targeted. Such children were assumed to belong to middle or higher socioeconomic level and thus in a context for optimal development. Quota sampling was used to include children in the following age categories (in months): 0–2, 3–8, 9–14, 15–20, 21–26, 27–32, 33–41, 42–53, and 54–65.

Before testing a target child, the mother was interviewed using a 10-point checklist which listed the exclusion criteria. Children, whose mother reported the presence of any of the following potential developmental risks were excluded: prematurely born, birth weight less than 2500 g, very tiny body at birth, instrumentally delivered, or delivered after 24 h of labor, born twins or





triplets, born with a chronic health problem, sick during the first year after birth, having observable impairments affecting sight or/and hearing, or/and mobility, having a mother who was seriously sick during pregnancy. Besides, anthropometric measurements were made to assess the nutritional status and exclude malnourished children. Weight was measured with a calibrated digital weighing scale; mid-upper-arm-circumference (MUAC) with a MUAC tape. Anthropometric indices related to length/height were not used for fear of measurement inaccuracy as some children were nervous while positioning them for measurement. Earlier studies have also used weight-for-age to determine child's nutritional status [7, 26] because the weight-for-age is considered as more comprehensive than the height-for-age [27]. Assessment was done (if the child was well) in the following sequence: developmental assessment, measuring weight, MUAC.

We dispatched questionnaire and study consent form to parents of private kindergarten attending children in Jimma town. The homes of parents who signed the consent form were visited. From 3502 children, only 1682 (mean age = 31.2, SD = 17.75 in months) who were eligible according to the inclusion criteria were tested. The age of the children ranges from four days to 73.3 months. Initially, 1552 children were tested at home from 11 January to 21 June 2011 and later, 130 children of lower ages (<10 months) were added. Two children of unknown nutritional status and eighty-three malnourished children were excluded during analysis based on weight for age Z-score (WAZ)  $\leq -2$ , or mid-upper-arm-circumference Z-score (MUACZ)  $\leq -2$  when WAZ was absent.

The study complied with the Helsinki Declaration [28] and was reviewed and approved by Ethical Clearance Board of Jimma University, Ethiopia, and Comité voor Medische Ethiek Universiteit Hasselt, Belgium. Written and oral consents of parents were obtained

and children were always tested in the presence of caregivers.

#### Assessment of feasibility and reliability

Feasibility of each test item (meaningfulness of test items, their practicality and ease of administration) was documented during data collection and discussed at final consensus meeting. Inter-rater and test-retest reliabilities were assessed for each test item. Ten female clinical nurses worked in pairs alternately as a tester or an observer. Independent scores were generated for each child by a tester and an observer. These scores by testers and observers were calculated as percentages of agreement to determine the reliability of the test items. Inter-rater reliability was tested on 409 children. Within an average interval of 14 days, 147 of them were tested for test-retest reliability. Inter-rater reliability was not calculated during a re-test condition.

#### Test item administration and scoring system

Test item administration and scoring is the same as in Denver II manual [29]. Each test item on Denver II is presented on a chart by a horizontal bar partitioned into 25, 50, 75 and 90 percentile ages of passing the items. To test a child, his or her age is calculated and a vertical age-line is drawn on the II chart. The testing starts from a test item completely to the left of the age-line. All test items passed by 75 % or more children of same age in the norming sample and by lower ages are counted for a child as *expected passes*. If a child passes three consecutive test items arranged on Denver II test chart, all items to the left are assumed to be passed because they are items achieved at a lower age. These items are called *implied passes*. If a child fails three consecutive test items, it is assumed that all other items arranged to the right on the Denver II chart are failed. These items are *implied failures*. Items passed

by a child through testing are *tested passes*. Implied passes and tested passes are added up as *actual passes*. A child's raw score on each test item is marked as tested pass, implied pass, tested failure, implied failure, refusal, or no opportunity. Categorical and numerical scores were derived for statistical analysis.

**Categorical score:** For each test item, a binary outcome variable (pass/fail) was created: *pass* (tested pass items) and *fail* (tested failure and refusal). “No opportunity” to perform the item, “implied passes” and “implied failures” were treated as missing values.

**Numerical score:** The ratio of *actual passes* to the *expected passes* was calculated as a *performance ratio* score.

#### Standardization

The objective of the standardization was to determine the ages at which 25, 50, 75 and 90 % of the children pass each of the adapted test items using binary logistic regression.

#### Data management and statistical analysis

Data within the adaptation process (except for reliability) were analyzed qualitatively. Whether or not a test item was culture-specific or cross-cultural was analyzed using theoretical information and discussion among the research team. Data collected during drafting and re-drafting were discussed at interdisciplinary team meetings comprising local and western professionals. Standardization data were entered into Epi-Data 3.1, double checked, cleaned and exported to SAS 9.3 and STATA 12.1 for analysis. The WAZ and MUACZ scores were calculated as anthropometric indices using WHO Anthro and AnthroPlus and children's nutritional status determined against WHO reference standard [30].

Predicted ages at which 25, 50, 75 and 90 % of the norming sample passed each test item were derived from the models and calculated as milestone ages. Using the categorical score “pass/fail”, binary logistic regression model was fitted for each test item by entering child age in days as a single covariate. Predicted probabilities of passing were calculated from alpha and beta coefficients. Goodness of the fit was assessed using Hosmer and Lemeshow test statistic at 5 % level of significance. Items with poor model fit ( $p$ -value <0.05) were refitted using cubic splines [31].

Age of attaining milestones by Denver II and Denver II-Jimma norming samples were compared on 90 percentile age. More than 10 % difference was considered clinically significant.

Reliability was assessed at item and domain levels. Item reliability was calculated as a percentage of agreement between a tester score and an observer score (inter-rater), and between the first test and re-test scores (test-retest) for the same child. Chance agreements were corrected using Cohen's kappa. Kappa values by Landis and Koch [32] were used for interpretation: value below 0.20 as slight; between 0.21 and 0.40 as fair; between 0.41 and 0.60 as moderate, between 0.61 and 0.80 as substantial, and between 0.81–1.00 as excellent agreement. Where kappa could not be calculated, percentages of agreement for events were determined: 70 % or higher was considered as acceptable.

Domain reliability was evaluated using intra-class correlation coefficients. First, performance ratio scores were generated for each of the four domains separately. Then, the correlations between tester and observer performance ratio scores at two testing moments (test and retest) were computed for each domain as inter-rater and test-retest intra-class correlation.

## Results

### Outcome of the adaptation

Of the 125 Denver II test items, 55 (20 PS, 18 FM, 15 LA, 2 GM) were theoretically identified as culture-specific. These 55 items were piloted through exploratory survey and discussed at a consensus meeting. Only 36 of them needed adaptation. The other 19 items were retained as was in the original (Fig. 1). A tryout revealed difficulties with eight (6 LA and 2 PS) test items. Further fine-tuning resulted in Draft II (36 adapted, 1 newly added, 89 original Denver II items). Inter-rater reliability of Draft II was excellent (kappa >0.83) for all tested items. For items with skewed data distribution kappa could not be computed. Their percentages of agreement, however, were all acceptable (71.4 to 95.2).

Some test items were found practically difficult to administer or still difficult for children to understand even after initial adaptation. Hence, to make sure that test items were feasible to administer, understandable for children and caregivers, further adaptations were made. One item from PS was adapted, and another re-adapted; and the adaptations of three LA items were dropped. This resulted in the Denver II-Jimma-Draft III, which comprises 36 adapted (18 personal social, 10 fine motor, 8 language), 1 newly added (toilet going), and 89 original Denver II items.

At the final consensus meeting following the standardization study, one gross motor (walk up steps) was adapted, the newly added item was dropped, and adaptation of one personal social item was dropped. This resulted in the final Denver II-Jimma having 36 adapted items (Table 1).

**Table 1** Descriptions of adaptation made to the Denver II test items to make Denver II- Jimma

Denver II test items adapted	Item code	Description of adaptation of the item
Work for toy <sup>a</sup>	PS5	<i>Rattle or culturally used tools such as small pebbles "calle" or "elela" were selected.</i>
Feed self	PS6	The items "cracker", "cookie" or "any finger food" were replaced by locally used food such a piece of "bread", "cake", "biscuit", "injera" or a piece of sugar-cane.
Play pat-a-cake	PS7	Replaced by "play a clapping game": a culturally equivalent game played by clapping hands.
Wave bye-bye	PS9	Modified Cultural difference in expressing goodbye: "saying" or "waving" goodbye
Imitate (household) activities	PS11	Activities such as "vacuuming", or "talking on the phone" were replaced by activities such as ...."cooking" or "washing clothes".
Drink from cup	PS12	....child can hold a regular cup or glass and drink from it without help replaced by a regular "cup or glass" or any suitable container used in the family.
Use spoon/fork	PS14	Modified as "eat using hand or spon/fork. The child uses a spoon or fork to eat. is modified as "the child is able to eat independently by using his/her fingers or a spoon or fork"
Remove garment	PS15	... items such as ..."jackets", "pants", are modified as... items such as "blouse", "dress" or "trousers"
Feed doll	PS16	The criterion to pass the test item " if the child places the bottle to the doll's mouth, or tries to put it to the mouth"... is modified as "if the child imitates putting food into the doll's mouth or if the child imitates breast feeding". The use of bottle feeding is being discouraged and many mothers do not practice. Since toy bottle was found strange for many children, performing the task is not expected to pass the test item .
Put on clothing	PS17	....clothing, such as "underpants", "socks" are modified as ... clothing, such as "blouse", "trousers", "dress", "skirt"
Brush teeth with help	PS18	Replaced by "wash mouth with help"... if the child "brushes his/her teeth with some help"...is modified as if the child "washes his/her mouth with some help"
Wash and dry hands	PS19	Criteria for passing the test item is modified as ... child can "wash both sides of hands properly" (because hands are culturally dried by dripping water off the hands). Use of soap and towels are not required to pass the item.
Name friend	PS20	Replaced by "name playmate"
Put on T-shirt	PS21	"pullover" ...is replaced by "T-shirt" or "blouse" ...
Dress, no help	PS22	First: "at least play-clothes" ...is modified as ..."his/her own clothes", because children may not have many alternative clothing. Finally: adaptation dropped
Play board/card games	PS23	Modified as "play social games" .... "board" or "card" games, ... "Candy Land" or "Old Maid"... is modified as "joins in simple "group games", like "hide and seek"
Brush teeth, no help	PS24	"brush teeth, no help" replaced by "clean face, no help". ... if the child "washes his/her own teeth" ...is replaced by... if the child "washes and dries" his/her face (eyes, nose, mouth and teeth) .... this is usually done by using water and fingers)
Prepare Cereal	PS25	First: ... if the child can prepare a bowl of cereal ..... is modified as ...if the child can prepare his/her own breakfast, including taking bread (or injera,...) from the shelf, taking a cup and pouring a liquid (water, milk, juice) in it. Finally: modified as "serve oneself 'injera'" (cultural food served with stew)
Toilet going <sup>b</sup>	PS26	First: Toilet going (new item): Ask the caregiver if the child can independently use latrine or other facilities available for the family. PASS if they report the child can independently use toilet or latrine or available facility Finally: item dropped
Regard raisin	FM6	Is replaced by "Regard coffee bean".
Rake raisin	FM9	Is replaced by "Rake coffee bean"
Thumb-finger grasp	FM12	... "raisin" which is used as an object for child to grasp is replaced by "coffee bean"
Scribbles	FM15	The instruction "...do not show him/her how to scribble"... is modified as... you may write your name using the pencil to let the child who have never ever seen a pencil before that it is something to write with
Dump raisin, demonstrated	FM16	Replaced by "dump coffee bean, demonstrated"
Copy 0	FM23	The instruction " you may show how to hold the pencil" is added to familiarize a child who has never seen a pencil before
Draw person – 3 parts, 6 parts	FM24 & FM 28	The instruction... "You may show how to hold the pencil"... is added to familiarize a child who has never seen a pencil before
Copy +	FM25	The instruction... "You may show how to hold the pencil"... is added familiarize a child who has never seen a pencil before



**Table 1** Descriptions of adaptation made to the Denver II test items to make Denver II- Jimma (*Continued*)

Copy □	FM29	The instruction... "You may show how to hold the pencil"... is added to familiarize a child who has never seen a pencil before
Body parts – 6	LA21	A testing object "white doll" replaced by a black or chocolate colour doll. The use of either a white doll or a chocolate colour doll based on need of child was suggested based on repeated observation.
Name colours – 1,4	LA27 & LA34	First: Criterion for passing test item modified as, " Child could pass the test if he/she refers to an object with the same colour: sky/water for blue, grass/tree for green, sun for yellow and blood for red. Finally: The "blue" and "green" colour cubes are replaced by "black" and "white" colour cubes.
Use of objects – 2, 3	LA28 & LA30	...."What is a 'pencil' used for?".... is replaced by ...."What is a 'bed' used for?"....
Define words – 5, 7	LA35 & LA39	The words "desk", "curtain", "lake", are replaced by the words "knife", "firewood", "river"
Opposites – 2	LA37	"If 'fire' is 'hot', 'ice' is..." was replaced by "if a stone is heavy, a feather/leaf is ....."
Walk up step <sup>c</sup>	GM20	Walk up steps where there are steps in homes, and/or walk up-ward on steep location/ climbs and passes over an elevated door-step

<sup>a</sup>The adaptation of this test item was dropped at a final consensus meeting

<sup>b</sup>Newly added test item removed at final consensus meeting

<sup>c</sup>The adaptation of this test item was added at a final consensus meeting

**Table 2** Characteristics of study participants

Characteristics	No. (%)	Characteristics	No. (%)
Children's (n = 1682)		Mother's (n = 1588)	
Sex		Ethnicity	
Male	833 (49.5)	Oromo	716 (45.1)
Female	849 (50.5)	Amhara	363 (22.9)
Nutritional status		Tigre	46 (2.9)
Normal range <sup>a</sup>		Gurage	140 (8.8)
Male	789 (46.9)	Dawuro	148 (9.3)
Female	808 (48)	Keficho	72 (4.5)
Malnourished <sup>b</sup>		Wolaita	30 (1.9)
Male	43 (2.6)	Others	62 (3.9)
Female	40 (2.4)	Missing/Unknown	11 (0.7)
Unknown status		Perceived socio-economic status	
Male	1 (0.06)	High	58 (3.7)
Female	1 (0.06)	Middle	1443 (90.9)
Mother's (n = 1588)		Low	65 (4.1)
Education level		Very low	0
Illiterate	147 (9.3)	Missing/unknown	22 (1.4)
Grades 1–8	546 (34.4)	Religion	
Grades 9–12	485 (30.5)	Islam	558 (35.1)
Certificate after grade 12	98 (6.2)	Orthodox Christian	725 (45.7)
Diploma	247 (15.6)	Protestant	271 (17.1)
Degree and above	55 (3.5)	Catholic	23 (1.4)
Missing/unknown	10 (0.6)	Others	6 (0.4)
		Missing	5 (0.3)

<sup>a</sup>(WAZ > -2 where both WAZ and MUACZ score are present; and MUACZ > -2 where WAZ score is missing); <sup>b</sup>(WAZ ≤ -2 where both WAZ and MUACZ score are present; and MUACZ ≤ -2 where WAZ score is missing)

WAZ Weight-for-age-Z-score, MUACZ Mid-upper-arm-circumference Z score

## Outcome of the standardization

### *Characteristics of the standardization sample*

Nearly equal number of boys and girls participated in the study. About 95 % of the caregivers rated themselves as belonging to middle or higher socio-economic standard.

The Oromo, as the largest ethnic group, seem to have been fairly represented (45.1 %). Only 9.3 % of mothers of children enrolled in the study are illiterate (Table 2).

### *The Denver II-Jimma Age Milestones*

Of the 126 test items separately fitted on logistic model, 66 items fitted well. Three items (PS1, LA1, GM1) could not be fitted because all tested children passed them. Fifty-seven items showed poor fit (13 PS, 18 FM, 11 LA, 15 GM). The model fitness for 39 of these were improved by refitting using cubic splines. On lots of test items, the Denver II-Jimma differed from Denver II on 50, 75 and 90 % ages of attaining milestones (Table 3).

The 90 % age of milestones attainment on Denver II-Jimma significantly differed from Denver II on 42 (33.6 %) items (9 PS, 6 FM, 15 LA, and 12 GM). Fifteen test items were attained at an earlier age and 27 items at a later age than they are achieved on the Denver II. The remaining 83 (66.4 %) milestones were achieved at a similar age (Table 3).

### *Reliability of the Denver II-Jimma*

Table 4 summarizes the results for the reliability of the Denver II-Jimma at individual test item and overall domain levels. Inter-rater reliability was excellent except for two test items which showed substantial agreement: ("PS5: work for toy", kappa = 0.74 and FM5: "follow 180 degrees", kappa = 0.78). Majority (above 90 %) of the test items have a substantial to excellent test-retest reliability. Only one test item (FM 8: "look for yarn", kappa = 0.33) showed unacceptable kappa values. The Denver II-Jimma also demonstrated very high intra-class correlations on all domains of development (Table 4).

### *Final consensus on Denver II-Jimma*

As bottle feeding is being discouraged in line with WHO's recommendation, it is agreed that the test item "Feed doll" should be administered without using a toy bottle. Local material "*Callee*" initially suggested to replace the object "rattle" for administering the item "work for toy" was so risky for babies because it is small and could be swallowed. Hence, the adaptation was dropped. A newly added test item ("toilet going") was found difficult to perform before the age of six years and was thus eliminated. A gross motor item "Walk up steps" was not possible to assess in homes lacking steps. In such cases, care givers were asked if a child is able to walk up-ward a steep position or cross elevated

doorstep. Hence, the Denver II-Jimma finally evolved as a-125-test item tool with 36 (28.8 %) adapted test items: 17 PS, 10 FM, 8 LA and one GM items.

## Discussion

In order to provide early intervention for children developmentally at risk, correct assessment of their developmental status is an essential first step. Since development is influenced by the sociocultural contexts, instrument assessing child development should take culture into account. The tools should also be psychometrically valid. While child development tools created in western cultural contexts are psychometrically valid, they may not be culturally relevant to use with African children. Many agree that culturally relevant developmental assessment tools should be either created [33] or adapted from tools developed in other cultures [5, 34]. Adapting an existing tool is less expensive and more suitable to maintain construct validity of a tool across different settings.

In this study the Denver II created in the Western socio-cultural context, was adapted and standardized on Ethiopian children in Jimma town. The Denver II-Jimma evolved as a culturally relevant tool, ready to use for children from birth to six years in the multicultural and multilingual communities in the Jimma Zone, south west of Ethiopia. In the adaptation process, 36 items of the 125 in the Denver II test were modified. No test item was dropped, and this would guarantee to maintain the objectives and content validity of the original tool. Content validation was conducted by going through each test item at different meetings by the multidisciplinary research team with knowledge of local and western cultures. First, the objective of testing each Denver II test item, specific skill or competence assessed was discussed. Then the equivalence of the adapted version of the test item with the original one was examined in line with the objective, skill or competence assessed. This process was meant to maintain both content and construct validity.

Adaptation was predominantly in personal social test items. Only one gross motor item was adapted. This is consistent with other studies [19, 34]. Personal social skills seem to be more prone to socio-cultural influences than gross motor skills.

Feasibility and reliability of all test items were ensured during the adaptation process through piloting and fine-tuning. Good inter-rater and test-retest reliabilities were demonstrated during testing at schools by kindergarten teachers, and, at home by clinical nurses indicating that the Denver II-Jimma is reliable to use at different settings by different professionals. A strong intra-class correlation across all the domains also shows good overall reliability. Similar to the Denver II [16], inter-rater reliability seems to be better than the test-retest reliability.



**Table 3** The Denver II-Jimma with its age norms (in months) for 25, 50, 75 and 90 % of children passing the test items within the different domains

Item code	Item label	25 %	50 %	75 %	90 %	Item code	Item label	25 %	50 %	75 %	90 %
Personal social domain						Fine motor domain					
PS1	Regard face	birth <sup>b</sup>	birth <sup>b,d</sup>	birth <sup>b,d</sup>	birth <sup>b,d</sup>	FM1	Follow to midline	birth <sup>b</sup>	0.1 <sup>f</sup>	0.1 <sup>e</sup>	0.2 <sup>e</sup>
PS2	Smile responsively	0.8	1.1 <sup>f</sup>	1.3 <sup>d</sup>	1.6 <sup>d</sup>	FM2	Follow past midline	1.3	1.5 <sup>f</sup>	2.2 <sup>e</sup>	2.7 <sup>d</sup>
PS3	Smile spontaneously	1.1	1.4 <sup>e</sup>	1.6 <sup>f</sup>	1.9 <sup>e</sup>	FM3	Grasp rattle	2.4	3.0 <sup>d</sup>	3.6 <sup>d</sup>	4.1 <sup>d</sup>
PS4	Regard own hand	1.6	1.8 <sup>e</sup>	3.2 <sup>e</sup>	4.1 <sup>d</sup>	FM4	Hands together	2.5	3.1 <sup>f</sup>	3.7 <sup>f</sup>	4.3 <sup>d</sup>
PS5	Work for toy	3.1	4.5 <sup>d</sup>	6.0 <sup>f</sup>	7.6 <sup>f</sup>	FM5	Follow 180°	3.1	3.7 <sup>f</sup>	4.3 <sup>f</sup>	4.9 <sup>d</sup>
PS6	<b>Feed self<sup>a</sup></b>	4.0	5.9 <sup>d</sup>	8.2 <sup>f</sup>	10.5 <sup>f</sup>	FM6	<b>Regard coffee bean<sup>a</sup></b>	4.1	4.5 <sup>f</sup>	5.0 <sup>f</sup>	5.4 <sup>d</sup>
PS7	<b>Play clapping game<sup>a</sup></b>	6.9	8.1 <sup>e</sup>	9.4 <sup>d</sup>	10.6 <sup>d</sup>	FM7	Reaches	4.2	4.7 <sup>d</sup>	5.2 <sup>d</sup>	5.6 <sup>d</sup>
PS8	Indicate wants	6.1	8.1 <sup>e</sup>	10.1 <sup>d</sup>	12.2 <sup>d</sup>	FM8	Look for yarn	4.9	5.4 <sup>d</sup>	6.0 <sup>d</sup>	6.6 <sup>d</sup>
PS9	<b>Wave bye-bye/Say good-bye<sup>a</sup></b>	8.2	10.0 <sup>f</sup>	11.8 <sup>f</sup>	13.6 <sup>d</sup>	FM9	<b>Rake coffee bean<sup>a</sup></b>	5.3	5.9 <sup>d</sup>	6.4 <sup>d</sup>	6.9 <sup>d</sup>
PS10	Play ball with examiner	10.4	12.2 <sup>f</sup>	14.0 <sup>f</sup>	15.8 <sup>d</sup>	FM10	Pass cube	6.2	7.4 <sup>f</sup>	8.5 <sup>f</sup>	9.7 <sup>f</sup>
PS11	<b>Imitate activities<sup>a</sup></b>	9.7	11.6 <sup>d</sup>	13.4 <sup>d</sup>	15.2 <sup>d</sup>	FM11	Take 2 cubes	5.8	6.9 <sup>f</sup>	8.2 <sup>f</sup>	9.5 <sup>d</sup>
PS12	<b>Drink from cup or glass<sup>a</sup></b>	9.7	11.8 <sup>d</sup>	13.9 <sup>d</sup>	16.0 <sup>d</sup>	FM12	<b>Thumb-finger grasp<sup>a</sup></b>	6.5	7.5 <sup>d</sup>	8.7 <sup>d</sup>	9.9
PS13	Help in house	12.9	15.9 <sup>f</sup>	18.4 <sup>f</sup>	20.9 <sup>f</sup>	FM13	Bang 2 cubes held in hands	7.4	9.3 <sup>f</sup>	11.3 <sup>f</sup>	13.3 <sup>f</sup>
PS14	<b>Eats using spoon/fork/fingers<sup>a</sup></b>	11.7	14.5 <sup>d</sup>	17.2 <sup>d</sup>	20.0	FM14	Put block in cup	8.7	10.2 <sup>d</sup>	11.7 <sup>d</sup>	13.1 <sup>d</sup>
PS15	<b>Remove garment<sup>a</sup></b>	15.5	20.4 <sup>f</sup>	24.5 <sup>f</sup>	28.6 <sup>f</sup>	FM15	<b>Scribbles<sup>a</sup></b>	10.7	13.4 <sup>d</sup>	16.0 <sup>d</sup>	18.6 <sup>f</sup>
PS16	<b>Feed doll<sup>a</sup></b>	14.4	19.5 <sup>f</sup>	24.5 <sup>f</sup>	29.7 <sup>f</sup>	FM16	<b>Dump coffee bean, demonstrated<sup>a</sup></b>	10.9	13.5 <sup>d</sup>	16.0 <sup>d</sup>	18.5 <sup>d</sup>
PS17	<b>Put on clothing<sup>a</sup></b>	25.3	31.8 <sup>f</sup>	38.1 <sup>f</sup>	44.5 <sup>f</sup>	FM17	Tower of 2 cubes	12.6	15.6 <sup>d</sup>	18.3 <sup>d</sup>	21.0 <sup>d</sup>
PS18	<b>Wash mouth with help<sup>a</sup></b>	20.4	24.8 <sup>f</sup>	29.1 <sup>d</sup>	33.4 <sup>d</sup>	FM18	Tower of 4 cubes	16.4	19.3 <sup>d</sup>	22.2 <sup>d</sup>	25.1 <sup>d</sup>
PS19	<b>Wash and dry hands<sup>a</sup></b>	22.9	27.4 <sup>f</sup>	32.4 <sup>f</sup>	37.4 <sup>d</sup>	FM19	Tower of 6 cubes	19.7	23.1 <sup>d</sup>	26.6 <sup>d</sup>	30.1 <sup>d</sup>
PS20	<b>Name playmate<sup>a</sup></b>	22.8	28.3 <sup>d</sup>	33.4 <sup>d</sup>	38.6 <sup>d</sup>	FM20	Imitate vertical line	25.9	32.4 <sup>f</sup>	37.7 <sup>f</sup>	43.0 <sup>f</sup>
PS21	<b>Put on t-shirt<sup>a</sup></b>	34.2	40.8 <sup>f</sup>	47.3 <sup>f</sup>	53.8 <sup>f</sup>	FM21	Tower of 8 cubes	22.1	29.4 <sup>f</sup>	35.0 <sup>d</sup>	40.6 <sup>d</sup>
PS22	<b>Dress, no help<sup>a</sup></b>	43.9	50.5 <sup>f</sup>	57.0 <sup>f</sup>	63.5 <sup>f</sup>	FM22	Thumb wiggle	27.0	32.8 <sup>d</sup>	38.4 <sup>d</sup>	44.2 <sup>d</sup>
PS23	<b>Play social games<sup>a</sup></b>	31.8	41.6 <sup>f</sup>	51.2 <sup>d</sup>	60.9 <sup>d</sup>	FM23	<b>Copy<sup>a</sup> O</b>	38.7	43.1 <sup>d</sup>	47.3 <sup>d</sup>	51.6 <sup>d</sup>
PS24	<b>Clean face, no help<sup>a</sup></b>	37.9	46.8 <sup>f</sup>	55.5 <sup>d</sup>	64.4 <sup>d</sup>	FM24	<b>Draw person—3 parts<sup>a</sup></b>	44.0	48.2 <sup>d</sup>	52.3 <sup>d</sup>	56.4 <sup>d</sup>
PS25	<b>Serve oneself injera<sup>a</sup></b>	31.7	46.5 <sup>d</sup>	61.2 <sup>f</sup>	76.0 <sup>d</sup>	FM25	Copy <sup>a</sup> +	33.8	40.4 <sup>d</sup>	46.9 <sup>d</sup>	53.4 <sup>d</sup>
PS26	Toilet-going <sup>c</sup>	60.1	70.4 <sup>d</sup>	80.6 <sup>d</sup>	90.8 <sup>d</sup>	FM26	Pick longer line	33.8	40.4 <sup>d</sup>	46.9 <sup>d</sup>	53.4 <sup>e</sup>
						FM27	Copy □ demonstrated	45.7	51.8 <sup>d</sup>	57.7 <sup>d</sup>	63.6 <sup>d</sup>
						FM28	Draw person—6 parts <sup>a</sup>	52.4	57.5 <sup>d</sup>	62.4 <sup>d</sup>	67.4 <sup>d</sup>
						FM29	Copy <sup>a</sup> □	52.3	58.2 <sup>d</sup>	64.0 <sup>d</sup>	69.8 <sup>d</sup>
Language domain						Gross motor domain					
LA1	Respond to bell	birth <sup>b</sup>	birth <sup>b,d</sup>	birth <sup>b,d</sup>	birth <sup>b,d</sup>	GM1	Equal movement	birth <sup>b</sup>	birth <sup>b,d</sup>	birth <sup>b,d</sup>	birth <sup>b,d</sup>
LA2	Vocalizes	birth <sup>b</sup>	birth <sup>b,d</sup>	0.3 <sup>f</sup>	1.0 <sup>f</sup>	GM2	Lift head	birth <sup>b</sup>	0.1 <sup>f</sup>	0.4 <sup>f</sup>	0.7 <sup>f</sup>
LA3	"Ooo"/ Aah	1.4	1.5 <sup>f</sup>	1.6 <sup>d</sup>	1.6 <sup>e</sup>	GM3	Head up 45°	1.6	2.4 <sup>f</sup>	2.3 <sup>f</sup>	2.6 <sup>d</sup>
LA4	Laugh	2.0	2.2 <sup>f</sup>	2.5 <sup>d</sup>	2.7 <sup>e</sup>	GM4	Head up 90°	2.9	3.2 <sup>f</sup>	3.6 <sup>f</sup>	3.9 <sup>d</sup>
LA5	Squeals	2.2	2.5 <sup>f</sup>	2.8 <sup>d</sup>	3.2 <sup>e</sup>	GM5	Sit head steady	3.0	3.3 <sup>f</sup>	3.6 <sup>f</sup>	3.9 <sup>d</sup>
LA6	Turn to rattling sound	3.7	4.3 <sup>f</sup>	4.8 <sup>d</sup>	5.4 <sup>d</sup>	GM6	Bear weight on legs	3.1	3.4 <sup>f</sup>	3.6 <sup>d</sup>	3.8 <sup>e</sup>
LA7	Turn to voice	4.5	5.2 <sup>f</sup>	5.8 <sup>d</sup>	6.5 <sup>d</sup>	GM7	Chest up-arm support	3.9	4.3 <sup>f</sup>	4.6 <sup>f</sup>	4.9 <sup>d</sup>
LA8	Single syllables	4.6	5.5 <sup>d</sup>	6.3 <sup>d</sup>	7.1 <sup>d</sup>	GM8	Roll over	3.8	4.4 <sup>f</sup>	4.9 <sup>f</sup>	5.4 <sup>d</sup>
LA9	Imitate speech sounds	5.2	6.5 <sup>f</sup>	7.8 <sup>f</sup>	9.1 <sup>d</sup>	GM9	Pull to sit, no head lag	4.4	4.9 <sup>f</sup>	5.5 <sup>f</sup>	6.0 <sup>d</sup>
LA10	Dada/Baba/Mama, non-specific	5.7	6.7 <sup>d</sup>	7.8 <sup>d</sup>	8.8 <sup>d</sup>	GM10	Sit, no support	5.4	6.0 <sup>d</sup>	6.6 <sup>d</sup>	7.3 <sup>d</sup>

**Table 3** The Denver II-Jimma with its age norms (in months) for 25, 50, 75 and 90 % of children passing the test items within the different domains (*Continued*)

LA11	Combine syllables	6.2	7.5 <sup>f</sup>	8.8 <sup>f</sup>	10.1 <sup>d</sup>	GM11	Stand, holding on	6,5	7.3 <sup>d</sup>	8.3 <sup>d</sup>	9.3 <sup>d</sup>
LA12	Jabbers	7.0	8.5 <sup>f</sup>	10.0 <sup>f</sup>	11.5 <sup>d</sup>	GM12	Pull to stand	7,2	8.0 <sup>d</sup>	8.9 <sup>d</sup>	9.9 <sup>d</sup>
LA13	Dada/Mama/baba, specific	8.8	10.2 <sup>f</sup>	11.6 <sup>d</sup>	12.9 <sup>d</sup>	GM13	Get to sitting	7,4	8.3 <sup>d</sup>	9.2 <sup>d</sup>	10.1 <sup>d</sup>
LA14	One word	10.1	11.9 <sup>d</sup>	13.6 <sup>d</sup>	15.3 <sup>d</sup>	GM14	Stand 2 s	8,6	9.7 <sup>d</sup>	10.7 <sup>d</sup>	11.6 <sup>d</sup>
LA15	2 words	11.5	13.8 <sup>f</sup>	15.9 <sup>d</sup>	18.0 <sup>d</sup>	GM15	Stand alone	9,8	11.3 <sup>d</sup>	12.3 <sup>d</sup>	13.4 <sup>d</sup>
LA16	3 words	13,5	15.6 <sup>f</sup>	17.7 <sup>f</sup>	19.8 <sup>d</sup>	GM16	Stoop and recover	11,6	13.2 <sup>d</sup>	14.8 <sup>f</sup>	16.4 <sup>f</sup>
LA17	6 words	16,5	19.1 <sup>f</sup>	21.8 <sup>f</sup>	24.4 <sup>f</sup>	GM17	Walk well	11,3	13.1 <sup>d</sup>	14.9 <sup>f</sup>	16.7 <sup>f</sup>
LA18	Point 2 pictures	20,8	24.0 <sup>f</sup>	27.2 <sup>f</sup>	30.3 <sup>f</sup>	GM18	Walk backwards	11,9	14.9 <sup>d</sup>	17.4 <sup>f</sup>	19.9 <sup>f</sup>
LA19	Combine words	18,7	21.3 <sup>d</sup>	23.9 <sup>d</sup>	26.4 <sup>d</sup>	GM19	Runs	14,4	16.6 <sup>d</sup>	18.9 <sup>d</sup>	21.2 <sup>d</sup>
LA20	Name 1 picture	20,1	23.2 <sup>f</sup>	26.2 <sup>f</sup>	29.3 <sup>d</sup>	GM20	<b>Walk up steps<sup>a</sup></b>	13,9	16.9 <sup>d</sup>	19.6 <sup>d</sup>	22.3 <sup>d</sup>
LA21	<b>Body parts 6<sup>a</sup></b>	20,3	23.0 <sup>f</sup>	25.6 <sup>f</sup>	28.2 <sup>d</sup>	GM21	Kick ball forward	14,2	17.6 <sup>d</sup>	21.0 <sup>d</sup>	24.4 <sup>d</sup>
LA22	Point 4 pictures	25,9	31.0 <sup>f</sup>	35.9 <sup>f</sup>	40.9 <sup>f</sup>	GM22	Jump up	24.0	27.0 <sup>f</sup>	31.0 <sup>f</sup>	35.0 <sup>f</sup>
LA23	Speech, half understandable	20,5	24.2 <sup>f</sup>	27.8 <sup>d</sup>	31.5 <sup>e</sup>	GM23	Throw ball overhand	16.9	22.1 <sup>d</sup>	27.2 <sup>f</sup>	32.4 <sup>d</sup>
LA24	Name 4 pictures	27,6	32.5 <sup>f</sup>	37.3 <sup>f</sup>	42.1 <sup>f</sup>	GM24	Broad jump	31.7	35.6 <sup>d</sup>	39.3 <sup>f</sup>	43.1 <sup>f</sup>
LA25	Know 2 actions	24,5	29.5 <sup>d</sup>	34.4 <sup>d</sup>	39.4 <sup>d</sup>	GM25	Balance each foot 1 s	23.4	28.9 <sup>d</sup>	33.0 <sup>d</sup>	37.1 <sup>d</sup>
LA26	Know 2 adjectives	30,9	35.1 <sup>d</sup>	39.1 <sup>d</sup>	43.3 <sup>d</sup>	GM26	Balance each foot 2 s	23.9	31.2 <sup>e</sup>	36.4 <sup>e</sup>	41.6 <sup>e</sup>
LA27	<b>Name 1 color<sup>a</sup></b>	40,3	45.1 <sup>d</sup>	49.8 <sup>f</sup>	54.6 <sup>f</sup>	GM27	Hops	31.9	38.5 <sup>d</sup>	45.0 <sup>d</sup>	51.5 <sup>d</sup>
LA28	<b>Use of 2 objects<sup>a</sup></b>	30,4	35.6 <sup>f</sup>	40.7 <sup>d</sup>	45.8 <sup>d</sup>	GM28	Balance each foot 3 s	29.8	36.1 <sup>d</sup>	42.2 <sup>e</sup>	48.4 <sup>e</sup>
LA29	Count 1 block	36,6	41.9 <sup>d</sup>	47.0 <sup>f</sup>	52.2 <sup>f</sup>	GM29	Balance each foot 4 s	33.8	40.0 <sup>e</sup>	46.0 <sup>e</sup>	52.1 <sup>e</sup>
LA30	<b>Use of 3 objects<sup>a</sup></b>	32,3	37.5 <sup>d</sup>	42.7 <sup>d</sup>	47.8 <sup>d</sup>	GM30	Balance each foot 5 s	39.2	44.8 <sup>e</sup>	50.2 <sup>e</sup>	55.8 <sup>e</sup>
LA31	Know 4 actions	32,4	38.5 <sup>e</sup>	44.4 <sup>f</sup>	50.4 <sup>d</sup>	GM31	Heel-to-toe walk	49.1	55.5 <sup>d</sup>	61.8 <sup>d</sup>	68.1 <sup>d</sup>
LA32	Speech all understandable	28,4	36.3 <sup>f</sup>	44.0 <sup>f</sup>	51.8 <sup>d</sup>	GM32	Balance each foot 6 s	41.8	47.9 <sup>e</sup>	53.8 <sup>e</sup>	59.7 <sup>e</sup>
LA33	Understand 4 prepositions	30,8	37.5 <sup>d</sup>	44.0 <sup>d</sup>	50.7 <sup>e</sup>						
LA34	<b>Name 4 colors<sup>a</sup></b>	51,9	56.8 <sup>f</sup>	61.5 <sup>f</sup>	66.3 <sup>f</sup>						
LA35	<b>Define 5 words<sup>a</sup></b>	47,4	54.3 <sup>f</sup>	61.0 <sup>d</sup>	67.8 <sup>d</sup>						
LA36	Know 3 adjectives	34,5	40.6 <sup>d</sup>	46.5 <sup>d</sup>	52.5 <sup>e</sup>						
LA37	<b>Count 5 blocks<sup>a</sup></b>	49,0	54.3 <sup>d</sup>	59.4 <sup>d</sup>	64.6 <sup>d</sup>						
LA38	Opposites-2	48,5	54.0 <sup>d</sup>	59.2 <sup>d</sup>	64.5 <sup>d</sup>						
LA39	<b>Define 7 words<sup>a</sup></b>	59,1	66.8 <sup>f</sup>	74.2 <sup>f</sup>	81.8 <sup>f</sup>						

PS personal social, FM fine motor-adaptive, LA language, GM gross motor items

<sup>a</sup>Adapted test items (written in bold)

<sup>b</sup>The child is able to perform or pass the task soon after birth

<sup>c</sup>Newly added test item removed at final consensus meeting

<sup>d</sup>Item achieved at no significantly different ages it is achieved in Denver II (achieved at similar age)

<sup>e</sup>Item achieved at a significantly earlier age than the age it is achieved in Denver II (achieved at earlier age)

<sup>f</sup>Item achieved at a significantly later age than the age it is achieved in Denver II (achieved at later age)

Milestones attainment on Denver II and Denver II-Jimma were compared on 90 percentile ages. Though there is no significant difference on majority (66.4 %) of the test items, a clinically significant difference was observed on 42 items. Such a difference was also reported in earlier studies [17–19, 34–39]. The difference was found for both the culture specific and the cross-cultural items. This finding of achieving milestones at different ages seems to justify the need to have separate normative standards for valid interpretation of test results from different socio-cultural contexts.

Differences are observed in the number of Denver II test items adapted in different settings. While 36 test items are adapted in the present study, only two items (personal social item “play-pat-a-cake’ and language item Baba or Mama, nonspecific’) were modified while standardizing and adapting Denver test to Tbilisi [40] children in Georgia. Only five test items (4 personal social and one language) were modified while adapting and standardizing Denver II on Sri Lankan children [19]. In Singapore, 77 Denver II items (67 %) were shared with the adapted and standardized Singapore version [17]. Such

**Table 4** Reliability of Denver II-Jimma at item level indicated by inter-rater<sup>a</sup> and test re- test<sup>b</sup> kappa values, and at domain-level indicated by inter-rater and test-retest intraclass correlation coefficients

Reliability Measures	PS (26 items)	FM (29 items)	LA (39 items)	GM (32 items)	Total (126 items)
<b>Inter-rater</b> (kappa values)					
Excellent (0.81–1.00)	21 (80.8 %)	24 (82.8 %)	34 (87.2 %)	27 (84.4 %)	124 (98.4 %)
	4 (15.4 %) <sup>***</sup>	4(13.8 %) <sup>***</sup>	5(12.8 %) <sup>***</sup>	5(15.6 %) <sup>***</sup>	
Substantial (0.61–0.80)	1(3.8 %)	1 (3.4 %)	-	-	2 (1.6 %)
Acceptable (0.41–0.60)	-	-	-	-	-
Poor (<0.41)	-	-	-	-	-
<b>Inter-rater</b> (ICC) <sup>c</sup> , [95 % CI]	0.983, [0.979-0.986]	0.982, [0.978-0.985]	0.951, [0.940-0.959]	0.967, [0.961-0.973]	
<b>Test retest</b> (kappa values)					
Excellent (0.81–1.00)	15 (57.7 %)	4(13.8 %)	14 (35.9 %)	14 (43.8 %)	69 (54.8 %)
	4 (15.4 %) <sup>***</sup>	5 (17.24 %) <sup>***</sup>	6 (15.4 %) <sup>***</sup>	5 (15.6 %) <sup>***</sup>	
		1(3.45) <sup>*</sup>	1(2.6 %) <sup>**</sup>		
Substantial (0.61–0.80)	5 (19.2 %)	13 (44.8 %)	16 (41 %)	11 (34.4 %)	45 (35.7 %)
Acceptable (0.41–0.60)	2 (7.7 %)	5 (17.24 %)	2 (5.1 %)	2 (6.3 %)	11 (8.7 %)
Poor (<0.41)		1 (3.45 %)	-	-	1 (0.8 %)
Test-retest (ICC) <sup>d</sup> , [95 % CI]	0.802, [0.721–0.859]	0.831, [0.736–0.888]	0.840, [0.773–0.887]	0.793, [0.711–0.852]	

PS Personal social, FM Fine motor-adaptive, LA Language, GM Gross motor, ICC intraclass correlation coefficient, CI confidence interval

<sup>\*\*\*</sup>Kappa value not calculated but percentage of agreement is 100

<sup>\*\*</sup>Kappa value not calculated but percentage of agreement is 93.3

<sup>\*</sup>Kappa value not calculated but percentage of agreement is 90.91

<sup>a</sup>Agreement between two measurements done independently at a time

<sup>b</sup>Agreement between measurements repeated at a different time

<sup>c</sup>One-way random effect model is used and shows very high correlation

<sup>d</sup>Two-way random effect model is used and shows high correlation

findings seem to show that the number of test items needing adaptation varies in different socio-cultural contexts.

There are also differences in ages of attaining milestones in different settings. With a difference of more than 10 % on 90 percentile passing age, the Singapore differed on more than 30 items (20.1 %); the Denver–Tbilisi on 25 items (24 %), the Denver II-Jimma on 42 items (33.6 %) with the original Denver II. A comparison of the Sri Lankan norm with the Singapore and the Denver II norms also showed a difference of more than one month in ages of attaining milestones in more than 75 % of items in all domains [19]. The differences in ages of attaining milestones in the present study produced findings that are expected and consistent to earlier studies.

Taking in account that the Denver II-Jimma should be an ‘ideal reference’ to detect children at developmental risk, and monitor the general recovery of the child during rehabilitation, much care was spent on the standardization. Standardization therefore was done on a large sample of healthy children by excluding those with obvious disabilities and at risk during pre and perinatal stages of development. Children from comparatively very low-income families were not included for fear that such children are at higher developmental risks related to malnutrition and developmentally non-stimulating home environment. Moreover, significantly

malnourished children were also excluded from the analysis since malnutrition affects development.

An important aspect of the adaptation process is the involvement of an interdisciplinary team comprising academicians and practitioners from both the western and the local cultures. They were found instrumental in understanding both contexts while making relevant adaptations. Such a team composition was either not reported or considered in other similar studies.

The study is not also without limitations. First, though the Denver II is valid and is still in use in the western world, it was standardized 24 years ago. This standard is, however, still in use. Therefore, this study compared the data from two different time points. Second, though it is claimed that adaptation improves sensitivity [40], the Denver II-Jimma could still be a subject of limitation of the Denver II: weak specificity [41]. With adaptation of the traditional scoring and interpretation, however, the Denver II is regarded as more suitable for children with medically complex conditions [42], and a valid tool, particularly in assessing the language and fine motor skills of children with neurodevelopment risks [43].

## Conclusion

This study demonstrated how a Western tool can be effectively adapted to a non-Western setting. With high

inter-rater and test retest reliability, the Denver II-Jimma quickly assesses development of under six children, and is easy to use by first-line health workers and kindergarten teachers at home, school or health centers. Difference in milestones achievement ages on the adapted tool and on its originating Western tool shows that creating a local standard using the adapted tool is necessary for a valid interpretation of results. The study was conducted on children of diverse cultural, linguistic and ethnic communities. Hence, the result could be generalized to many other populations of Ethiopian children. However, some minor modifications may be needed in certain contexts which significantly differ from the present study setting. Future research has to examine if the tool can be used in other similar settings.

#### Abbreviations

FM, gross motor; GM, gross motor; LA, language; MUAC, Mid-upper circumference; MUACZ, Mid-upper arm circumference z score; PS, personal social; WAZ, weight-for-age z score.

#### Acknowledgements

We thank Dr. Tsinuel Girma, a pediatrician at Jimma University Specialized Referral Teaching Hospital, Ethiopia and Dr. Piet Vandebriel, child psychiatrist at PXL, Belgium for their contributions in team meetings; Nick Wirix, Ms in physiotherapy, Inne Snijers, Katrien Bielen, Ms in pediatric rehabilitation and Martijn Orye, Ms in rehabilitation and physiotherapy, Belgium for collecting data and writing reports during pilot studies; all caregivers, and kindergarten teachers who participated in the study. Our gratitude also goes to Dr. Gillian Lancaster for sharing ideas on creation of the Denver II-Jimma test chart. Our heartfelt appreciation is extended to scholars who developed the Denver II test. Our special gratitude is to Jimma University and VLIR /Maamse Interuniversitaire Raad (Flemish Interuniversity Council) partnership program for funding the research project as part of a PhD scholarship.

#### Funding

This study is part of the JU-IUC (Jimma University-Institutional University Co-operation) partnership program funded by VLIR-UOS (Maamse Interuniversitaire Raad). The funder had no role in study design, data collection and analysis, decision to prepare or publish the manuscript.

#### Authors' contributions

TG has conducted the statistical analysis, interpreted the data, and drafted the manuscript. MW took part in design of the study and supervision of data collection. JV and JL contributed to conception and design of the study, supervision of data collection and qualitative data analysis during the different phases of the study. HT and WK contributed to quantitative data analysis and critically reviewing the manuscript. PK contributed to conception and design of the study, interpretation of the data and critically reviewing the manuscript. MG contributed to conception and design of the study, interpretation of data and drafting of the manuscript. All authors read and approved the final manuscript.

#### Competing interests

The authors declare that they have no competing interests.

#### Ethics approval and consent to participate

Ethical approval was obtained from Jimma University, Ethiopia (RPGC/217/2010 dated 03/02/2010), and Hasselt University, Belgium (CME 2010/306). Written and oral consents of parents were obtained and children were always tested in the presence of caregivers.

#### Financial disclosure

The authors have no financial relationships relevant to this article to disclose.

#### Author details

<sup>1</sup>Department of Psychology, College of Behavioral Sciences and Education, Jimma University, Jimma, Oromia, Ethiopia. <sup>2</sup>Department of Population and Family Health, Jimma University, Jimma, Oromia, Ethiopia. <sup>3</sup>Department of Healthcare, PXL University College, Hasselt, Belgium. <sup>4</sup>Biocartis NV, Mechelen, Belgium. <sup>5</sup>Department of Epidemiology & Biostatistics, Jimma University, Jimma, Oromia, Ethiopia. <sup>6</sup>Department of Food Safety and Food Quality, University of Ghent, Ghent, Belgium. <sup>7</sup>REVAL Rehabilitation Research Center, Biomedical Research Institute, Faculty of Medicine and Life Sciences, University of Hasselt, Hasselt, Belgium.

Received: 1 February 2016 Accepted: 9 July 2016

Published online: 28 July 2016

#### References

1. Grantham-McGregor S, Cheung YB, Cueto S, Glewwe P, Richter L, Strupp B, Group tI. Developmental potential in the first 5 years for children in developing countries. *Lancet*. 2007;6(369(9555)):60–70.
2. Boivin MJ, Green SD, Davies AG, Giordani B, Mokili JKL, Cutting W. A preliminary evaluation of the cognitive and motor effects of pediatric HIV infection in Zairian children. *Health Psychol*. 1995;14(1):13–21.
3. Drotar D, Olness K, Wiznitzer M, Schatschneider C, Marum L, Guay L, Fagan J, Hom D, Svilar G, Ndugwa C, et al. Neurodevelopmental outcomes of Ugandan infants with HIV infection: an application of growth curve analysis. *Health Psychol*. 1999;18(2):114–21.
4. Holding PA, Taylor HG, Kazungu SD, Mkala T, Gona J, Mwamuye B, Mbonani L, Stevenson J. Assessing cognitive outcomes in a rural African population: development of a neuropsychological battery in Kilifi District, Kenya. *J Int Neuropsychol Soc*. 2004;10(2):246–60.
5. Nampijja M, Apule B, Lule S, Akurut H, Muhangi L, Elliott AM, Alcock KJ. Adaptation of Western measures of cognition for assessing 5-year-old semi-urban Ugandan children. *Br J Educ Psychol*. 2010;80(1):15–30.
6. Servili C, Medhin G, Hanlon C, Tomlinson M, Worku B, Baheretibeb Y, Dewey M, Alem A, Prince M. Maternal common mental disorders and infant development in Ethiopia: the P-MaMiE Birth Cohort. *BMC Public Health*. 2010;10:693.
7. Aboud FE, Alemu T. Nutrition, maternal responsiveness and mental development of Ethiopian children. *Soc Sci Med*. 1995;41(5):725–32.
8. Bhargava A. Modeling the effects of maternal nutritional status and socioeconomic variables on the anthropometric and psychological indicators of Kenyan infants from age 0–6 months. *Am J Phys Anthropol*. 2000;111(1):89–104.
9. Drewett R, Wolke D, Asefa M, Kaba M, Tessema F. Malnutrition and mental development: is there a sensitive period? A nested case–control study. *J Child Psychol Psychiatry*. 2001;42(2):181–7.
10. Neumann C, McDonald MA, Sigman M, Bwibo N, Marquardt M. Relationships between morbidity and development in mildly to moderately malnourished Kenyan toddlers. *Pediatrics*. 1991;88(5):934–42.
11. Stoltzfus RJ, Kvalsvig JD, Chwaya HM, Montresor A, Albonico M, Tielsch JM, Savioli L, Pollitt E. Effects of iron supplementation and anthelmintic treatment on motor and language development of preschool children in Zanzibar: double blind, placebo controlled study. *BMJ*. 2001;323(7326):1389–93.
12. Hokjindee U, Chongsuvivatwong V, Lim A, Pruphetkaew N. Denver Developmental Screening Test (DDST) survey and degree of malnutrition among children born to HIV infected mothers under the Prevention of Mother to-Child-Transmission (PMTCT) Program. *J Med Assoc Thai*. 2010; 93(12):1458–62.
13. Lohausa A, Kellerb H, Lammb B, Teuberta M, Fassbendera I, Freitagc C, Goertzd C, Graf d F, Kolling d T, Spanglerc S, et al. Infant development in two cultural contexts: Cameroonian Nso farmer and German middle-class infants. *J Reprod Infant Psychol*. 2011;29(2):148–61.
14. Nanthamongkolchai S, Ngaosusit C, Munsawaengsub C. Influence of parenting styles on development of children aged three to six years old. *J Med Assoc Thai*. 2007;90:971–6.
15. Frankenburg WK, Dodds JB. The Denver developmental screening test. *J Pediatr*. 1967;71(2):181–91.
16. Frankenburg WK, Dodds J, Archer P, Shapiro H, Bresnick B. The Denver II: a major revision and restandardization of the Denver developmental screening test. *Pediatrics*. 1992;89(1):91–7.
17. Lim HC, Chan T, Yoong T. Standardization and adaptation of the Denver Developmental Screening Test (DDST) and Denver II for use in Singapore children. *Singap Med J*. 1994;35:156–60.

18. Chikvinidze G, Geladze N, Natriashvili G, Tsigroshvili Z. Standardization and adaptation of the Denver Developmental Screening Test (DDST) for use in Tbilisi children. In: Tbilisi State Medical University Annals of Biomedical Research and Education. 2003. p. 3.
19. Wijedasa DN. Developmental screening in context: adaptation and standardization of the Denver Developmental Screening Test-II (DDST-II) for Sri Lankan children. *Child Care Health Dev.* 2012;38(6):889–99.
20. Frankenburg WK. Developmental surveillance and screening of infants and young children. *Pediatrics.* 2002;109(1):144–5.
21. Gladstone MLG, Umar E, Nyirenda M, Kayira E, et al. The Malawi Developmental Assessment Tool (MDAT): the creation, validation, and reliability of a tool to assess child development in rural African settings. *PLoS Med.* 2010;7(5), e1000273.
22. Central Statistical Agency-Ethiopia. The population projection based on the results of the May 2007 National Population and Housing Census. 2013.
23. WHO. ICF-CY: International Classification of Functioning, Disability and Healthy, Children & Youth Version. Geneva: World Health Organization; 2007.
24. Burton AW, Miller DE. Movement Skills Assessment. Champaign, IL: Human Kinetics; 1988.
25. Berry JW, Poortinga YH, Segall MH, Dasen PR. *Cross-Cultural Psychology: Research and Applications.* Secondthth ed. Cambridge: University Press; 2002.
26. Md O, Monteiro C, Akre J, Clugston G. The worldwide magnitude of protein-energy malnutrition: an overview from the WHO global database on child growth. *Bull World Health Organ.* 1993;71(6):703–12.
27. WHO Working Group. Use and interpretation of anthropometric indicators of nutritional status. *Bull World Health Organ.* 1986;64:924–41. World Health Organization.
28. World Medical Association. World medical association declaration of Helsinki: ethical principles for medical research involving human subjects. *Bull World Health Organ.* 2001;79(4):373–4.
29. Frankenburg WK, Dobbs JB, Archer P, Shapiro H, Bresnick B. Denver II Technical Manual. In: Denver Developmental Materials. USA: Denver, CO; 1992.
30. WHO. WHO Anthro Software (version 3.2.2, January 2011) and macros. Geneva, Switzerland: WHO Anthro; 2011.
31. Harrell FE. Regression modeling strategies with applications to linear models, logistic regression and survival analysis. 2012.
32. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1997;33(1):159–74.
33. Vameghi RHN, Sajedi F, Kazemnejad A. Production of a native developmental screening test: the Iranian experience. *Child Care Health Dev.* 2009;36(3):340–5.
34. Gladstone M, Lancaster G, Jones AP, Maleta K, Mtitimila E, Ashorn P, Smyth RL. Can Western developmental screening tools be modified for use in a rural Malawian setting? *Arch Dis Child.* 2008;93(1):23–9.
35. Ueda R. Standardization of the Denver Developmental Screening Test on Tokyo children. *Dev Med Child Neurol.* 1978;20:647.
36. Bryant GM, Davies KJ, Newcombe RG. Standardisation of the Denver developmental screening test for Cardiff children. *Dev Med Child Neurol.* 1979;21(3):353–64.
37. Shapira H, Harel S. Standardization of the Denver developmental screening test for Israeli children. *Isr J Med Sci.* 1983;19(3):246–51.
38. Alnaquib NM, Frankenburg WK, Mirza H, Yazdi AW, Al-Noori S. The standardization of the Denver developmental screening test on Arab children from the Middle East and North Africa. *J Med Liban.* 1999;47:95–106.
39. Solomons HC. Standardization of the Denver Developmental Screening Test on infants from Yucatan, Mexico. *Int J Rehabil Res.* 1982;5:179.
40. Chikvinidze G, Geladze N, Natriashvili G, Tsigroshvili Z. Standardization and adaptation of the Denver Developmental Screening Test (DDST) for use in Tbilisi children. *Ann Biomed Res Educ.* 2003;3(3):216–8.
41. Glascoe F, Byrne K, Ashford L, Johnson K, Chang B, Strickland B. Accuracy of the Denver-II in developmental screening. *Pediatrics.* 1992;89(6):1221–5.
42. Ware C, Sloss CF, Chugh CS, Budd KS. Adaptations of the Denver II scoring system to assess the developmental status of children with medically complex conditions. *Children's Health Care.* 2002;31(4):255–72.
43. Schatz J, McClellan CB, Puffer ES, Johnson K, Roberts CW. Neurodevelopmental screening in toddlers and early preschoolers with sickle cell disease. *J Child Neurol.* 2008;23(1):44–50.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

