

# Can Western developmental screening tools be modified for use in a rural Malawian setting?

M J Gladstone,<sup>1</sup> G A Lancaster,<sup>2</sup> A P Jones,<sup>3</sup> K Maleta,<sup>4</sup> E Mtitimila,<sup>5</sup> P Ashorn,<sup>5</sup> R L Smyth<sup>6</sup>

► Additional figures and a table are published online only at <http://adc.bmj.com/content/vol93/issue1>

<sup>1</sup> Department of Paediatrics, College of Medicine, Blantyre, Malawi; <sup>2</sup> Postgraduate Statistics Centre, Department of Mathematics and Statistics, Lancaster University, Lancaster, UK; <sup>3</sup> Centre for Medical Statistics and Health Evaluation, University of Liverpool, Liverpool, UK; <sup>4</sup> Department of Community Health, College of Medicine, Blantyre, Malawi; <sup>5</sup> Department of International Health, University of Tampere Medical School, Finland and Department of Paediatrics, Tampere University Hospital, Finland; <sup>6</sup> Institute of Child and Reproductive Health, University of Liverpool, Liverpool, UK

Correspondence to: Dr Melissa Gladstone, Institute of Child Health, University of Liverpool, Royal Liverpool Children's Hospital, Eaton Rd, Liverpool L12 2AP, UK; [mgladstone@btinternet.com](mailto:mgladstone@btinternet.com)

Accepted 4 March 2007  
Published Online First  
22 March 2007

## ABSTRACT

**Objective:** To create a more culturally relevant developmental assessment tool for use in children in rural Africa.

**Design:** Through focus groups, piloting work and validation, a more culturally appropriate developmental tool, based on the style of the Denver II, was created. Age standardised norms were estimated using 1130 normal children aged 0–6 years from a rural setting in Malawi. The performance of each item in the tool was examined through goodness of fit on logistic regression, reliability and interpretability at a consensus meeting. The instrument was revised with removal of items performing poorly.

**Results:** An assessment tool with 138 items was created. Face, content and respondent validity was demonstrated. At the consensus meeting, 97% (33/34) of gross motor items were retained in comparison to 51% (18/35) of social items, and 86% (69/80) of items from the Denver II or Denver Developmental Screening Test (DDST) were retained in comparison to 69% (32/46) of the newly created items, many of these having poor reliability and goodness of fit. Gender had an effect on 23% (8/35) of the social items, which were removed. Items not attained by 6 years came entirely from the Denver II fine motor section (4/34). Overall, 110 of the 138 items (80%) were retained in the revised instrument with some items needing further modification.

**Conclusions:** When creating developmental tools for a rural African setting, many items from Western tools can be adapted. The gross motor domain is more culturally adaptable, whereas social development is difficult to adapt and is culturally specific.

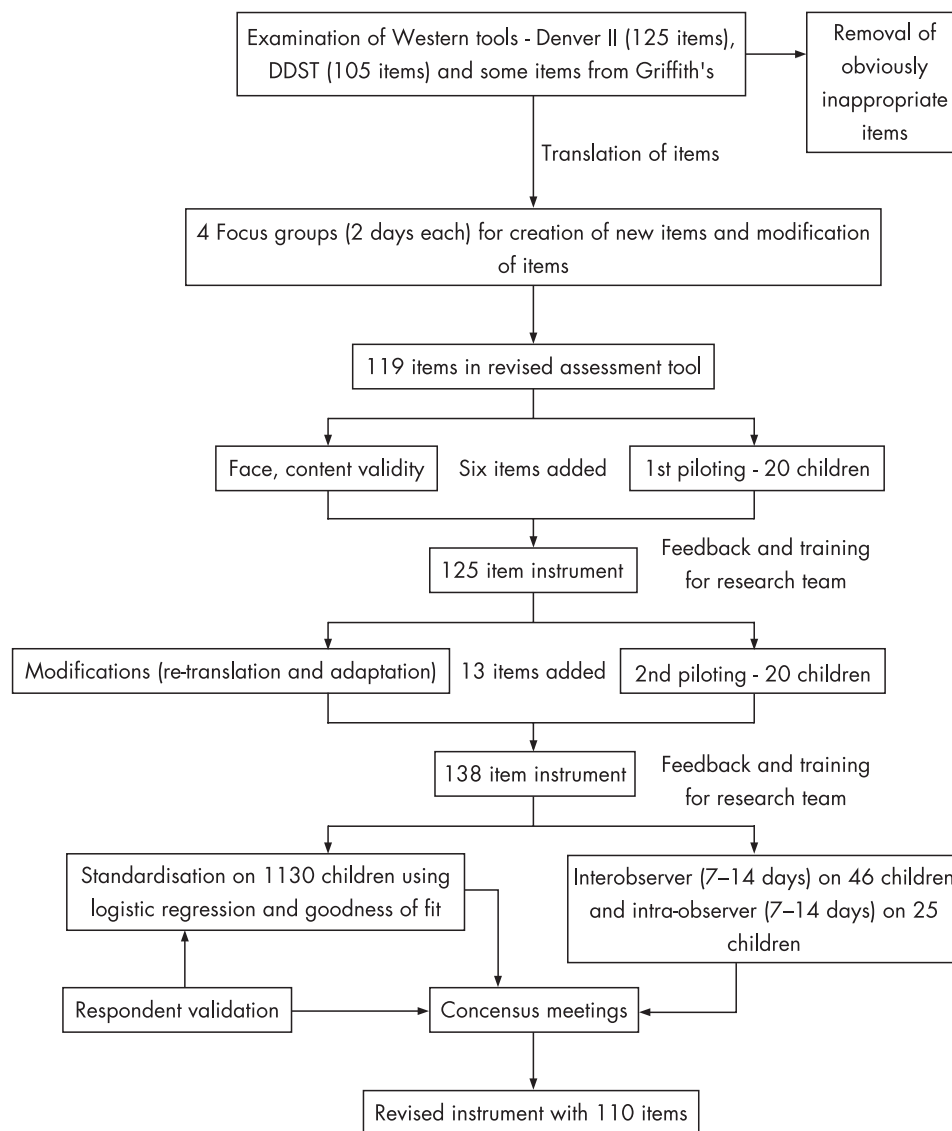
Eighty percent of the world's disabled population live in low income countries, many of these in Africa.<sup>1</sup> The World Health Organization has made early identification of children with disabilities a high priority, especially as early rehabilitation may reduce the impact of impairments.<sup>2–3</sup> To identify these children and provide basic services, developmental milestones need to be clearly identified. Furthermore, clinical studies investigating interventions in children require normal parameters.

When child development is assessed in clinical studies in developing countries, Western developmental tools are often utilised.<sup>4–5</sup> These include the Bayley scales,<sup>6</sup> the Griffith's,<sup>7</sup> the McCarthy scales<sup>8</sup> and the Denver II,<sup>9</sup> all designed and validated in Western countries. These tools may be tailored for use in non-Western settings. Often translation (changing of the language used) is all that is carried out.<sup>10–11</sup> If this is not accompanied by a process of adaptation, translation alone may not allow completely for local expressions and customs,

therefore leading to misinterpretation of results.<sup>12</sup> In other settings, tools are adapted and items are modified and in some cases new items are created for use within a Western tool.<sup>13</sup> Sometimes these tools are piloted (tried out before use)<sup>14</sup> and validated (assessed that they are measuring what they are supposed to be measuring) in the local population.<sup>15</sup> Even these adapted tools, however, are of limited value without normal ranges for their defined population. Standardisation studies (finding norms for a population) have taken place in many non-Western countries mainly using the Denver Developmental Screening Test (DDST) in a translated and occasionally adapted form,<sup>16–18</sup> but none of these studies was in Africa. Only two studies have attempted standardisation in Africa, one using a translated form of the Bayley scales with an urban black South African population<sup>19</sup> and the other on a limited age range in a rural Nigerian population.<sup>20</sup>

It is clear that Western developmental assessment tools may include tasks and materials which are completely alien to other cultures. These tools may therefore fail to identify and assess children adequately in cultural settings other than those for which they were created.<sup>21</sup> This may be less of a problem when comparing groups of children, but when Western tools are used alone as an outcome measure, culture may have an effect. In theoretical studies, culture has been demonstrated to have an influence on child development, particularly in the area of social development.<sup>22–24</sup> Cognitive abilities such as memory, categorisation techniques and pattern recognition have also been reported to be influenced by culture.<sup>25–28</sup> Even gross motor development may possibly be affected by culture.<sup>29–31</sup>

In this study, we aimed to create a simple, culturally appropriate developmental assessment tool adapted and modified from Western tools and standardised for use in rural Malawi. The first stage in the development of this tool was to identify which items from Western tools (eg, the DDST or Denver II) were not relevant to the age-appropriate experiences of rural Malawian children. These items were then replaced with ones more appropriate to this cultural context. We did this firstly by holding focus groups to agree which items should be replaced and to create alternative items. All items (both retained and new) were then validated and standardised in a large population study. The performance of all items was examined in a consensus meeting and a revised instrument proposed.



**Figure 1** The process of creating a more culturally appropriate developmental assessment tool.

## METHODS

### Setting and study population

This study was a substudy of the Lungwena Child Survival Study (LCSS), a prospective family cohort study looking at gestational health and the growth, development, morbidity and mortality of rural Malawian infants and children. Lungwena is an area in southern Malawi where a government health centre serves an approximately 100 km<sup>2</sup> rural area with some 17 000 people in 23 villages. Most of the inhabitants are Muslims of the Yao tribe. The literacy rate is low and subsistence farming and fishing are the main occupations. The original cohort for the LCSS was enrolled between June 1995 and August 1996. All pregnant women presenting for antenatal care were eligible for the LCSS and 97% of the population of pregnant women in the area, at that time, were enrolled in the study. Details of recruitment, collection of background data and follow-up have been described previously.<sup>32 33</sup>

The population of children used for this study is the original LCSS cohort of children aged 3.5–6.5 years and younger siblings aged 0–3.5 years. Out of the 1237 LCSS children and siblings available, 1197 were seen, with 40 families either refusing to take part or not being available. The ages of the children were

known from LCSS birth data or from the “health passport” given to mothers at the birth of their baby where the date of birth is recorded and which almost all mothers carry with them for all health appointments. A quota sampling strategy was used as in the DDST and Denver II<sup>34</sup> with target numbers of children being sought in each of 33 age groups (see supplementary table A). A total of 67 children were excluded due to premature birth (34 weeks or less measured by fundal height at the antenatal clinic),<sup>32</sup> twin birth or significant disability including severe malnutrition (weight for height z score of less than  $-2$ ), leaving 1130 children in the final analysis.

The LCSS received approval from the National Health Science Research Committee in Malawi (HSRC 93/94). Informed verbal consent was sought from each mother at the beginning of the LCSS and again before a development assessment was carried out.

### Creation of the developmental assessment tool

The Denver II, DDST and Griffith’s instruments were examined by the Malawian research team. Items considered to be culturally appropriate were included and translated, whereas those considered inappropriate (such as “prepares cereal” or “plays board/card games”) were removed. New items and modifications to

**Table 1** Examples of specific items added or removed during the process of creating new more culturally appropriate tool

Domain of development	DDST or Denver II items removed at stage 1	New items replaced at stage 1	Denver II/Griffith's items removed or needing modification after consensus meeting	New items removed or needing modification after consensus meeting
Gross motor	<ol style="list-style-type: none"> <li>1. Head up to 45°</li> <li>2. Walks up steps</li> <li>3. Pedals tricycle</li> </ol>	<ol style="list-style-type: none"> <li>1. Dances to music</li> <li>2. Kneels well to be respectful</li> <li>3. Climbs and gets down</li> <li>4. Carries object on head with one hand</li> <li>5. Pushes wheel with stick</li> <li>6. Carries water on head – no hands</li> </ol>	<ol style="list-style-type: none"> <li>1. Pulls self to stand (not understood by research assistants)</li> <li>2. Touches toes with legs straight (difficult to teach)</li> </ol>	<ol style="list-style-type: none"> <li>1. Dances (not specific enough)</li> <li>2. Carries object on head with one hand (need to specify exact object)</li> <li>3. Carries water on head with no hands (need to specify size)</li> </ol>
Fine motor	<ol style="list-style-type: none"> <li>1. Looks for yarn</li> <li>2. Thumb wiggle</li> <li>3. Thumb finger grasp</li> </ol>	<ol style="list-style-type: none"> <li>1. Bangs drum/pot</li> <li>2. Moulds ball with clay</li> <li>3. Makes doll of clay</li> <li>4. Sorts objects into groups</li> <li>5. Picks longest stick</li> </ol>	<ol style="list-style-type: none"> <li>1. Draws a person with 3 parts (many not passing by age 6)</li> <li>2. Draws a person with 6 parts (many not passing by age 6)</li> </ol>	<ol style="list-style-type: none"> <li>1. Moulds ball with clay (task not normally done by children)</li> <li>2. Bangs drum/pot (translation “plays drum” incorrect)</li> <li>3. Builds house of corn cobs (sex specific)</li> </ol>
Language	<ol style="list-style-type: none"> <li>1. Responds to bell</li> <li>2. Squeals</li> <li>3. Points to pictures/name pictures</li> <li>4. Counts 1 block</li> <li>5. Uses plurals</li> </ol>	<ol style="list-style-type: none"> <li>1. Sings</li> <li>2. Knows today and tomorrow</li> <li>3. Knows name of own village</li> <li>4. Uses respectful terms</li> <li>5. Knows father's name</li> <li>6. Can count bottle tops to 10</li> </ol>	<ol style="list-style-type: none"> <li>1. Knows 3 colours (colours used differently in Malawi)</li> <li>2. Knows opposites (not specific enough –question needed clarifying)</li> </ol>	<ol style="list-style-type: none"> <li>1. Uses respectful terms (poor reliability)</li> <li>2. Knows father's name (poor reliability)</li> <li>3. Can count bottle tops to 10 (counting learned at school – not many children in school/nursery aged under 5)</li> </ol>
Social	<ol style="list-style-type: none"> <li>1. Plays pat-a-cake</li> <li>2. Plays ball with examiner</li> <li>3. Resists toy pull</li> <li>4. Plays peek-a-boo</li> <li>6. Uses spoon/fork</li> <li>7. Feeds doll</li> <li>8. Brushes teeth with help</li> <li>9. Names a friend</li> <li>10. Plays board games</li> <li>12. Prepares cereal</li> <li>13. Plays interactive games, eg, tag</li> <li>14. Buttons up</li> </ol>	<ol style="list-style-type: none"> <li>1. Shares things</li> <li>2. Plays with friends</li> <li>3. Spends most of time on mum's back</li> <li>4. Eats with family in a group</li> <li>5. Can plant seeds by self</li> <li>6. Washes dishes</li> <li>7. Washes clothes</li> <li>8. Sweeps</li> <li>9. Goes to the toilet with help</li> <li>10. Goes to the toilet by self</li> <li>11. Makes porridge by self</li> <li>12. Adds wood to fire</li> <li>13. Can build fire by self</li> </ol>	<ol style="list-style-type: none"> <li>1. Helps in house</li> <li>2. Puts on clothing with help (poor reliability)</li> <li>3. Feeds self (specify what type of food)</li> <li>4. Drinks from cup by self (not specific enough regarding how well or not child needed to do this)</li> </ol>	<ol style="list-style-type: none"> <li>1. Spends most of time on mum's back (family dependent)</li> <li>2. Shy with strangers (poor reliability)</li> <li>3. Sweeps (sex specific)</li> <li>4. Washes dishes (sex specific)</li> <li>5. Goes to the toilet with help (not specific enough, eg, older children need help using pit latrines)</li> </ol>

Western test items were then created through discussions with a series of focus groups. Key informants were the eight local research workers. They were all women of child-bearing age with at least 8 year's education and research experience of at least 5 years. Themes relating to developmental milestones were discussed and ideas from these sessions were used to create new items. Illustrations were made for most items in the instrument and used as prompts for the research workers. Some came with permission from *Disabled village children*.<sup>35</sup>

Face validity<sup>36</sup> and content validity were assessed by all research assistants, five Malawian paediatricians, a language expert from the University of Malawi and six medical students at the College of Medicine, Malawi. Once the new instrument was created, the team was trained in its use and it was piloted in two stages. At each stage, feedback and training were given and problematic items were re-adapted or re-translated. The process of creating and refining the more culturally appropriate tool is shown in fig 1.

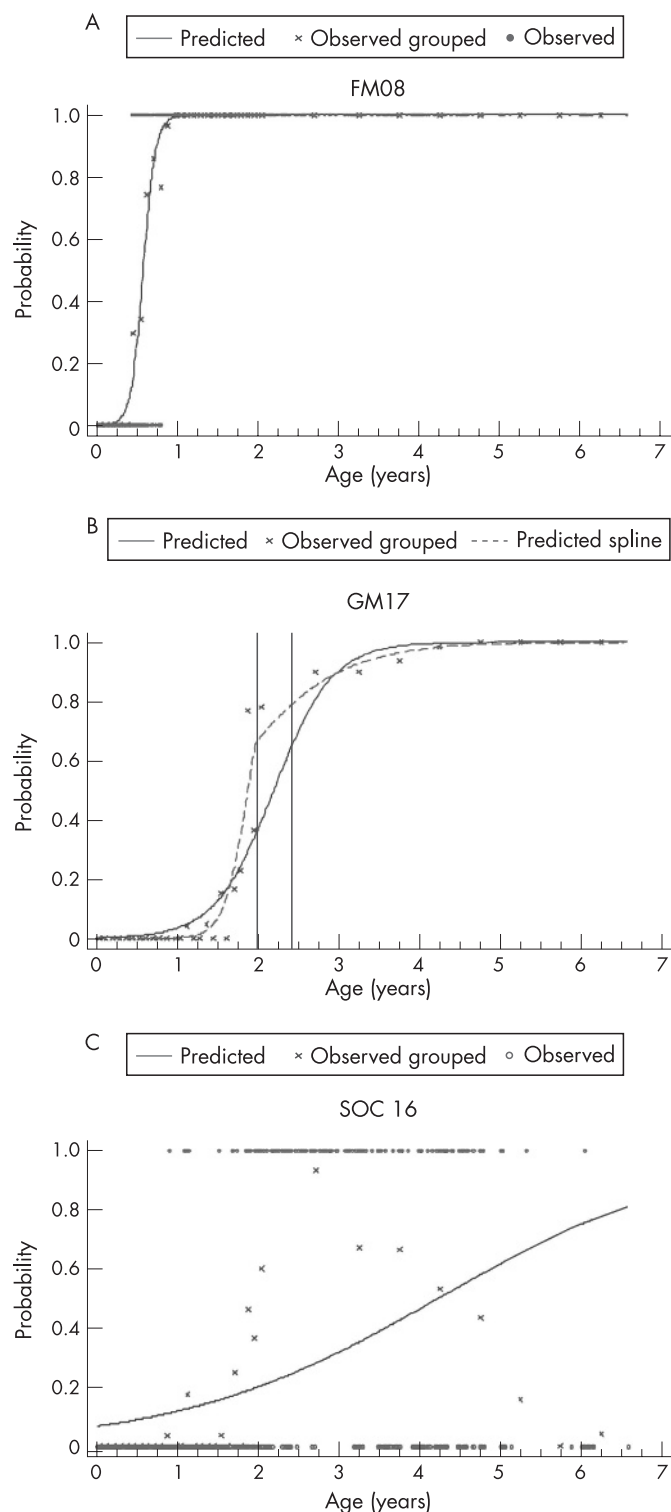
### Standardisation using a normal population sample

Overall, 1197 children were assessed on one occasion between February 2000 and April 2001 on a home visit by research

assistants. The assessment took approximately 35 min to complete and where possible, items were directly observed. In a few cases a report was given, for example “does he go to the toilet by himself?”. Items were scored as either pass or fail, or “don't know” if the child was uncooperative or unwell. Items were asked until the child failed seven items in a row.

Data entry and analysis were carried out using Microsoft Excel 6.0, SPSS 11.1, Stats-direct and STATA computer programs. Each child in the study was identified by a code. Data were checked prior to analysis and any outlying results were reviewed.

Standardisation is the process of determining normal age ranges for which children pass the items for a developmental assessment tool. A logistic regression analysis was carried out with decimal age and sex as explanatory variables. The observed and predicted probability of passing was determined and graphs were drawn for each item. The goodness of fit of the graph was visually assessed and discrepancies reviewed. To determine statistically whether or not the fitted curve was a sufficiently good representation of the data, a goodness of fit statistic was calculated.<sup>37</sup> If this was significant at the 5% level, indicating a poor fit, then the data were re-examined and refitting was done



**Figure 2** Examples of (A) a good logistic fit for fine motor question 8 “Transfers objects from hand to hand”, (B) spline fit for gross motor question 17 “Walks backwards”, and (C) a poorly worded question (social question (SOC) 16 “Can put clothes on with help”).

using triple split spline regression. The ages corresponding to the 35th and 65th percentiles were calculated from the original fit to determine the cut-points. For some items that performed less well, the cut-points were chosen by viewing the graphs to facilitate a good fit. Three logistic curves were then fitted, one for each region, based on the split.<sup>38 39</sup> Any items with

significant gender effects were removed or considered for further modification to ensure the tool was applicable to all children irrespective of gender. Using the predicted probabilities found from the logistic regression analyses, the ages corresponding to 25%, 50%, 75% and 90% of the children passing were determined for each item. These were then used to plot the age norms of achievement of each milestone in a box-type representation.

### Reliability of the items

Reliability for each item was tested by using two subsamples of 60 (inter-observer) and 28 (intra-observer) randomly selected children who were seen at 7 and 14 days after initial assessment. Of the 60 children, 46 completed the follow-up using two different examiners (inter-observer), while 25 of the 28 children used the same examiner (intra-observer). All items in the tool were assessed for both types of reliability. Kappa statistics ( $\kappa$ ) with 95% confidence intervals (CI) were used to calculate the degree of observer agreement for each question. Positive values of 0 to <0.2 indicate poor agreement, >0.2 to 0.4 fair agreement, >0.4 to 0.6 moderate agreement, >0.6 to 0.8 good agreement and >0.8 to 1 very good agreement.<sup>40</sup>

Respondent validation was carried out after the preliminary analysis. This method of validation involves the reporting of findings back to the participants. Findings were fed back at the end of the study to the Lungwena Health Centre Management Committee. This consisted of four chiefs, one overall representative and three women representatives, all from the local area.

### Consensus meeting

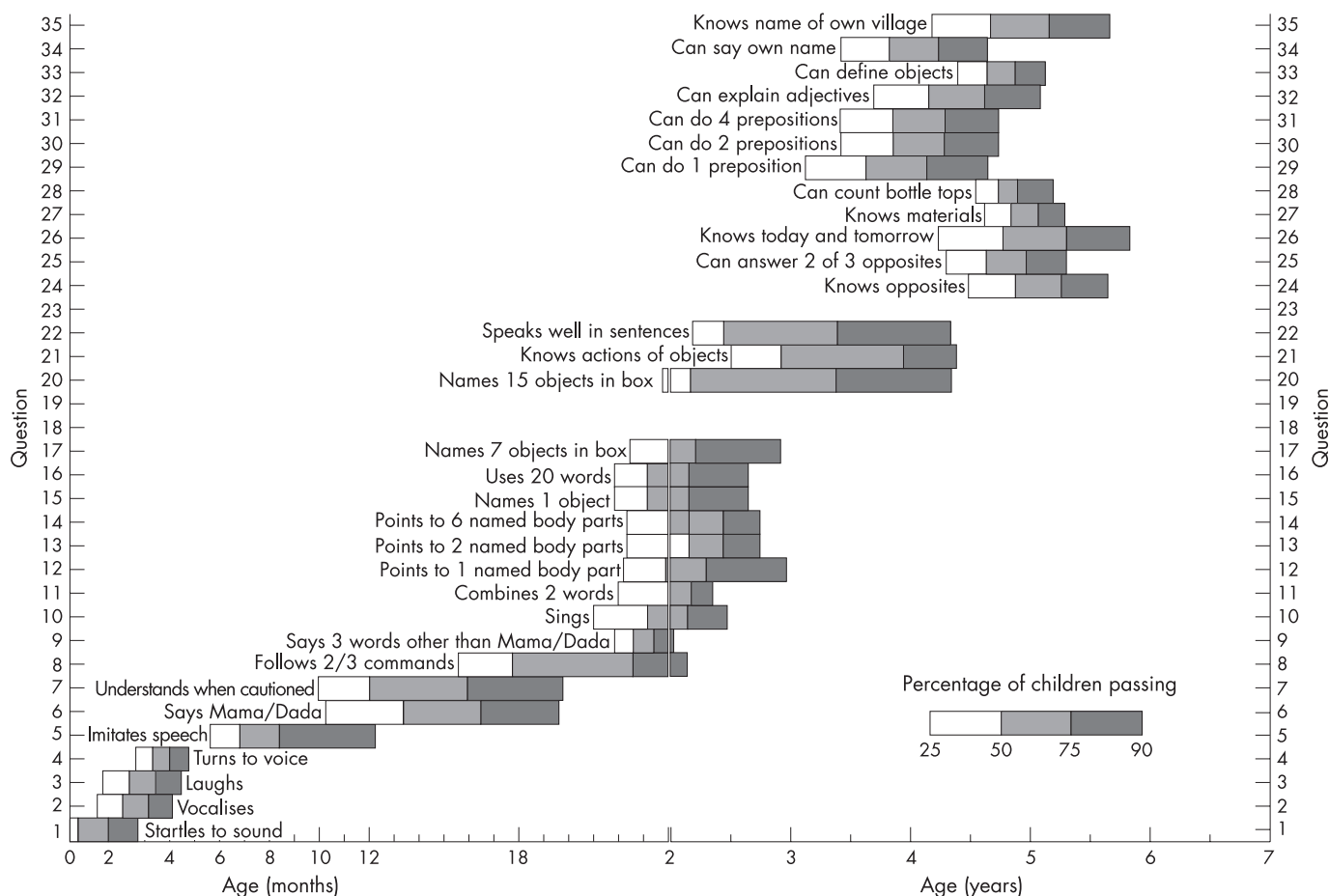
Once all the items were analysed, an expert panel (MG, AJ, EM and GL), which included a Malawian paediatrician, met to review the results and decide which items should remain, which should be modified and which should be removed. Items were judged on their graphical representation, and goodness of fit on logistic regression, reliability and subjective ratings of “interpretability” by participants and researchers.

### RESULTS

A tool with 138 items (34 gross motor (GM), 34 fine motor (FM), 35 language and 35 social items) was created. An example of the tool is shown in supplemental fig B (see supplementary data). Most (58%) items were from the DDST and Denver II, with a small percentage (9%) from the Griffith’s instrument. Many items in the GM (82%), language (77%) and FM (70%) sections were directly translated from Western tests with modifications mainly in the FM section. Only 37% of the social items were taken from Western tests. The first two columns of table 1 provide examples of many of the items that were removed from the DDST or Denver II, and shows the newly created items which replaced them.

The face validity and content validity of the tool were tested. The modified instrument appeared to those questioned to cover development in children in ways that were important, and it was judged to examine in a fully comprehensive and logical fashion the domains of child development for children in Malawi. It was therefore considered to have good face validity and content validity. Most items were found to be acceptable for studying children’s development in this setting through respondent validity. The pictures as prompts were found to be particularly helpful to the researchers in the field.

Examples of graphs created through logistic regression during the standardisation procedure and where triple split joined



**Figure 3** Example of developmental milestones achieved by Lungwena children in the area of language development. Age ranges are given for percentage of children passing an item.

regression was used, are shown in fig 2. In terms of goodness of fit on logistic regression and on spline regression, social items had the highest number of poor fits (51%, 18/35), sex being an independent predictor in some of these (23%, 8/35) (see table 2). A larger proportion of the newly created items had a poor fit on logistic regression (15%, 7/46) and an effect of sex (17%, 8/46) than those from the Western tools. The few items not attained by 6 years came from the fine motor area of development and included “draws a man with 6 parts” and “draws a square”. The results of the Lungwena milestones for the language section of development are shown in fig 3. The other areas of development are described in supplemental fig C (see supplementary data).

Reliability results are also shown in table 2.

For inter-observer reliability, 82% (113/138) of the questions had moderate to very good reliability ( $\kappa > 0.4$ ). There are no figures in the Denver technical manual for inter-observer reliability for comparison. Intra-observer reliability demonstrated moderate to very good reliability ( $\kappa > 0.4$ ) for 75% (106/138) of the questions. This compares well with Denver II figures,<sup>34</sup> where 81% of their items had a  $\kappa > 0.4$ . In relation to the domains of development, GM items had the best overall inter-observer (29/34 items) and intra-observer (32/34 items) reliability with  $\kappa > 0.4$ . Items from the social area performed less well, with only 74% (26/35) of the items on inter-observer

**Table 2** Decisions regarding suitability of items within each domain of development and the source of the question

	Domain of development				Source of question			
	GM	FM	Lang	Social	Denver or DDST	Griffith's	Newly derived	Total
No. of items in total	34	34	35	35	80	12	46	138
Inter-observer items with $\kappa > 0.4$	29	28	30	26	72	8	33	113
Intra-observer items with $\kappa > 0.4$	32	29	24	21	68	9	29	106
Poor fit on logistic regression	1	0	3	9	6	0	7	13
Sex effect on logistic regression	0	3	0	8	3	0	8	11
Not attained by 6 years	0	4	0	0	4	0	0	4
Poor interpretability or translation	1	2	9	9	7	4	10	21
To be included or modified for revised version of the tool	33	27	32	18	69	9	32	110

DDST, Denver Developmental Screening Test; FM, fine motor; GM, gross motor; Lang, language.

reliability and 60% (21/35) of the intra-observer items having a  $\kappa > 0.4$ . In relation to the source of the item, more of the locally-derived items had poor inter-observer (33/46) and intra-observer (15/46) reliability ( $\kappa < 0.4$ ) in comparison to those items derived from the Denver II (12/80 and 8/80).

After a consensus meeting, 110 of the 138 items (80%) were retained in the revised instrument, with some needing further modification. Only 69% (32/46) of the newly created items were retained in comparison to 86% (69/80) of the DDST or Denver items used (see table 2). The results of this meeting giving examples of items removed are detailed in the last two columns of table 1.

## DISCUSSION

We have demonstrated that many items from Western tools can work well when adapted and translated for other settings. They have already had their own rigorous reliability and validity studies carried out in the West and therefore are more likely to be robust in use. However, through our focus group, validation and piloting work, we have also demonstrated that in all domains of Western tests (such as the DDST), there are some items which are culturally inappropriate for a rural African population. For example, questions such as “prepares cereal” or “plays board games/card games” are uncommon activities for children in rural Africa. Also, the pink doll in the DDST kit was terrifying to most children when used in piloting; many children had never seen anything like it and many screamed. It would have been unlikely that we would have been able to get them to sit down and “feed the doll”. Some of the naming questions in the Language section of the DDST or Denver II have pictures of objects that children, at least in the part of rural Africa studied, have never seen before, such as a horse and a car. This makes it difficult for them to name them, especially as many children have also never seen a book at their age, or pictorial representations of many objects.

In the creation of new items, however, many newly created items were less reliable, more sex-specific and had poorer goodness of fit in logistic regression. This was most evident in the social domain and least evident for gross motor skills. Social skills seem to have the least “universality” and in measuring them, we need to question the appropriateness of the concepts being measured in such different settings. When measuring “social skills” we may be determining the ability of the child to have learned important skills instilled by parents and carers in particular cultural settings, but this can only be measured if pertinent skills are tested for. The difficulty when creating new social items for a tool such as this, is that the items must be specific enough to distinguish between the developmental age ranges of children, but also be clear and easy to explain in a developmental tool. This will continue to be a challenge.

It was not a primary aim of our study to compare our results with the Denver II or DDST. A formal statistical comparison has not been possible; however, when comparing our charts with those from the Denver II or DDST on gross comparison, it does seem that there are obvious differences in milestones with children from the West. For example, the item “combines two words” in the Denver II is attained at between 17 and 21 months, whereas in our sample this was obtained at between 21 months and 2 years 4 months. This demonstrates the importance and necessity of creating norms for a given African population, as they are likely to be different from those in the West.

A second phase of work is currently underway using the methodology that we have formulated in this first study to

## What is already known on this subject

- ▶ The ages of attainment of developmental milestones can differ with the cultural background of the child, although there are very few data from Africa.
- ▶ Developmental assessment in African settings is often carried out using translated and adapted Western tools, with a dearth of tools adapted, validated and standardised specifically for rural African children.

## What this study adds

- ▶ A procedure using validity, reliability, goodness of fit on logistic regression and expert consensus has been devised for assessing newly created questions or items adapted or taken from Western developmental assessment tools.
- ▶ Many items from Western developmental assessment tools can work in a rural African setting. Gross motor items are most reliable, whereas items within the social area of development need to account for cultural differences.

refine a further tool with a larger standardisation sample. This work will include creating a scoring system, and carrying out more detailed reliability measurements and further validity tests of between-group and construct validity. Once this new version has been created and has undergone the strict procedures that we have instituted in our methodology, we hope to have created a tool that may benefit community health workers in other rural settings in Africa after local validation. The complete tool may also be used by research workers who are investigating developmental outcomes as part of their intervention strategies.

**Acknowledgements:** Our thanks to all the staff and research assistants at the Lungwena Child Survival Study site in Lungwena, Malawi, to Professor R Broadhead, head of the Department of Paediatrics at the College of Medicine in Blantyre, Malawi, to Professor MB Duggan, professor in paediatrics and community health who worked at the College of Medicine in Malawi when the study was carried out for her help and suggestions, and Dr L Rosenbloom, recently retired consultant paediatric neurologist at the Royal Liverpool Children’s Hospital, for his help and suggestions. Our thanks also to Dr W Frankenberg and H Shapiro for comments as well as statistical advice when writing up the study.

**Funding:** Funding was supplied by the Academy of Finland, the Foundation for Paediatric Research in Finland, the Medical Research Fund of Tampere University Hospital, the Alexander Wernher Piggot Memorial Trust and the EWG Memorial Trust. There was no involvement from the study sponsors in the study design, collection, analysis, interpretation, writing of reports or decisions to submit.

**Competing interests:** None.

## REFERENCES

1. WHO. Disability, including prevention, management and rehabilitation. Report by the Secretariat, 2005. [http://www.who.int/gb/ebwha/pdf\\_files/WHA58/A58\\_17-en.pdf](http://www.who.int/gb/ebwha/pdf_files/WHA58/A58_17-en.pdf) (accessed 21 October 2007).
2. Committee on Children with Disabilities. Developmental surveillance and screening of infants and young children. *Pediatrics* 2001;108:192–5.
3. Durkin MS, Davidson LL, Desai P, et al. Validity of the ten-question screen for childhood disability: results from population based studies in Bangladesh, Jamaica and Pakistan. *Epidemiology* 2005;5:283–9.
4. Powell CA, Walker SP, Himes JH, et al. Relationships between physical growth, mental development and nutritional supplementation in stunted children: the Jamaica study. *Acta Paediatr* 1995;84:22–9.
5. Joos SK, Pollitt E, Mueller WH, et al. The Bacon Chow Study: maternal nutritional supplementation and infant behavioural development. *Child Dev* 1983;54:669.
6. Bayley N. *Bayley Scales of Infant Development manual*. 2nd ed. San Antonio: Harcourt Brace, 1993.
7. Griffiths R. *The abilities of young children*. Amersham: ARICD, 1984.

8. **McCarthy D.** *McCarthy Scales of Children's Abilities*. New York: Psychological Corporation, 1972.
9. **Frankenburg WK,** Dodds J, Archer P, *et al.* The Denver II: a major revision and restandardization of the Denver Developmental Screening Test. *Pediatrics* 1992;**89**(1):91–7.
10. **Boivin MJ,** Green SD, Davies AG, *et al.* A preliminary evaluation of the cognitive and motor effects of pediatric HIV infection in Zairian children. *Health Psychol* 1995;**14**(1):13–21.
11. **Boivin MJ,** Giordani B, Ndanga K, *et al.* Effects of treatment for intestinal parasites and malaria on the cognitive abilities of schoolchildren in Zaire, Africa. *Health Psychol* 1993;**12**(3):220–6.
12. **Geisinger K.** Cross-cultural normative assessment: translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychol Assess* 2006;**6**(4):304–12.
13. **Drotar D,** Olness K, Wiznitzer M, *et al.* Neurodevelopmental outcomes of Ugandan infants with HIV infection. An application of growth curve analysis. *Health Psychol* 1999;**18**(2):114–21.
14. **Stoltzfus RJ,** Kvalsvig JD, Chawaya HM, *et al.* Effects of iron supplementation and antihelminthic treatment on motor and language development of preschool children in Zanzibar: double blind, placebo controlled study. *BMJ* 2001;**323**:1389–93.
15. **Holding PA,** Taylor HG, Kazungu SD, *et al.* Assessing cognitive outcomes in a rural African population: development of a neuropsychological battery in Kilifi district, Kenya. *J Int Neuropsychol Soc* 2004;**10**(2):246–60.
16. **Al-Naquib N,** Frankenburg WK, Mirza H, *et al.* The standardization of the Denver Developmental Screening Test on Arab children from the Middle East and north Africa. *J Med Liban* 1999;**47**(2):95–106.
17. **Lim HC,** Chan T, Yoong T. Standardisation and adaptation of the Denver Developmental Screening Test (DDST) and Denver II for use in Singapore children. *Singapore Med J* 1994;**35**(2):156–60.
18. **Song J,** Zhu YM, Gu XY. Restandardization of Denver Developmental Screening Test for Shanghai children. *Chin Med J (Engl)* 1982;**95**(5):375–80.
19. **Richter LM,** Griesel R, Rose C. The Bayley Scales of Infant Development: a South African standardization. *S Afr J Occup Ther* 1992;**22**(1):14–25.
20. **Ogunnaik OA,** Houser RFJ. Yoruba toddlers' engagement in errands and cognitive performance on the Yoruba Mental Subscale. *Int J Behav Dev* 2002;**26**(2):145–53.
21. **Olade RA.** Evaluation of the Denver Developmental Screening Test as applied to African children. *Nurs Res* 1984;**33**(4):204–7.
22. **Levine RA,** Dixon S, Levine S, *et al.* *Child care and culture: lessons from Africa*. Cambridge: Cambridge University Press, 1994.
23. **Whiting B,** Whiting WJ. *Children of six cultures. A psycho-cultural analysis*. Harvard, MA: Harvard University Press, 1975.
24. **Kambalmetore S,** Hartley S, Lansdown R. An exploration of the Malawian perspective on children's everyday skills: implications for assessment. *Disabil Rehabil* 2000;**22**(17):802–7.
25. **Serpell R.** How specific are perceptual skills? A cross-cultural study of pattern reproduction. *Br J Psychol* 1979;**70**:365–80.
26. **Rogoff B,** Waddell KJ. Memory for information organized in a scene by children from two cultures. *Child Dev* 1982;**53**:1224–8.
27. **Curran HV.** *Nigerian children: developmental perspectives*. London: Routledge & Kegan Paul, 1984.
28. **Glick J.** Cognitive development in cross-cultural perspective. In: F Horowitz, ed. *Review of child development research*. Chicago: University of Chicago Press, 1975:595–654.
29. **Super CM.** Environmental effects on motor development: the case of "African infant precocity". *Dev Med Child Neurol* 1976;**18**:561–7.
30. **Warren N.** African infant precocity. *Psychol Bull* 1972;**78**:353.
31. **Geber M.** Developpement psycho-moteur de l'enfant africain. *Courier* 1956;**6**:17.
32. **Vaahtera M,** Kulmala T, Ndeka M, *et al.* Ante and perinatal predictors of infant mortality in rural Malawi. *Arch Dis Child Fetal Neonatal Ed* 2000;**82**:F200–4.
33. **Maleta K,** Virtanen S, Espo M, *et al.* Timing of growth faltering in rural Malawi. *Arch Dis Child* 2003;**88**:574–8.
34. **Frankenburg WK,** Dodds JB, Archer P, *et al.* *Denver II technical manual*. Denver, CO: Denver Developmental Materials, 1990.
35. **Werner D.** *Disabled village children - a guide for community health workers, rehabilitation workers, and families*. Berkeley, CA: The Hesperian Foundation, 1987.
36. **Eiser C,** Morse R. Quality of life measures in chronic diseases of childhood. *Health Technol Assess* 2001;**5**(4).
37. **Lemeshow S,** Hosmer D. A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol* 1982;**115**(1):92–106.
38. **Greenland S.** Dose-response and trend analysis in epidemiology: alternatives to categorical analysis. *Epidemiology* 1995;**6**(4):356–64.
39. **Pastor R,** Guallar E. Use of two-segmented logistic regression to estimate change-points in epidemiological studies. *Am J Epidemiol* 1998;**148**(7):631–42.
40. **Landis R,** Koch G. The measurement of observer agreement for categorical data. *Biometrics* 1977;**33**:159–74.

### Access a vast information database with Toll-Free linking

"Toll-free" linking gives you immediate access to the full text of many of the cited articles in a paper's reference list—FOR FREE. With the support of HighWire's vast journal catalogue, a huge reference library is now open to you. If HighWire hosts the journal, you can view the full text of the referenced article, completely free of charge by following the Free Full Text links in the references.