

Addressing measurement challenges in early childhood development research in low- and middle-income countries

The British Academy-led GCRF/ DFID Early Childhood Development (ECD) programme looked at how to measure ECD outcomes in low- and middle-income countries (LMICs). Tools for ECD measurement inherently reflect the cultural beliefs, values and practices of child-rearing that are prevalent in different social and cultural contexts and developing a widely applicable measure poses a number of challenges.ⁱⁱ As most commonly used measures are developed primarily in high-income countries, there is a significant risk of cultural imposition in expanding ECD research agendas in LMIC contexts if these are not implemented or adapted sensitively. There is a need to minimise the risks in adapting these measures in different contexts, including addressing:

- i. *construct bias* (where the instrument measures something different than its intended target);
- ii. *method bias* (where the test procedure exposes children to situations, stimuli and options that are not familiar to them); and
- iii. *item bias* (where items do not measure in the same ways across groups due to translational problems).ⁱⁱⁱ

Developing context-specific measures or, at least, carefully adapting existing ones is crucial for gaining a culturally informed understanding of ECD in LMICs.

Equally important is the *process^{iv}* of measurement as there is a general lack of technical capacity required for administering assessment tools in LMICs. The ECD programme inevitably confronted a number of measurement challenges, and the project teams deployed a range of strategies to address these. This note presents the key insights generated in relation to such challenges and responses.

Key Insights from the ECD Programme

Finding appropriate solutions to the measurement challenges in LMICs is an interdisciplinary enterprise. Several projects brought together practitioners and academic experts with diverse disciplinary and methodological backgrounds including education, psychology, anthropology,

public health and child health for adapting measures for physiological and behavioural change.

Research teams engaged with national experts and other stakeholders in deciding upon the appropriate assessment tools to be used and/or adapted, and conducted qualitative exploratory studies to make culturally-informed choices. In Kenya, for example, one study collected information about locally used terms and language related to play and child stimulation. As contexts are different, adapting measurement tools is critical to ensure that data collected reflect the children's development rather than differences in the child's environment compared to other contexts. One of the main factors to consider is the availability of normative standards for the particular setting. Normative standards are helpful in establishing whether or not a specific population is performing below, at, or above expected levels for that country or area. Normative standards also minimise the risk of reporting 'false positives', where children may be reported as falling below the expected standards because the tool is not normed for that particular context.

In Zimbabwe and South Africa, the research team observed the participants' surroundings and communal places including homes, ECD centres, clinics, and markets to understand the context of children's language and cognitive experiences for effectively adapting assessment tools. The research conducted in Zimbabwe was the first study to use the Griffiths III tool for assessment of a large number of individuals in a community setting. The Griffiths III was developed primarily for detailed one-on-one child assessment but the research team extended this to community level: population-wide data from the tool is rare and the results from such an exercise would be invaluable in terms of health benefits for the children and understanding child development at a community rather than individual level. In South Africa both the language and cognitive measures used (Hopkins Verbal Learning complemented by the BOT-2 test)^v were adapted for the Ndumo area and for pre-school children. The goals of adaptation were to ensure relevant constructs were tested fairly and accurately. The process of adaptation began with critical analysis of the test, vocabulary adjustment, piloting in a clinical context and then testing in the field. The research team found the Hopkins Verbal Learning test easy to adapt: the word lists are able to be developed specifically for the target language and age, it does not have visual instruments to alter and it allows repeated assessments for re-evaluation.

The adaptations the team made to the child development assessment tools (Griffiths III Child Development Assessment Tool and Hopkins Verbal Test^{vi}) constitute new knowledge. The adaptations in language and memory tests can now be applied to similar groups, fitting similar

profiles. Test adaptation is time consuming and needs to be based on large numbers, therefore accumulated information from each research project informs test adaptations and standardisation for future studies.

An important insight generated by the ECD programme was the value of ensuring culturally appropriate measurement within ECD agendas. Various concepts and their constructs could function as a trojan horse for structural inequalities at local levels. By bringing to the fore the contested nature of the notions of early childhood care and education (ECCE) in Bihar and Tamil Nadu in India, the study illustrated the tendency of ECCE measures, such as an idea of 'school readiness', to advance values and practices of dominant caste/class as the standard model.

It is important to recognise that the sub-components and dimensions of the existing tools are often large, each consisting of lists of corresponding items and indicators. Constraints in linguistic and cultural translation of lengthy tools, and the capacity and resources in LMICs to administer these with very young children, require adapting these measures carefully while keeping them *short* and *simple*. A study on scaling up ECD interventions in Rwanda, for example, adapted a questionnaire for measuring parental self-efficacy (the TOPSE questionnaire)^{vii} by working with the developer of the tool to ensure that it was suitable for rural contexts. It was also the first time that this tool had been translated into Kinyarwanda. Reducing the dimensions of the standard instrument made it logistically convenient with regards to travel, storage and processing, and appropriate for the Rwandan context. Yet such a strategy is not necessarily the most efficient in every scenario.

One of the challenges faced in administering measurement instruments was the lack of capacity among local teams. In order to address this issue, project teams provided training to local team members, caregivers and community members. In Kenya, contextualisation included fieldwork training where the project team and trainees engaged with mothers of children under the age of 3 to learn and test how assessment tests would be conducted in the local context. Materials for the assessments were procured from the local market to make the test tools locally appropriate. In some instances, the lack of certain measures made a case for developing new context-specific measures to cover dimensions that were not captured before or for which the adaptation did not prove to be effective. In Bangladesh, for example, the research team studying the intergenerational impact of poverty reduction programmes on ECD chose to use a non-verbal spatial reasoning task, a mathematics task and two executive functioning tasks, instead of translating and aligning a cognitive skills test. The team initially developed a large number of items using existing assessments which were reduced after piloting and adjusting for statistical validity and the range of children's ability. These

instruments were coded on the electronic tablets that field enumerators could easily use in remote villages. Developing such a measurement can prove to be an efficient, time-saving and cost-effective strategy in low-resource settings.

The study on the intergenerational ECD impact of educational scholarships in Ghana contributed to the development of pre-school social, motivational and cognitive measures that (a) could highlight skills critical to success in schooling, and (b) are cheaper, effective to administer and appropriate for contexts such as Ghana. The details are set out in Box 1.

Box 1: Adapting and developing game-based measures of cognitive development in Ghana

The study adapted game-based measures of child development that were initially designed by the Harvard Laboratory of Development Studies for its [study](#) of child cognitive development in India. Under the ECD Programme, these measures were adapted and piloted for six months to suit the cultural context of rural and urban Ghana. Led by an expert national surveyor, the field team in Ghana video-recorded piloting sessions, and frequently reported on the field performance of the measures to the primary investigators. The measures were implemented once the project team had confidence that they were sufficiently informative about child development in the Ghanaian context.

Several logistical challenges were faced in adapting and developing these game-based measures to the Ghanaian context: the distance to research sites was often long, assessment materials were heavy, and there was a high turn-over of the field team. Where possible, the project team used private transport, light-weight game materials, and digital screens (which were suitable for children from the age of five) instead of physical games. There were also challenges in administering the tests in rural areas as community members would gather and try to influence children's test performance. Field teams adapted various strategies to manage crowd and external interventions during the test. Teams found that 20-30 minutes of testing were optimal for infants with 45-60 minutes recommended for three- and five-year-old children. Longer testing time would prove too fatiguing to the children to provide reliable results. To ensure consistent and reliable results, children were tested on their birthdays.

Recommendations

The insights generated by the ECD Programme offer several implications for future research programmes.

- The time scale for adapting a measurement in a given cultural context can be lengthy and requires support, including building local capacity. Promoting ECD in LMICs must be

accompanied with long-term investments which avoid cultural impositions and compromising the validity and reliability of the measures.

- Researchers and practitioners should engage with the local context, national experts and stakeholders in the process of selecting, developing, adapting, testing/piloting and revising various measures of ECD.
- There is a need to recognise the risks in presenting, through the measures, certain groups' values and practices as *universal*, whether at community or global levels. Such assumptions can work to obscure the rich cultural knowledge of traditionally marginalised groups.
- The issue of translating the instruments, which are often developed in English, into local language(s) also requires strong linguistic and cultural familiarity so that the questions and cultural references make sense in a research context.
- Successful adaptation of measurement tools requires making choices around the various domains, dimensions and items to be included according to the purpose of the assessment. Instruments should be both concise and comprehensible in order to achieve the benchmarks of validity and reliability in cultural contexts different to that of their origin whilst also keeping costs of administration low.
- Developing interdisciplinary collaborations with experts in partner countries is critical to ensuring the success of measurement tools, whether newly developed or adapted from existing models.

i This note was written by Dr Arif Naveed, University of Bath. The British Academy is very grateful for his work. This note is based on British Academy-funded projects 'Economic conditions in early childhood and the intergenerational transmission of poverty in Bangladesh' (Professor Oriana Bandiera FBA), 'The impact of parental education on cognitive development in early childhood: evidence from the long-term tracking of a randomized evaluation in Ghana' (Professor Esther Duflo FBA), 'Testing the feasibility of incorporating support for early childhood development into the Baby Friendly Community Initiative in Kenya', (Professor Paula Griffiths), 'Examining the contexts, practices and costs of early childhood care and education in India: responsive models for child development' (Dr Jyotsna Jha, Dr Arathi Sriprakash), 'Scaling-up early child development interventions in Rwanda' (Professor Patricia Justino) and 'Providing the evidence base and tools for prioritising and implementing paediatric schistosomiasis control to enhance early childhood development' (Professor Francisca Mutapi). The projects are part of the British Academy-led DFID/GCRF programme on [Early Childhood Development](#), which investigated the dynamics of early childhood development in low- and middle-income countries.

ii See for details:

Fernald, L. C. H., Prado, E., Kariger, P. and Raikes, A. A toolkit for measuring early childhood development in low – and middle – income countries. The World Bank. 2017.

Greenfield, Patricia M., L. Monique Ward, and Jennifer Jacobs. You can't take it with you: Why ability assessments don't cross cultures." *American Psychologist* 52 (10). 1997, p. 1115–24.

Rosselli, Monica and Alfredo Ardila. The impact of culture and education on non-verbal neuropsychological measurements: A critical review.' *Brain and Cognition* 52 (3), 2003, pp. 326–33.

Cole, Michael. 'Culture-Free Versus Culture-Based Measures of Cognition.' In *The Nature of Cognition*, edited by Robert J. Sternberg. Cambridge: MIT Press. 1999, pp. 654–64.

iii See Fernald, L. C. H., Prado, E., Kariger, P. and Raikes, A. 2017, pp.68-69.

iv See for details:

Carter, Julie A., Janet A. Lees, Gladys M. Murira, Joseph Gona, Brian G. R. Neville, and Charles R. J. C. Newton. 2005. "Issues in the Development of Cross-Cultural Assessments of Speech and Language for Children." *International Journal of Language & Communication Disorders* 40 (4): 385–401.

Malda, Maike, Fons J. R. van de Vijver, Krishnamachari Srinivasan, Catherine Transler, Prathima Sukumar,

and Kirthi Rao. "Adapting a Cognitive Test for a Different Culture: An Illustration of Qualitative Procedures." *Psychology Science Quarterly* 50 (4). 2008, pp. 451–68.

^v The Bruininks-Oseretsky Test of Motor Proficiency, Second Edition (**BOT-2**) **Brief form** is an individually administered test that uses engaging, goal-directed activities to measure a wide array of motor skills in individuals aged 4 through 21. The Hopkins Verbal Learning Test consists of a 12-item word list, composed of four words from each of the three semantic categories. The subject is instructed to listen carefully as the examiner reads the word list and attempt to memorize the words.^{vi}

^{vii} The Tool to Measure Parenting Self-Efficacy (TOPSE) is a UK-developed instrument that assesses six parenting domains: emotion and affection, play and enjoyment, empathy and understanding, control, discipline and boundary setting, pressures of parenting self-acceptance and learning and knowledge.